



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ
ΧΩΡΟΤΑΞΙΑΣ ΠΟΛΕΟΔΟΜΙΑΣ & ΠΕΡΙΦΕΡΕΙΑΚΗΣ ΑΝΑΠΤΥΞΗΣ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Γεώργιου Κουζιώκα του Νικολάου

ΣΥΣΤΗΜΑ ΛΗΨΗΣ ΑΠΟΦΑΣΕΩΝ ΣΤΗ ΔΙΟΙΚΗΣΗ ΔΗΜΟΣΙΩΝ
ΚΑΙ ΔΙΚΑΣΤΙΚΩΝ ΥΠΗΡΕΣΙΩΝ ΜΕ ΧΡΗΣΗ
ΤΕΧΝΗΤΗΣ ΝΟΗΜΟΣΥΝΗΣ, GIS ΚΑΙ ΤΗΛΕΠΙΣΚΟΠΗΣΗΣ

Email: gekouzio@uth.gr

Βόλος, 2018

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ
ΧΩΡΟΤΑΞΙΑΣ ΠΟΛΕΟΔΟΜΙΑΣ
& ΠΕΡΙΦΕΡΕΙΑΚΗΣ ΑΝΑΠΤΥΞΗΣ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

**ΣΥΣΤΗΜΑ ΛΗΨΗΣ ΑΠΟΦΑΣΕΩΝ ΣΤΗ ΔΙΟΙΚΗΣΗ ΔΗΜΟΣΙΩΝ
ΚΑΙ ΔΙΚΑΣΤΙΚΩΝ ΥΠΗΡΕΣΙΩΝ ΜΕ ΧΡΗΣΗ ΤΕΧΝΗΤΗΣ
ΝΟΗΜΟΣΥΝΗΣ, GIS ΚΑΙ ΤΗΛΕΠΙΣΚΟΠΗΣΗΣ**

Γεώργιος Νικ. Κουζιώκας

Τριμελής Συμβουλευτική Επιτροπή:

- Κωνσταντίνος Περάκης, Καθηγητής, ΤΜΧΠΠΑ, Πανεπιστημίου Θεσσαλίας
(Επιβλέπων Καθηγητής)
- Μαρί Νοέλ Ντυκέν, Καθηγήτρια, ΤΜΧΠΠΑ, Πανεπιστημίου Θεσσαλίας
(Μέλος της Τριμελούς Συμβουλευτικής Επιτροπής)
- Αλέξανδρος Χατζηγεωργίου, Καθηγητής, Τμήμα Εφαρμοσμένης
Πληροφορικής, Πανεπιστημίου Μακεδονίας
(Μέλος της Τριμελούς Συμβουλευτικής Επιτροπής)

Βόλος, 2018

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

Γεώργιος Νικ. Κουζιώκας

Σύστημα Λήψης Αποφάσεων στη Διοίκηση Δημοσίων και Δικαστικών Υπηρεσιών με
Χρήση Τεχνητής Νοημοσύνης, GIS και Τηλεπισκόπησης

Διδακτορική Διατριβή

που υποβλήθηκε στο Τμήμα Μηχανικών Χωροταξίας,
Πολεοδομίας και Περιφερειακής Ανάπτυξης της Πολυτεχνικής Σχολής
του Πανεπιστημίου Θεσσαλίας
Ημερομηνία Προφορικής Εξέτασης: 17-10-2018

Επταμελής Εξεταστική Επιτροπή

- Ντυκέν Μαρί Νοέλ, Καθηγήτρια Πανεπιστημίου Θεσσαλίας
- Περάκης Κωνσταντίνος, Καθηγητής, Πανεπιστημίου Θεσσαλίας
- Χαλκιάς Χρίστος, Καθηγητής Χαροκόπειου Πανεπιστημίου
- Χατζηγεωργίου Αλέξανδρος, Καθηγητής Πανεπιστημίου Μακεδονίας
- Σταθάκης Δημήτριος, Αναπληρωτής Καθηγητής Πανεπιστημίου Θεσσαλίας
- Δασυγένης Μηνάς, Επίκουρος Καθηγητής Πανεπιστημίου Δυτικής Μακεδονίας
- Καλογήρου Σταμάτης, Επίκουρος Καθηγητής Χαροκόπειου Πανεπιστημίου

© ΓΕΩΡΓΙΟΣ Ν. ΚΟΥΖΙΩΚΑΣ

© ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

Διδακτορική Διατριβή

«ΣΥΣΤΗΜΑ ΛΗΨΗΣ ΑΠΟΦΑΣΕΩΝ ΣΤΗ ΔΙΟΙΚΗΣΗ ΔΗΜΟΣΙΩΝ ΚΑΙ ΔΙΚΑΣΤΙΚΩΝ ΥΠΗΡΕΣΙΩΝ ΜΕ ΧΡΗΣΗ ΤΕΧΝΗΤΗΣ ΝΟΗΜΟΣΥΝΗΣ, GIS ΚΑΙ ΤΗΛΕΠΙΣΚΟΠΗΣΗΣ»

Ο συγγραφέας της διδακτορικής διατριβής δηλώνει ότι η Διδακτορική Διατριβή εκπονήθηκε από τον ίδιο και αντιπροσωπεύει τις προσωπικές του απόψεις και αναφέρει πλήρως όλες τις πηγές στις οποίες ανέτρεξε για την εκπόνηση της Διδακτορικής Διατριβής.

Περίληψη

Η παρούσα διδακτορική διατριβή ερευνά την ανάπτυξη μεθοδολογίας για την εφαρμογή της τεχνητής νοημοσύνης και της πληροφορικής, σε συνδυασμό με την χρήση των γεωγραφικών συστημάτων πληροφοριών και της τηλεπισκόπησης, στη λήψη αποφάσεων στην διοίκηση δικαστικών και δημοσίων υπηρεσιών, όπως οι Αστυνομικές Αρχές, Εισαγγελία, Δικαστήρια, Εκπαιδευτικές Μονάδες, Δημοτικές Αρχές, Χωροταξικές και Κοινωνικο-οικονομικές υπηρεσίες, οργανισμοί περιβαλλοντικοί και διαχείρισης υδάτων και άλλες δημόσιες υπηρεσίες. Η ανάπτυξη της μεθοδολογίας βασίζεται στην εφαρμογή και σύγκριση μεθόδων τεχνητής νοημοσύνης, όπως: Τεχνητών Νευρωνικών Δικτύων, Μηχανών Διανυσμάτων Υποστήριξης, Τυχαίων Δασών Δέντρων Απόφασης, Νοημοσύνης Σμήνους Σωματιδίων για την πρόβλεψη παραγόντων σημαντικών στη διοίκηση των ανωτέρω υπηρεσιών, όπως της εγκληματικότητας, περιβαλλοντικών παραγόντων, όπως το όζον, κοινωνικο-οικονομικών παραγόντων όπως η ανεργία, αλλά και για την ταξινόμηση δορυφορικών εικόνων. Στην παρούσα διατριβή, προτείνεται μία νέα μεθοδολογία επονομαζόμενη ως προβλεπτική διοίκηση, βασισμένη στην Τεχνητή Νοημοσύνη. Επίσης, αναπτύχθηκε μία νέα μεθοδολογία για την χωρική πρόβλεψη παραγόντων όπως η εγκληματικότητα, βασισμένη σε τεχνικές τεχνητής νοημοσύνης και στα γεωγραφικά συστήματα πληροφοριών. Επιπλέον, αναπτύχθηκαν δύο νέες υβριδικές μέθοδοι ταξινόμησης δορυφορικών εικόνων στη δημόσια διοίκηση, βασισμένες σε έναν νέο προτεινόμενο υβριδικό πυρήνα Μηχανών Διανυσμάτων Υποστήριξης και σε ένα νέο υβριδικό μοντέλο Τεχνητών Νευρωνικών Δικτύων και Τυχαίων Δασών Δέντρων Απόφασης. Τα μοντέλα Τεχνητής Νοημοσύνης βελτιστοποιήθηκαν με αλγόριθμους αυτοματοποιημένης μηχανικής μάθησης AutoML (Automated Machine Learning), όπως η Βελτιστοποίηση Σμήνους Σωματιδίων και η βελτιστοποίηση Bayes. Ως βασικός στόχος της προτεινόμενης μεθοδολογίας στην παρούσα διατριβή, είναι η εφαρμογή κατάλληλων μεθόδων τεχνητής νοημοσύνης για την αποτελεσματικότερη αντιμετώπιση θεμάτων λήψης αποφάσεων των εμπλεκόμενων υπηρεσιών. Η εφαρμογή του προτεινόμενου συστήματος για την λήψη αποφάσεων στην δημόσια διοίκηση μέσω της πρόβλεψης διαφόρων παραγόντων υψηλής σημασίας, όπως

της εγκληματικότητας, περιβαλλοντικών και κοινωνικοοικονομικών παραγόντων, βοηθάει στην βέλτιστη λήψη αποφάσεων σε διάφορους τομείς, όπως στην πρόληψη και αντιμετώπιση της εγκληματικότητας, στην ρύπανση του περιβάλλοντος, στην ασφάλεια των συγκοινωνιών και στον οικονομικό και χωρικό σχεδιασμό.

Λέξεις κλειδιά: Τεχνητή Νοημοσύνη, Τεχνητά Νευρωνικά Δίκτυα, Πρόβλεψη Εγκληματικότητας Περιβαλλοντικών και Κοινωνικο-οικονομικών Παραγόντων, Λήψη Αποφάσεων στη Δημόσια Διοίκηση, Γεωγραφικά Συστήματα Πληροφοριών, Νοημοσύνη Σμήνους Σωματιδίων, Τυχαία Δάση Δέντρων Απόφασης, Τηλεπισκόπηση, Ταξινόμηση Δορυφορικών Εικόνων, Μηχανές Διανυσμάτων Υποστήριξης.

Abstract

This PhD thesis investigates the development of a methodology for the application of artificial intelligence and the information and communication technology in combination with the use of geographic information systems and remote sensing in decision-making in public management and judicial management services such as the Police Authorities, the Prosecutor's Office, Courts, Spatial Planning and Socio-Economic organizations, Educational Units, Municipal Authorities, environmental and water management organizations and other public services. The development of the methodology is based on the application of artificial intelligence methods, such as Artificial Neural Networks, Support Vector Machines, Random Forests of Decision Trees and Particle Swarm Intelligence in order to predict factors important in public management such as crime, environmental pollution factors such as ozone, socio-economic factors such as unemployment. In the present thesis, a new methodology is proposed, referred to as Predictive Management based on Artificial Intelligence. A new methodology has also been developed for the spatial prediction of decision factors such as crime, based on artificial intelligence techniques and geographic information systems. In addition, two new hybrid methods for classification of satellite images in public management were developed, based on a new proposed hybrid Support Vector Machine kernel and a new hybrid model of Artificial Neural Networks and Random Forest Decision Trees. The artificial Intelligence models were optimized by using Automated Machine Learning algorithms, such as Particle Swarm Optimization and Bayesian Optimization. The main objective of the proposed methodology in this thesis is the application of appropriate artificial intelligence methods to better address the decision-making issues of the public services involved. The implementation of the proposed system in decision-making in the public management through the prediction of various factors of major importance, such as crime, environmental and socio-economic factors, helps to optimize decision-making in various areas, such as crime prevention, environmental pollution, transportation safety and economic and spatial planning.

Keywords: Artificial Intelligence, Artificial Neural Networks, Crime Environmental and Socio-Economic Factors Forecasting, Decision Making in Public Management, Geographic Information Systems, Particle Swarm Intelligence, Random Forest Decision Trees, Remote Sensing, Satellite Image Classification, Support Vector Machines.

Περιεχόμενα

Περίληψη	7
Abstract.....	9
Κατάλογος Εικόνων.....	21
Κατάλογος Πινάκων	30
Κατάλογος Αρκτικόλεξων	36
Ευχαριστίες.....	33

ΚΕΦΑΛΑΙΟ 1

ΕΙΣΑΓΩΓΗ

1.1 Αντικείμενο – Στόχοι	40
1.2 Δομή της διατριβής.....	43

ΚΕΦΑΛΑΙΟ 2

ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ

2.1 Τεχνητή Νοημοσύνη	45
-----------------------------	----

Γενικά στοιχεία	45
2.2 Τεχνητά Νευρωνικά Δίκτυα.....	46
2.2.1 Γενικά στοιχεία	46
2.2.2 Βιολογικός και Τεχνητός Νευρώνας.....	47
2.2.3 Συναρτήσεις Μεταφοράς των Νευρωνικών Δικτύων.....	49
2.2.4 Δομικά στοιχεία Τεχνητού Νευρωνικού Δικτύου	51
2.2.5 Κατηγορίες Τεχνητών Νευρωνικών Δικτύων	52
2.2.6 Κανόνες ανανέωσης βαρών των νευρωνικών δικτύων.....	53
2.2.7 Αλγόριθμος οπισθόδρομης διάδοσης σφάλματος (backpropagation)	55
2.2.8 Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης.....	56
2.2.9 Νευρωνικά Δίκτυα Μη Γραμμικής Αυτοπαλινδρόμησης Εξωγενή Παράγοντα.....	57
2.3 Μηχανές Διανυσμάτων Υποστήριξης	58
2.3.1 Γενικά.....	58
2.3.2 Πυρήνες Μηχανών Διανυσμάτων Υποστήριξης	60
2.4 Τυχαία Δάση Δέντρων Απόφασης	61
2.4.1 Γενικά.....	61
2.4.2 Τεχνική Bagging στα Τυχαία Δάση	63
2.5 Γεωγραφικά Συστήματα Πληροφοριών	64
2.6 Τηλεπισκόπηση.....	66

ΚΕΦΑΛΑΙΟ 3

ΜΕΘΟΔΟΛΟΓΙΑ

3.1	Μεθοδολογία προσέγγισης	67
3.2	Σχετική έρευνα	69
3.3	Η παρούσα έρευνα	71
3.4	Παράγοντες στη λήψη αποφάσεων στη δημόσια διοίκηση	74
3.4.1	Παράγοντες εγκληματικότητας στην δικαστική διοίκηση	74
3.4.2	Παράγοντες σημαντικοί στην περιβαλλοντική διοίκηση	75
3.4.3	Παράγοντες κοινωνικοοικονομικοί σημαντικοί στη Δημόσια διοίκηση.....	77
3.5	Προτεινόμενη μεθοδολογία λήψης αποφάσεων - Προβλεπτική Διοίκηση	78

ΚΕΦΑΛΑΙΟ 4

ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΣΤΗΝ ΠΡΟΒΛΕΨΗ ΤΗΣ ΕΓΚΛΗΜΑΤΙΚΟΤΗΤΑΣ ΣΤΗΝ ΔΗΜΟΣΙΑ ΚΑΙ ΔΙΚΑΣΤΙΚΗ ΔΙΟΙΚΗΣΗ

4.1	Συνοπτική Περιγραφή της Έρευνας	84
4.2	Εισαγωγή – Σχετική Έρευνα	85
4.3	Θεωρητικό πλαίσιο	86
4.3.1	Τεχνητά Νευρωνικά Δίκτυα.....	86
4.3.2	Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης.....	87

4.3.3	Αλγόριθμος Levenberg Marquardt.....	87
4.4	Μεθοδολογία.....	88
4.5	Αποτελέσματα	89
4.5.1	Συλλογή Δεδομένων – Περιοχή Μελέτης.....	89
4.5.2	Μοντέλα Πρόβλεψης Τεχνητών Νευρωνικών Δικτύων	91
4.5.3	Βελτιστοποίηση με Αναζήτηση Πλέγματος.....	92
4.5.4	Βελτιστοποίηση Σμήνους Σωματιδίων (Particle Swarm Optimization)	96
4.5.5	Αλγόριθμος Βελτιστοποίησης Νευρωνικού Δικτύου με Σμήνος Σωματιδίων	99
4.5.6	Μοντέλα Μηχανών Διανυσμάτων Υποστήριξης στην Παλινδρόμηση 102	
4.5.7	Βελτιστοποίηση Bayes στα Μοντέλα Μηχανών Διανυσμάτων Υποστήριξης.....	104
4.5.8	Βελτιστοποιημένα Μοντέλα Μηχανών Διανυσμάτων Υποστήριξης 107	
4.5.9	Σύγκριση Προβλέψεων Βέλτιστων Μοντέλων SVM και FFNN	118
4.6	Συμπεράσματα και Συζήτηση	119

ΚΕΦΑΛΑΙΟ 5

Η ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΚΑΙ ΤΑ GIS ΣΤΗΝ ΧΩΡΙΚΗ ΠΡΟΒΛΕΨΗ ΤΗΣ ΕΓΚΛΗΜΑΤΙΚΟΤΗΤΑΣ ΣΤΗ ΔΗΜΟΣΙΑ ΚΑΙ ΔΙΚΑΣΤΙΚΗ ΔΙΟΙΚΗΣΗ

5.1	Συνοπτική Περιγραφή της Έρευνας	122
-----	---------------------------------------	-----

5.2	Εισαγωγή – Σχετική Έρευνα	122
5.3	Θεωρητικό πλαίσιο	125
5.3.1	Ανάλυση με τα γεωγραφικά συστήματα πληροφοριών.....	125
5.3.2	Ανάλυση θερμών σημείων.....	125
5.3.3	Χωρική κατανομή	126
5.3.4	Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης.....	128
5.3.5	Ο αλγόριθμος Levenberg-Marquardt	128
5.4	Μεθοδολογία.....	130
5.5	Αποτελέσματα	132
5.5.1	Συλλογή Δεδομένων – Περιοχή Μελέτης.....	132
5.5.2	Χωρικές μέθοδοι ανάλυσης.....	133
5.6	Τεχνητά νευρωνικά δίκτυα	133
5.6.1	Εφαρμογή	133
5.6.2	Ορισμός συνόλων εκπαίδευσης, επικύρωσης και δοκιμών.....	134
5.6.3	Τοπολογία βέλτιστου νευρωνικού μοντέλου πρόβλεψης.....	134
5.6.4	Βελτιστοποίηση Νευρωνικού Δικτύου με Σμήνος Σωματιδίων	137
5.6.5	Προβλέψεις του Προτεινόμενου Μοντέλου	139
5.6.6	Συμπεράσματα – Συζήτηση	140

ΚΕΦΑΛΑΙΟ 6

ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΣΤΗΝ ΠΡΟΒΛΕΨΗ ΠΕΡΙΒΑΛΛΟΝΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ ΣΤΗΝ ΔΗΜΟΣΙΑ ΔΙΟΙΚΗΣΗ

6.1	Συνοπτική Περιγραφή της μελέτης.....	144
6.2	Εισαγωγή – Σχετική Έρευνα	145
6.3	Θεωρητικό πλαίσιο	147
6.3.1	Τεχνητά Νευρωνικά Δίκτυα.....	147
6.3.2	Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης.....	147
6.3.3	Αλγόριθμος Κλιμακούμενων Συζυγών Κλίσεων	148
6.3.4	Περιβαλλοντικοί ρύποι.....	148
6.4	Μεθοδολογία Έρευνας	149
6.5	Αποτελέσματα	149
6.5.1	Συλλογή Δεδομένων – Περιοχή Μελέτης.....	149
6.5.2	Διανύσματα Εισόδου - Εξόδου.....	151
6.5.3	Βελτιστοποίηση Τοπολογίας Νευρωνικού Μοντέλου.....	151
6.5.4	Βελτιστοποίηση Νευρωνικού Δικτύου με Σμήνος Σωματιδίων	155
6.5.5	Βελτιστοποιημένα Μοντέλα Μηχανών Διανυσμάτων Υποστήριξης	156
6.5.6	Προτεινόμενο Υβριδικό Μοντέλο	166
6.5.7	Σύγκριση Βέλτιστων Μοντέλων SVM, FFNN και του Υβριδικού Μοντέλου	167
6.6	Συμπεράσματα και Συζήτηση	168

ΚΕΦΑΛΑΙΟ 7

ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΣΕ ΠΡΟΒΛΕΨΕΙΣ ΚΟΙΝΩΝΙΚΟ-ΟΙΚΟΝΟΜΙΚΩΝ ΠΑΡΑΓΟΝΤΩΝ

7.1	Συνοπτική Περιγραφή της μελέτης.....	171
7.2	Εισαγωγή – Σχετική Έρευνα	172
7.3	Θεωρητικό πλαίσιο	173
7.3.1	Τεχνητά Νευρωνικά Δίκτυα.....	173
7.3.2	Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης.....	174
7.3.3	Αλγόριθμος εκπαίδευσης Levenberg Marquardt	174
7.3.4	Ανεργία.....	175
7.4	Μεθοδολογία Έρευνας	176
7.5	Αποτελέσματα	177
7.5.1	Συλλογή Δεδομένων – Περιοχή Μελέτης.....	177
7.5.2	Μοντέλα Πρόβλεψης Τεχνητών Νευρωνικών Δικτύων	178
7.5.3	Βελτιστοποίηση Τοπολογίας του Νευρωνικού Δικτύου	179
7.5.4	Βελτιστοποίηση Νευρωνικού Δικτύου με Σμήνος Σωματιδίων	183
7.5.5	Βελτιστοποιημένα Μοντέλα Μηχανών Διανυσμάτων Υποστήριξης	184
7.5.6	Σύγκριση Προβλέψεων Βέλτιστων Μοντέλων SVM και FFNN	193
7.6	Συμπεράσματα και Συζήτηση	194

ΚΕΦΑΛΑΙΟ 8

ΤΗΛΕΠΙΣΚΟΠΗΣΗ ΚΑΙ ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΣΤΗΝ ΤΑΞΙΝΟΜΗΣΗ ΔΟΡΥΦΟΡΙΚΩΝ ΕΙΚΟΝΩΝ ΣΤΗ ΔΗΜΟΣΙΑ ΔΙΟΙΚΗΣΗ

8.1	Περιγραφή της Έρευνας	197
8.2	Εισαγωγή	199
8.2.1	Ταξινόμηση Δορυφορικών Εικόνων	199
8.2.2	Μηχανές Διανυσμάτων Υποστήριξης	199
8.2.3	Τυχαία Δάση Δέντρων Απόφασης	201
8.2	Σχετική Έρευνα	202
8.3	Μεθοδολογία	203
8.4	Συλλογή και προεπεξεργασία δεδομένων	205
8.5	Ανάπτυξη Μοντέλων Μηχανών Διανυσμάτων Υποστήριξης	206
8.6	Ανάπτυξη μοντέλων με Τυχαία Δάση Δέντρων Απόφασης	211
8.7	Το προτεινόμενο υβριδικό μοντέλο ταξινόμησης	214
8.7.1	Στάδιο 1ο: Υπολογισμός δείκτη Gini και πίνακα εγγύτητας	215
8.7.2	Στάδιο 2ο: Τροφοδότηση Νευρωνικού Δικτύου και Μηχανών Διανυσμάτων Υποστήριξης με διάνυσμα εισόδου τον πίνακα εγγύτητας φύλλων	218
8.8	Μετρικές αξιολόγησης ταξινόμησης αλγορίθμων τεχνητής νοημοσύνης	226
8.8.1	Ακρίβεια (precision)	226

8.8.2	Ανάκληση (Recall)	227
8.8.3	Ειδικευση (Specificity)	227
8.8.4	F-measure	227
8.8.5	Συντελεστής Kappa	228
8.9	Σύγκριση μοντέλων με τις μετρικές ταξινόμησης – Συζήτηση.....	229
8.10	Συμπεράσματα.....	231

ΚΕΦΑΛΑΙΟ 9

ΣΥΜΠΕΡΑΣΜΑΤΑ - ΣΥΖΗΤΗΣΗ

9.1	Συμπεράσματα - Συζήτηση	232
9.2	Περιορισμοί – Επεκτάσεις,,.....	239

ΚΕΦΑΛΑΙΟ 10

ΒΙΒΛΙΟΓΡΑΦΙΑ

10.1	Ελληνική βιβλιογραφία	242
10.2	Ξενόγλωσση βιβλιογραφία	244

Κατάλογος Εικόνων

Εικόνα 1.1.	Βαθμολογία χρήσης Τεχνολογίας Πληροφοριών & Επικοινωνιών στη δικαιοσύνη στα κράτη μέλη της Ε.Ε. (min=0, max=4). Πηγή: Dubois et al. (2013).	41
Εικόνα 2.1.	Μοντέλο ανθρώπινου νευρώνα.	47
Εικόνα 2.2.	Απλοποιημένη αναπαράσταση τεχνητού νευρώνα σε αντιστοιχία με τον ανθρώπινου νευρώνα.	48
Εικόνα 2.3.	Γραφική παράσταση της εφαπτομενικής – σιγμοειδούς συνάρτησης μεταφοράς.	49
Εικόνα 2.4.	Γραφική παράσταση της λογιστικής – σιγμοειδούς συνάρτησης μεταφοράς.	50
Εικόνα 2.5.	Γραφική παράσταση της γραμμικής συνάρτησης μεταφοράς.	50
Εικόνα 2.6.	Γραφική παράσταση της βηματικής συνάρτησης.	51
Εικόνα 2.7.	Η τοπολογία ενός τυπικού πολυεπίπεδου νευρωνικού δικτύου εμπρόσθιας διάδοσης.	57
Εικόνα 2.8.	Παράδειγμα τοπολογίας νευρωνικού μοντέλου Μη Γραμμικής Αυτοπαλινδρόμησης με Εξωγενή Παράγοντα.	58
Εικόνα 2.9.	Απλοποιημένη δομική αναπαράσταση Μηχανών Διανυσμάτων Υποστήριξης.	59

- Εικόνα 2.10.** Μοντελοποίηση του πραγματικού κόσμου σε επίπεδα στα GIS. 65
- Εικόνα 3.1.** Επισκόπηση της γενικής μεθοδολογίας προσέγγισης του ερευνητικού προβλήματος. 69
- Εικόνα 3.2.** Ποσοστιαία μεταβολή των δεικτών εγκληματικότητας από το έτος 2016 στο έτος 2017 που προτείνονται από την Αστυνομία του Τορόντο. 75
- Εικόνα 3.3.** Ένταση εκπομπών SO_x και NO_x μεταξύ των χωρών του ΟΟΣΑ το 2008 και η % αλλαγή από το 1990 στις συνολικές εκπομπές SO_x και NO_x (Πηγή: Directorate, 2008). 76
- Εικόνα 3.4.** Το προτεινόμενο μοντέλο προβλεπτικής διοίκησης με την χρήση βελτιστοποιημένων μοντέλων Τεχνητής Νοημοσύνης με μεθόδους Αυτόματης Μηχανικής Μάθησης (AutoML). Οι n παράγοντες (Factors) εισόδου την χρονική στιγμή k συμβολίζονται ως $F_1(k)$, $F_2(k)$, ..., $F_n(k)$, ενώ οι m παράμετροι (Parameters) εξόδου (εξαρτημένες μεταβλητές) της επόμενης χρονικής στιγμής $k+1$ συμβολίζονται ως $P_1(k+1)$, $P_2(k+1)$, ..., $P_n(k+1)$. 80
- Εικόνα 3.5.** Συνοπτική παρουσίαση των σταδίων της προτεινόμενης μεθοδολογίας για την λήψη αποφάσεων στη δημόσια και δικαστική διοίκηση με την εφαρμογή της Τεχνητής Νοημοσύνης. 83
- Εικόνα 4.1.** Επισκόπηση των σταδίων της εφαρμοσθείσας μεθοδολογίας για την πρόβλεψη των ποσοστών εγκληματικότητας με την εφαρμογή μεθόδων τεχνητής νοημοσύνης. 89
- Εικόνα 4.2.** Το σχήμα αναπαριστά την τοπολογία του βέλτιστου μοντέλου τεχνητού νευρωνικού δικτύου με δύο κρυφά επίπεδα (hidden layers). Το πρώτο κρυφό επίπεδο αποτελείται από 15 νευρώνες και χρησιμοποιεί την εφαπτομενική – σιγμοειδής συνάρτηση μεταφοράς, ενώ το δεύτερο

- επίπεδο αποτελείται από 27 νευρώνες και χρησιμοποιεί την γραμμική συνάρτηση μεταφοράς. 95
- Εικόνα 4.3.** Η γραφική παράσταση απόδοσης του βέλτιστου μοντέλου τεχνητού νευρωνικού δικτύου, σύμφωνα με το ελάχιστο Μέσο Τετραγωνικό Σφάλμα (Mean Squared Error) συναρτήσει των εποχών. 95
- Εικόνα 4.4.** Η γραφική παράσταση της κίνησης ενός σωματιδίου στο χώρο αναζήτησης, στον αλγόριθμο Βελτιστοποίησης Σμήνους Σωματιδίων. Η καλύτερη λύση του μεμονωμένου σωματιδίου αναπαρίσταται ως Pbest (Particle best), η καλύτερη λύση του σμήνους αναπαρίσταται ως Gbest (Global best) στη χρονική στιγμή k. 99
- Εικόνα 4.5.** Η γραφική παράσταση της μεταβολής του καλύτερου κόστους συναρτήσει των επαναλήψεων (iterations) του αλγορίθμου Βελτιστοποίησης Σμήνους Σωματιδίων κατά την βελτιστοποίηση των βαρών του νευρωνικού δικτύου. 101
- Εικόνα 4.6.** Η γραφική παράσταση της αρχιτεκτονικής των Μηχανών Διανυσμάτων υποστήριξης, όπου αναπαρίστανται το διάνυσμα εισόδου, το διάνυσμα εξόδου, τα επίπεδα εισόδου, εξόδου καθώς και το κρυφό επίπεδο. 102
- Εικόνα 4.7.** Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον Γκαουσιανό πυρήνα (Gaussian Kernel) Μηχανών Διανυσμάτων Υποστήριξης. 111
- Εικόνα 4.8.** Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον ακτινικό πυρήνα (RBF Kernel) Μηχανών Διανυσμάτων Υποστήριξης. 113

- Εικόνα 4.9.** Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον πολυωνυμικό πυρήνα (Polynomial Kernel) Μηχανών Διανυσμάτων Υποστήριξης. 115
- Εικόνα 4.10.** Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον γραμμικό πυρήνα (Linear Kernel) Μηχανών Διανυσμάτων Υποστήριξης. 117
- Εικόνα 4.11.** Συγκριτική γραφική παράσταση των προβλέψεων του βέλτιστου μοντέλου νευρωνικού δικτύου, του βέλτιστου μοντέλου SVM και της πραγματικής τιμής για τα τελευταία 36 χρονικά βήματα (μήνες) του δείγματος. 119
- Εικόνα 5.1** Επισκόπηση των σταδίων της μεθοδολογίας που ακολουθήθηκε για την χωρική πρόβλεψη της εγκληματικότητας. 131
- Εικόνα 5.2** Το σχήμα αναπαριστά την τοπολογία του βέλτιστου μοντέλου τεχνητού νευρωνικού δικτύου για την πρόβλεψη των συντεταγμένων των θερμών σημείων με την χρήση δύο κρυφών επιπέδων (hidden layers). Το πρώτο κρυφό επίπεδο αποτελείται από 22 νευρώνες και χρησιμοποιεί την λογιστική – σιγμοειδής συνάρτηση μεταφοράς, ενώ το δεύτερο επίπεδο αποτελείται από 24 νευρώνες και χρησιμοποιεί την εφαπτομενική – σιγμοειδής συνάρτηση μεταφοράς. 137
- Εικόνα 5.3** Απεικόνιση χάρτη σύγκρισης μεταξύ των κέντρων, των πραγματικών και των προβλεφθέντων θερμών σημείων από το βέλτιστο μοντέλο πρόβλεψης. Τα πραγματικά κέντρα hotspot απεικονίζονται με κόκκινο χρώμα και αυτά που προβλέφθηκαν με μπλε χρώμα. 140

- Εικόνα 6.1** Αναπαράσταση της διακύμανσης των τιμών της ημερήσιας συγκέντρωση του όζοντος (ppm) στο δείγμα στην περιοχή μελέτης. 150
- Εικόνα 6.2** Το σχήμα αναπαριστά την τοπολογία του βέλτιστου μοντέλου τεχνητού νευρωνικού δικτύου για την πρόβλεψη των τιμών του όζοντος O_3 με την χρήση δύο κρυφών επιπέδων (hidden layers). Το πρώτο κρυφό επίπεδο αποτελείται από 22 νευρώνες και χρησιμοποιεί την γραμμική συνάρτηση μεταφοράς, ενώ το δεύτερο επίπεδο αποτελείται από 25 νευρώνες και χρησιμοποιεί την γραμμική συνάρτηση μεταφοράς. 153
- Εικόνα 6.3** Το σχήμα αναπαριστά την τοπολογία του βέλτιστου μοντέλου τεχνητού νευρωνικού δικτύου για την πρόβλεψη των τιμών του όζοντος O_3 , τα επίπεδα εισόδου, εξόδου, τα κρυφά επίπεδα και τους νευρώνες των κρυφών επιπέδων. 154
- Εικόνα 6.4.** Η γραφική παράσταση απόδοσης του βέλτιστου μοντέλου νευρωνικού δικτύου. 155
- Εικόνα 6.5.** Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον Γκαουσιανό πυρήνα (Gaussian Kernel) Μηχανών Διανυσμάτων Υποστήριξης. 159
- Εικόνα 6.6.** Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον ακτινικό πυρήνα (RBF Kernel) Μηχανών Διανυσμάτων Υποστήριξης. 161
- Εικόνα 6.7.** Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον

- πολυωνυμικό πυρήνα (Polynomial Kernel) Μηχανών Διανυσμάτων Υποστήριξης. 163
- Εικόνα 6.8.** Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον γραμμικό πυρήνα (Linear Kernel) Μηχανών Διανυσμάτων Υποστήριξης. 165
- Εικόνα 6.9.** Συγκριτική γραφική παράσταση των προβλέψεων του βέλτιστου μοντέλου νευρωνικού δικτύου, του βέλτιστου μοντέλου SVM, του υβριδικού μοντέλου και της πραγματικής τιμής για τα τελευταία 30 χρονικά βήματα (ημέρες) του δείγματος. 168
- Εικόνα 7.1** Ανεργία στην χώρας της Ευρωπαϊκής Ένωσης από το 2004 έως το 2015. Πηγή: Eurostat 175
- Εικόνα 7.2** Επισκόπηση σταδίων εφαρμοσθείσας μεθοδολογίας για την ανάπτυξη και εφαρμογή των μοντέλων τεχνητής νοημοσύνης. 177
- Εικόνα 7.3** Το σχήμα αναπαριστά την μεταβολή των τιμών της ανεργίας (σε χιλιάδες) στο Ηνωμένο Βασίλειο από το 1971 έως το 2013. 178
- Εικόνα 7.4** Το σχήμα αναπαριστά την τοπολογία του βέλτιστου μοντέλου τεχνητού νευρωνικού δικτύου για την πρόβλεψη των τιμών της ανεργίας με την χρήση δύο κρυφών επιπέδων (hidden layers). Το πρώτο κρυφό επίπεδο αποτελείται από 14 νευρώνες και χρησιμοποιεί την γραμμική συνάρτηση μεταφοράς, ενώ το δεύτερο επίπεδο αποτελείται από 9 νευρώνες και χρησιμοποιεί την εφαστομενική – σιγμοειδής συνάρτηση μεταφοράς. 182
- Εικόνα 7.5** Η εικόνα αναπαριστά σχηματικά την τοπολογία και τις συνδέσεις των νευρώνων και του αριθμού των νευρώνων των κρυφών επιπέδων του

	βέλτιστου μοντέλου τεχνητού νευρωνικού δικτύου για την πρόβλεψη των τιμών της ανεργίας.	182
Εικόνα 7.6	Η γραφική παράσταση απόδοσης του βέλτιστου μοντέλου νευρωνικού δικτύου, σύμφωνα με το Μέσο Τετραγωνικό Σφάλμα (Mean Squared Error).	183
Εικόνα 7.7	Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον Γκαουσιανό πυρήνα (Gaussian Kernel) Μηχανών Διανυσμάτων Υποστήριξης.	186
Εικόνα 7.8	Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον ακτινικό πυρήνα (RBF Kernel) Μηχανών Διανυσμάτων Υποστήριξης.	188
Εικόνα 7.9	Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον πολωνυμικό πυρήνα (Polynomial Kernel) Μηχανών Διανυσμάτων Υποστήριξης.	190
Εικόνα 7.10.	Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον γραμμικό πυρήνα (Linear Kernel) Μηχανών Διανυσμάτων Υποστήριξης.	192
Εικόνα 7.11	Συγκριτική γραφική παράσταση των προβλέψεων του βέλτιστου μοντέλου νευρωνικού δικτύου, του βέλτιστου μοντέλου SVM και της	

	πραγματικής τιμής για τα τελευταία 20 χρονικά βήματα (έτη) του δείγματος.	194
Εικόνα 8.1	Επισκόπηση της ακολουθείσας μεθοδολογίας για την ταξινόμηση της δορυφορικής εικόνας.	203
Εικόνα 8.2	Δορυφορική εικόνα Landsat 7 χωρικής ανάλυσης 30μ.	206
Εικόνα 8.3	Μεθοδολογία δημιουργίας βέλτιστου μοντέλου πρόβλεψης Μηχανών Διανυσμάτων Υποστήριξης στην ταξινόμηση της δορυφορικής εικόνας.	207
Εικόνα 8.4	Μεθοδολογία δημιουργίας βέλτιστου μοντέλου πρόβλεψης Τυχαίων Δασών στην ταξινόμηση της δορυφορικής εικόνας.	211
Εικόνα 8.5	Διάγραμμα σπουδαιότητας μεταβλητών εισόδου στον καθορισμό των κλάσεων μέσω του διαχωρισμού των κόμβων σε θυγατρικούς στα δέντρα στα Τυχαία Δάση.	217
Εικόνα 8.6	Το 2 ^ο στάδιο της προτεινόμενης μεθοδολογίας για την δημιουργίας βέλτιστου μοντέλου πρόβλεψης.	218
Εικόνα 8.7	Διάγραμμα που αναπαριστά την αρχιτεκτονική του βέλτιστου τεχνητού νευρωνικού δικτύου ταξινόμησης που αποτελείται από 85 νευρώνες στο κρυφό επίπεδο και την εφαπτομενική – σιγμοειδής ως συνάρτηση μεταφοράς, ενώ στο επίπεδο εξόδου υλοποιεί την Softmax ως συνάρτηση μεταφοράς.	223
Εικόνα 8.8	Διάγραμμα μεταξύ της 1ης και της 8ης διάστασης που αναπαριστά τις ομάδες με βάση την εγγύτητα φύλλων στις 8 διαστάσεις.	224
Εικόνα 8.9	Διάγραμμα μεταξύ της 1ης και της 2ης διάστασης που αναπαριστά τις ομάδες με βάση την εγγύτητα φύλλων στις 8 διαστάσεις.	224
Εικόνα 8.10	Διάγραμμα μεταξύ της 3ης και της 7ης διάστασης που αναπαριστά τις ομάδες στις 8 διαστάσεις.	225

Εικόνα 8.11 Ταξινόμηση της δορυφορικής εικόνας σύμφωνα με το βέλτιστο υβριδικό μοντέλο. 226

Κατάλογος Πινάκων

Πίνακας 3.1	Βασικό σετ περιβαλλοντικών δεικτών που προτείνεται από τον Ευρωπαϊκό Οργανισμό Περιβάλλοντος (European Environment Agency) (EEA, 2005).....	77
Πίνακας 3.2	Σετ προτεινόμενων δεικτών – παραγόντων δημόσιας διοίκησης.....	77
Πίνακας 4.1.	Τα βέλτιστα νευρωνικά μοντέλα που βρέθηκαν για κάθε αλγόριθμο που δοκιμάστηκε χρησιμοποιώντας διαφορετικές τοπολογίες και συναρτήσεις ενεργοποίησης.....	94
Πίνακας 4.2.	Τα πέντε καλύτερα νευρωνικά μοντέλα του αλγορίθμου βελτιστοποίησης PSO που βρέθηκαν χρησιμοποιώντας διαφορετικές τοπολογίες παραμέτρων για την βελτιστοποίηση των βαρών του νευρωνικού δικτύου.....	101
Πίνακας 4.3.	Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον Γκαουσιανό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.....	109
Πίνακας 4.4.	Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον ακτινικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.	111

Πίνακας 4.5.	Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον πολυωνυμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης. Όπου d , ο βαθμός του πολυωνύμου.	113
Πίνακας 4.6.	Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον γραμμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.	115
Πίνακας 4.7.	Τα βέλτιστα μοντέλα των Μηχανών Διανυσμάτων Υποστήριξης (SVM) που βρέθηκαν με τον αλγόριθμο Bayes για κάθε πυρήνα σύμφωνα με την ελάχιστη τιμή της αντικειμενικής συνάρτησης.	117
Πίνακας 4.8.	Συγκριτικός πίνακας που απεικονίζει, τα βελτιστοποιημένα μοντέλα των Μηχανών Διανυσμάτων Υποστήριξης (SVM) με τον αλγόριθμο Bayes και των νευρωνικών δικτύων εμπρόσθιας διάδοσης (FFNN) με τον αλγόριθμο PSO.	118
Πίνακας 5.1.	Δεδομένα εγκληματικότητας που χρησιμοποιήθηκαν στην έρευνα αυτή.	132
Πίνακας 5.2.	Τα καλύτερα μοντέλα νευρωνικών δικτύων που ερευνήθηκαν σύμφωνα με τη ικανότητα γενίκευσής τους με διαφορετικούς αλγορίθμους εκπαίδευσης, τοπολογίες και συναρτήσεις ενεργοποίησης.	136
Πίνακας 5.3.	Τα πέντε καλύτερα νευρωνικά μοντέλα του αλγορίθμου βελτιστοποίησης PSO που βρέθηκαν χρησιμοποιώντας διαφορετικές τοπολογίες παραμέτρων για την βελτιστοποίηση των βαρών του νευρωνικού δικτύου.	139
Πίνακας 5.4.	Πίνακας σύγκρισης της υπάρχουσας βιβλιογραφίας σχετικά με την χρήση μεθόδων χωρικής ανάλυσης και νευρωνικών δικτύων για την πρόβλεψη της εγκληματικότητας χωρικά.	142

Πίνακας 6.1.	Τα βέλτιστα νευρωνικά μοντέλα που βρέθηκαν για κάθε αλγόριθμο που δοκιμάστηκε χρησιμοποιώντας διαφορετικές τοπολογίες και συναρτήσεις ενεργοποίησης.....	152
Πίνακας 6.2.	Τα πέντε καλύτερα νευρωνικά μοντέλα του αλγορίθμου βελτιστοποίησης PSO που βρέθηκαν χρησιμοποιώντας διαφορετικές τοπολογίες παραμέτρων για την βελτιστοποίηση των βαρών του νευρωνικού δικτύου.....	156
Πίνακας 6.3.	Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον Γκαουσιανό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.....	157
Πίνακας 6.4.	Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον ακτινικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.	159
Πίνακας 6.5.	Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον πολυωνυμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.....	161
Πίνακας 6.6.	Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον γραμμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.	163
Πίνακας 6.7.	Τα βέλτιστα μοντέλα των Μηχανών Διανυσμάτων Υποστήριξης (SVM) που βρέθηκαν με τον αλγόριθμο Bayes για κάθε πυρήνα σύμφωνα με την ελάχιστη τιμή της αντικειμενικής συνάρτησης.....	165
Πίνακας 6.8.	Συγκριτικός πίνακας που απεικονίζει, τα βελτιστοποιημένα μοντέλα των Μηχανών Διανυσμάτων Υποστήριξης (SVM) με τον αλγόριθμο Bayes και των νευρωνικών δικτύων εμπρόσθιας διάδοσης (FFNN) με τον αλγόριθμο PSO, ως προς την ακρίβεια πρόβλεψης.	167

Πίνακας 7.1.	Τα βέλτιστα νευρωνικά μοντέλα που βρέθηκαν για κάθε αλγόριθμο που δοκιμάστηκε χρησιμοποιώντας διαφορετικές τοπολογίες και συναρτήσεις ενεργοποίησης.....	181
Πίνακας 7.2.	Τα πέντε καλύτερα νευρωνικά μοντέλα του αλγορίθμου βελτιστοποίησης PSO που βρέθηκαν χρησιμοποιώντας διαφορετικές τοπολογίες παραμέτρων για την βελτιστοποίηση των βαρών του νευρωνικού δικτύου.....	184
Πίνακας 7.3.	Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον Γκαουσιανό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.....	185
Πίνακας 7.4.	Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον ακτινικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.	187
Πίνακας 7.5.	Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον πολωνυμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.....	189
Πίνακας 7.6.	Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον γραμμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.	191
Πίνακας 7.7.	Τα βέλτιστα μοντέλα των Μηχανών Διανυσμάτων Υποστήριξης (SVM) που βρέθηκαν με τον αλγόριθμο Bayes για κάθε πυρήνα σύμφωνα με την ελάχιστη τιμή της αντικειμενικής συνάρτησης.....	193
Πίνακας 7.8.	Συγκριτικός πίνακας που απεικονίζει, τα βελτιστοποιημένα μοντέλα των Μηχανών Διανυσμάτων Υποστήριξης (SVM) με τον αλγόριθμο Bayes και των νευρωνικών δικτύων εμπρόσθιας διάδοσης (FFNN) με τον αλγόριθμο PSO.....	193

Πίνακας 8.1.	Πίνακας σύγκρισης των καλύτερων μοντέλων κάθε πυρήνα των Μηχανών Διανυσμάτων Υποστήριξης βάσει του σφάλματος διασταυρωμένης επικύρωσης 10 φύλλων, για την ταξινόμηση της δορυφορικής εικόνας.....	209
Πίνακας 8.2.	Πίνακας σύγκρισης των καλύτερων μοντέλων κάθε πυρήνα των Μηχανών Διανυσμάτων Υποστήριξης στο σύνολο δοκιμής για την ταξινόμηση της δορυφορικής εικόνας.....	210
Πίνακας 8.3.	Πίνακας σύγκρισης μοντέλων ταξινόμησης με Τυχαία Δάση βάσει του σφάλματος διασταυρωμένης επικύρωσης 10 φύλλων.....	212
Πίνακας 8.4	Πίνακας σύγκρισης μοντέλων ταξινόμησης δορυφορικής εικόνας με Τυχαία Δάση στο σύνολο δοκιμής.	213
Πίνακας 8.5	Πίνακας σύγκρισης των καλύτερων μοντέλων κάθε πυρήνα των Μηχανών Διανυσμάτων Υποστήριξης του υβριδικού μοντέλου βάσει του σφάλματος διασταυρωμένης επικύρωσης 10 φύλλων, για την ταξινόμηση της δορυφορικής εικόνας.....	219
Πίνακας 8.6	Πίνακας σύγκρισης των καλύτερων μοντέλων κάθε πυρήνα των Μηχανών Διανυσμάτων Υποστήριξης του υβριδικού μοντέλου βάσει του σφάλματος διασταυρωμένης επικύρωσης 10 φύλλων.....	220
Πίνακας 8.7	Πίνακας σύγκρισης των καλύτερων μοντέλων του νευρωνικού δικτύου του υβριδικού μοντέλου βάσει του σφάλματος επικύρωσης.	222
Πίνακας 8.8	Πίνακας σύγκρισης στην ταξινόμηση της δορυφορικής εικόνας με το βέλτιστο προτεινόμενο υβριδικό μοντέλο νευρωνικού δικτύου.	229
Πίνακας 8.9	Πίνακας μετρικών για την σύγκριση των μοντέλων τεχνητής νοημοσύνης.	230
Πίνακας 9.1	Πίνακας σύνοψης των μεθόδων που μελετήθηκαν και σε ποιες δημόσιες υπηρεσίες μπορούν να αξιοποιηθούν.....	235

Πίνακας 9.2	Οι τομείς της δημόσιας διοίκησης, τα είδη αποφάσεων, που μπορεί να εφαρμοστεί η προτεινόμενη μεθοδολογία και η σχετική βιβλιογραφία	236
--------------------	--	-----

Κατάλογος Αρκτικόλεξων

ΑΕΠ – Εθνικό Ακαθάριστο Προϊόν

ΒΣΟ – Βελτιστοποίηση Σμήνους Σωματιδίων

ΜΔΥ – Μηχανές Διανυσμάτων Υποστήριξης

ΤΝΔ – Τεχνητά Νευρωνικά Δίκτυα

ΤΔΔΑ – Τυχαία Δάση Δέντρων Απόφασης

ACA – Ant Colony Algorithm

ANN – Artificial Neural Network

AutoML – Automated Machine Learning

BCA – Bee Colony Algorithm

BFGS-QN – BFGS Quasi-Newton

CDF – Cumulative Distribution Function

ESTF – Elliot Sigmoid Transfer Function

FFMLP – Feed-Forward Multi-Layer Perceptron

FFNN – Feed-Forward Neural Network

GDALR – Gradient Descent with Adaptive Learning Rate

GDMALRB – Gradient Descent with Momentum and Adaptive Learning Rate
Backpropagation

GDP – Growth Domestic Product

GIS – Geographic Information Systems

GP – Gaussian Process

LM – Levenberg – Marquardt

LSTF – Log-Sigmoid Transfer Function

LTF – Linear Transfer Function

MSE – Mean Squared Error

MLP – Multilayer Perceptron

NARX – Nonlinear Autoregressive Network with Exogenous Input

NDVI – Normalized Difference Vegetation Index

ODS – Ozone Depleting Substances

PDF – Probability Density Function

PLTF – Positive Linear Transfer Function

PPM – Parts Per Million

PSO – Particle Swarm Optimization

RB – Resilient Backpropagation

RBF – Radial Basis Function

RF – Random Forest

RMSE – Root Mean Square Error

RNN – Recurrent Neural Networks

SCG – Scaled Conjugate Gradient

SMBO – Sequential Model-Based Optimization

SVM – Support Vector Machine

TSTF – Tan-Sigmoid Transfer Function

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή Δρ. Κωνσταντίνο Περάκη και τον καθηγητή του Τμήματος Εφαρμοσμένης Πληροφορικής του Πανεπιστημίου Μακεδονίας Δρ. Αλέξανδρο Χατζηγεωργίου για την πολύτιμη καθοδήγηση τους καθ' όλη την διάρκεια της εκπόνησης της παρούσας διδακτορικής διατριβής, την καθηγήτρια Δρ. Μαρί Νοέλ Ντυκέν για την συμβολή της στην διατριβή, καθώς και τα μέλη της επταμελούς εξεταστικής επιτροπής.

Επίσης, θα ήθελα να ευχαριστήσω τους γονείς μου, για την στήριξή τους καθ' όλη την διάρκεια της εκπόνησης της διατριβής.

Γεώργιος Ν. Κουζιώκας

Βόλος, Οκτώβριος 2018

ΚΕΦΑΛΑΙΟ 1: Εισαγωγή

1.1 Αντικείμενο – Στόχοι

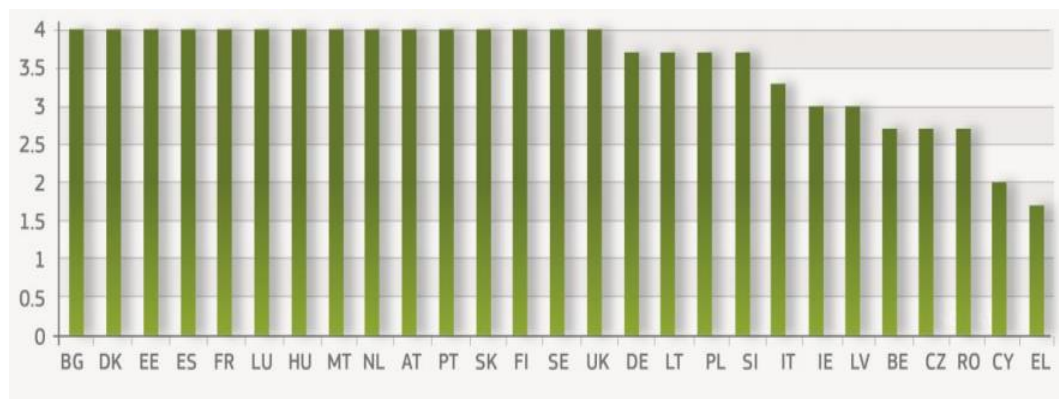
Η παρούσα διατριβή ερευνά την ανάπτυξη μεθοδολογίας για την εφαρμογή σύγχρονων μεθοδολογικών πλαισίων και συστημάτων με την χρήση της τεχνολογίας των πληροφοριών και επικοινωνιών και συγκεκριμένα με την εφαρμογή αλγορίθμων τεχνητής νοημοσύνης, σε συνδυασμό με την χρήση των γεωγραφικών συστημάτων πληροφοριών και της τηλεπισκόπησης στην λήψη αποφάσεων στην διοίκηση δικαστικών και δημοσίων υπηρεσιών.

Οι βασικοί στόχοι της μεθοδολογίας είναι η ανάπτυξη κατάλληλων μεθόδων με την χρήση της τεχνητής νοημοσύνης για την αποτελεσματικότερη διοίκηση των εμπλεκόμενων υπηρεσιών. Η ανάπτυξη μεθόδων πρόβλεψης της εγκληματικότητας βασισμένων στην τεχνητή νοημοσύνη και συγκεκριμένα η πρόβλεψη των θερμών σημείων (hotspots) με την μεγαλύτερη πυκνότητα των τελεσθέντων αδικημάτων, έχει ως σκοπό να βοηθήσει στην επικέντρωση των προσπάθειών των δικαστικών αρχών για την πρόληψη και καταπολέμηση της εγκληματικότητας και στην βέλτιστη κατανομή των αστυνομικών δυνάμεων και στην λήψη αποφάσεων στις εμπλεκόμενες δημόσιες και δικαστικές υπηρεσίες. Η μεθοδολογία που αναπτύσσεται στην παρούσα διατριβή, θα βοηθήσει στην βελτιστοποίηση της λήψης αποφάσεων σε διάφορες δημόσιες υπηρεσίες όπως για παράδειγμα οι αστυνομικές αρχές, οι δημοτικές αρχές, χωροταξικοί και περιβαλλοντικοί οργανισμοί, εκπαιδευτικοί οργανισμοί, στα πλαίσια της αναβάθμισης και του εκσυγχρονισμού της δημόσιας διοίκησης.

Η παρούσα διατριβή ερευνά την χρήση Τεχνητών Νευρωνικών Δικτύων αλλά και άλλων μεθόδων τεχνητής νοημοσύνης όπως οι Μηχανές Διανυσμάτων Υποστήριξης και τα Τυχαία Δάση σε συνδυασμό με μεθόδους χωρικής ανάλυσης στην πρόβλεψη της

εγκληματικότητας καθώς και άλλων παραγόντων (πχ περιβαλλοντικοί δείκτες ρύπανσης, ανεργία) στη λήψη αποφάσεων από τις εμπλεκόμενες δικαστικές και άλλες δημόσιες υπηρεσίες, ώστε να λαμβάνονται οι βέλτιστες αποφάσεις. Μελετήθηκε η χρήση διαφόρων τύπων νευρωνικών δικτύων και διαφόρων αλγορίθμων εκπαίδευσης νευρωνικών δικτύων και κατά πόσο επιδρούν διάφοροι παράγοντες και παράμετροι στην βελτιστοποίηση των προβλέψεων (πχ τύποι εγκλημάτων, κοινωνικοοικονομικοί παράγοντες, κλπ.) για την λήψη αποφάσεων. Επίσης, η προτεινόμενη μεθοδολογία εφαρμόστηκε σε πραγματικά δεδομένα στα οποία πραγματοποιήθηκε προεπεξεργασία ώστε να εφαρμοστούν οι αλγόριθμοι τεχνητής νοημοσύνης καθώς και οι χωρικές και τηλεπισκοπικές μέθοδοι.

Η προοπτική της αξιοποίησης της Τεχνολογίας των Πληροφοριών και Επικοινωνιών (Information and Communications Technology – ICT) στην στη δημόσια διοίκηση και την δικαιοσύνη έχει αναβαθμιστεί τα τελευταία χρόνια στην Ευρώπη. Η Ελλάδα όμως υστερεί κατά πολύ στην χρήση της Τεχνολογίας των Πληροφοριών & Επικοινωνιών στη δικαιοσύνη σε σχέση με τις άλλες ευρωπαϊκές χώρες (συγκεκριμένα βρίσκεται στην τελευταία θέση) όπως δείχνουν τα επίσημα στατιστικά στοιχεία της Ευρωπαϊκής Επιτροπής (Εικόνα 1.1).



Εικόνα 1.1 Βαθμολογία χρήσης Τεχνολογίας Πληροφοριών & Επικοινωνιών στη δικαιοσύνη στα κράτη μέλη της Ε.Ε. (min=0, max=4). Πηγή: Dubois et al. (2013).

Η σημασία της πορείας των ερευνών πάνω στην τεχνητή νοημοσύνη, φαίνεται και από την πρόσφατη απόφαση της Ευρωπαϊκής Επιτροπής να διαθέσει 20 δισεκατομμύρια

ευρώ μέχρι το τέλος του 2020, για να εφαρμοστεί μία σειρά μέτρων, ώστε να τεθεί η τεχνητή νοημοσύνη στην υπηρεσία των Ευρωπαίων πολιτών και επιπλέον για να προάγει την ανταγωνιστικότητα της Ευρώπης στον τομέα αυτό.¹

Η παρούσα έρευνα, μελετά την ανάπτυξη μεθοδολογιών λήψης αποφάσεων βασισμένων στην τεχνητή νοημοσύνη, τα γεωγραφικά συστήματα πληροφοριών και την τηλεπισκόπηση. Η μελέτη αυτή, μπορεί να διαδραματίσει σημαντικό ρόλο στη δημόσια διοίκηση και στη διαμόρφωση πρακτικών και διαδικασιών σε δημόσιους οργανισμούς και Δημόσιες Αρχές (Αστυνομικές αρχές, δημοτικές αρχές, κλπ.), όπου απαιτείται η αξιολόγηση δεδομένων, χωρικών ή μη. Επίσης, μπορεί να συμβάλλει στην αποτελεσματικότερη καταπολέμηση της εγκληματικότητας (πχ επικέντρωση των προσπαθειών δικαστικών και αστυνομικών αρχών στα χωρικά σημεία με την μεγαλύτερη πυκνότητα περιστατικών εγκληματικότητας (Hotspots)). Η χρήση της Τεχνολογίας των Πληροφοριών και Επικοινωνιών στην δημόσια διοίκηση και την δικαιοσύνη, είναι το κλειδί για τον εκσυγχρονισμό των πρακτικών λήψης αποφάσεων (Αποστολάκης κ.α., 2008), και η παρούσα εργασία σκοπεύει να συμβάλλει προς αυτή την κατεύθυνση, όπως επιτάσσεται και από την Ευρωπαϊκή Ένωση. Μέχρι στιγμής σύμφωνα με την βιβλιογραφία, δεν υπάρχει σχετική έρευνα που να εφαρμόζει την τεχνητή νοημοσύνη στη λήψη αποφάσεων στη δημόσια διοίκηση, σε συνδυασμό με τεχνολογίες όπως τα γεωγραφικά συστήματα πληροφοριών και η τηλεπισκόπηση και αυτό αποτελεί άλλον έναν παράγοντα πρωτοτυπίας της παρούσας έρευνας.

Λαμβάνοντας υπόψη τα παραπάνω, τα ερευνητικά ερωτήματα της διατριβής αφορούν, το πώς μπορεί να σχεδιαστεί μία τέτοια μεθοδολογία ώστε να περιλαμβάνει την χρήση των δυνατοτήτων που προσφέρει η τεχνητή νοημοσύνη, τα γεωγραφικά συστήματα πληροφοριών και η τηλεπισκόπηση, ώστε να συμβάλει στην αποτελεσματικότερη λήψη αποφάσεων σε δημόσιες και δικαστικές υπηρεσίες. Επίσης, το κατά πόσο είναι δυνατή η πρόβλεψη της εγκληματικότητας καθώς και άλλων παραμέτρων καθοριστικών στην λήψη αποφάσεων στην δημόσια διοίκηση (π.χ. παραγόντων ρύπανσης του περιβάλλοντος, κοινωνικο-οικονομικών παραγόντων όπως η

¹ http://europa.eu/rapid/press-release_IP-18-3362_en.htm

ανεργία, κλπ.), με την χρήση τεχνητών νευρωνικών δικτύων και άλλως μεθόδων Τεχνητής Νοημοσύνης καθώς και ποια είναι η βέλτιστη τοπολογία των αλγορίθμων αυτών για να επιτευχθούν τα βέλτιστα αποτελέσματα προβλέψεων.

1.2 Δομή της διατριβής

Η διατριβή αποτελείται από εννέα κεφάλαια το περιεχόμενο των οποίων περιγράφεται συνοπτικά στην συνέχεια.

Στο 1ο κεφάλαιο, αναφέρονται κάποια εισαγωγικά στοιχεία για το αντικείμενο της παρούσας διατριβής, τον σκοπό και τους στόχους, την συνεισφορά της και την διάρθρωση των κεφαλαίων της.

Στο 2^ο κεφάλαιο, αναπτύσσεται το θεωρητικό πλαίσιο της παρούσας διατριβής, με ανασκόπηση της σχετικής βιβλιογραφίας, περιγράφοντας τις μεθοδολογίες τεχνητής νοημοσύνης που χρησιμοποιήθηκαν όπως τα Τεχνητά Νευρωνικά Δίκτυα, τα Τυχαία Δάση και οι Μηχανές Διανυσμάτων Υποστήριξης. Επίσης, περιγράφονται οι τύποι των τεχνητών νευρωνικών δικτύων, οι παράγοντες της τοπολογίας τους που ερευνήθηκαν, όπως οι τύποι των συναρτήσεων ενεργοποίησης, οι νευρώνες των κρυφών επιπέδων.

Στο 3^ο κεφάλαιο, αναπτύσσεται η προτεινόμενη μεθοδολογία για την εφαρμογή της τεχνητής νοημοσύνης στην δημόσια και δικαστική διοίκηση με την χρήση επίσης, των γεωγραφικών συστημάτων πληροφοριών και της τηλεπισκόπησης.

Στο 4^ο κεφάλαιο, εφαρμόζονται οι τεχνικές τεχνητής νοημοσύνης σε πραγματικά δεδομένα για τη πρόβλεψη της εγκληματικότητας. Γίνεται αναφορά στην σχετική έρευνα, τη μεθοδολογία που ακολουθήθηκε, τα αποτελέσματα και η συζήτηση καθώς και η σύγκριση με άλλες σχετικές έρευνες.

Στο 5^ο κεφάλαιο, εφαρμόζονται οι τεχνικές τεχνητής νοημοσύνης και τα γεωγραφικά συστήματα πληροφοριών σε πραγματικά δεδομένα για την ανάπτυξη μίας νέας μεθοδολογίας για τη χωρική πρόβλεψη της εγκληματικότητας. Γίνεται αναφορά στην σχετική έρευνα, τη μεθοδολογία που ακολουθήθηκε, τα αποτελέσματα και η συζήτηση, καθώς και η σύγκριση της έρευνας αυτής με άλλες σχετικές έρευνες.

Στο 6^ο κεφάλαιο, ερευνάται η εφαρμογή και σύγκριση μοντέλων τεχνητής νοημοσύνης στην πρόβλεψη περιβαλλοντικών δεδομένων στη δημόσια διοίκηση, όπως παραγόντων ατμοσφαιρικής ρύπανσης. Στο κεφάλαιο αυτό, προτείνεται μία νέα υβριδική μέθοδος πρόβλεψης, η οποία συγκρινόμενη με τις άλλες τεχνικές τεχνητής νοημοσύνης που εφαρμόστηκαν, παρήγαγε αποτελέσματα βελτιωμένης ακρίβειας. Επίσης, επισκοπείται η σχετική έρευνα, και αναλύεται η μεθοδολογία που ακολουθήθηκε, τα αποτελέσματα και η συζήτηση, καθώς και πραγματοποιείται σύγκριση της έρευνας αυτής με άλλες σχετικές έρευνες.

Στο 7^ο κεφάλαιο, ερευνάται η εφαρμογή και σύγκριση μοντέλων τεχνητής νοημοσύνης στην πρόβλεψη κοινωνικο-οικονομικών παραγόντων στη δημόσια διοίκηση. Αναλύονται οι μεταβλητές που επηρεάζουν τους κοινωνικο-οικονομικούς παράγοντες που ερευνώνται, σύμφωνα με την βιβλιογραφία, ώστε να χρησιμοποιηθούν ως διανύσματα εισόδου στα μοντέλα τεχνητής νοημοσύνης. Επίσης, πραγματοποιείται μία ανασκόπηση της σχετικής έρευνας, αναλύεται η μεθοδολογία που ακολουθήθηκε, αναλύονται τα αποτελέσματα και η σύγκριση της έρευνας αυτής με άλλες σχετικές έρευνες.

Στο 8^ο κεφάλαιο, ερευνάται η εφαρμογή και σύγκριση μοντέλων τεχνητής νοημοσύνης και μεθόδων τηλεπισκόπησης στην ταξινόμηση δορυφορικών εικόνων στη δημόσια διοίκηση. Στο κεφάλαιο αυτό, προτείνεται μία νέα υβριδική προσέγγιση για την βελτίωση της ακρίβειας της ταξινόμησης, η οποία συγκρινόμενη με τις άλλες τεχνικές τεχνητής νοημοσύνης που εφαρμόστηκαν, παρήγαγε αποτελέσματα βελτιωμένης ακρίβειας. Επίσης, πραγματοποιείται ανασκόπηση της σχετικής έρευνας, και η σύγκριση της έρευνας αυτής με άλλες σχετικές έρευνες.

Στο 9^ο κεφάλαιο, πραγματοποιείται η συζήτηση για την εφαρμογή της Τεχνητής Νοημοσύνης, των Γεωγραφικών Συστημάτων Πληροφοριών και της Τηλεπισκόπησης με την προτεινόμενη μεθοδολογία, στη δημόσια και δικαστική διοίκηση και αποτιμούνται η συνεισφορά της παρούσας διατριβής, οι περιορισμοί και οι μελλοντικές της επεκτάσεις.

ΚΕΦΑΛΑΙΟ 2: Θεωρητικό Πλαίσιο

2.1 Τεχνητή Νοημοσύνη

Γενικά στοιχεία

Η Τεχνητή Νοημοσύνη (TN) αποτελεί κλάδο της πληροφορικής που ασχολείται με τη σχεδίαση και την υλοποίηση ευφύων υπολογιστικών συστημάτων που μιμούνται στοιχεία του ανθρώπινου εγκεφάλου όπως, μάθηση, προσαρμοστικότητα, επίλυση προβλημάτων, κλπ. Εμπλέκει πολλές επιστήμες πως της πληροφορικής, της ψυχολογίας, της νευρολογίας, της γλωσσολογίας κλπ. και έχει ως στόχο τη σύνθεση μίας ευφυούς συμπεριφοράς (Βλαχάβας κ.α., 2006; Zhang et al., 1998). Τα Τεχνητά Νευρωνικά Δίκτυα αποτελούν υπολογιστικά συστήματα που προσομοιώνουν τη δομή των νευρωνικών δικτύων του ανθρώπινου εγκεφάλου. Επεξεργάζονται δεδομένα από την είσοδο τους και μέσω των νευρωνικών συνδέσεων του δικτύου, αναπροσαρμόζουν τα συναπτικά βάρη, προκειμένου να παράγουν την κατάλληλη έξοδο (Zhang et al., 1998).

Ιστορικά, οι McCulloch and Pitts (1943), προτείνουν για πρώτη φορά ένα νέο μοντέλο με τεχνητούς νευρώνες που μπορούν να μαθαίνουν και να υπολογίζουν μία συνάρτηση. Στη συνέχεια, ο Alan Turing (1948), που θεωρείται ως ο εμπνευστής της τεχνητής νοημοσύνης, προτείνει το τεστ της μίμησης (τεστ Τούρινγκ), ώστε να χρησιμοποιηθεί για την αναγνώριση των ευφύων μηχανών. Αργότερα, ο Weizenbaum (1966), δημοσιεύει την έρευνά του σχετικά με επικοινωνία ανθρώπου – μηχανής σε φυσική γλώσσα και δημιουργεί το ELIZA.

Στην παρούσα διατριβή, θα ασχοληθούμε με διαφόρους αλγορίθμους τεχνητής νοημοσύνης που θεωρούνται οι πιο κατάλληλοι σε προβλέψεις χρονοσειρών (Koskela et

al., 1996; Tang and Fishwick, 1993; Zhang, 2001), όπως τα Τεχνητά Νευρωνικά Δίκτυα και συγκεκριμένα τα Τεχνητά Νευρωνικά Δίκτυα Εμπρόσθιας Διάδοσης (FeedForward Neural Networks). Επίσης, θα χρησιμοποιηθούν οι Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines), και τα Τυχαία Δάση Δέντρων Απόφασης (Random Forest Decision Trees). Αυτές οι τεχνικές μηχανικής μάθησης θα συγκριθούν μεταξύ τους ως προς την απόδοση τους στην πρόβλεψη χρονοσειρών παραγόντων σημαντικών στην δημόσια διοίκηση. Επίσης, τα Τεχνητά Νευρωνικά Δίκτυα Εμπρόσθιας Διάδοσης, οι Μηχανές Διανυσματικής Υποστήριξης και τα Τυχαία Δάση θα εφαρμοστούν στην ταξινόμηση δορυφορικών εικόνων και θα συγκριθούν ως προς την ακρίβειά τους με μια νέα υβριδική μέθοδο που προτείνεται στην παρούσα έρευνα. Στις παρακάτω ενότητες, θα περιγραφούν συνοπτικά τα είδη αυτά της τεχνητής νοημοσύνης.

2.2 Τεχνητά Νευρωνικά Δίκτυα

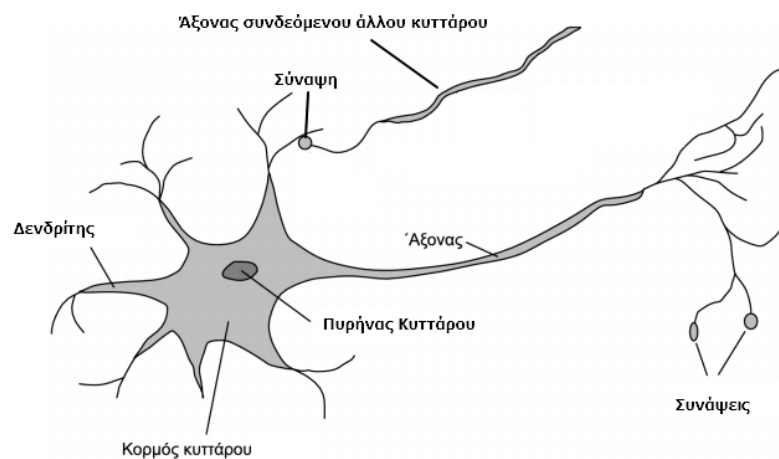
2.2.1 Γενικά στοιχεία

Τα Τεχνητά Νευρωνικά Δίκτυα είναι υπολογιστικά συστήματα που μιμούνται τη δομή των νευρωνικών δικτύων του ανθρώπινου εγκεφάλου. Τα τεχνητά νευρωνικά δίκτυα επεξεργάζονται δεδομένα από τις τιμές εισόδου. Οι πληροφορίες διοχετεύονται μέσω των συνδέσεων του νευρωνικού δικτύου προκειμένου να παράγουν την έξοδο σύμφωνα με τις τιμές εισόδου (Zhang et al., 1998).

Το πλεονέκτημα τους είναι ότι μπορούν να μοντελοποιήσουν και μη γραμμικές σχέσεις μεταξύ εισόδου και εξόδου (Svozil et al., 1997). Τα τεχνητά νευρωνικά δίκτυα (Artificial Neural Networks) έχουν προταθεί ως μαθηματικό μοντέλο προσομοίωσης της λειτουργίας του ανθρώπινου εγκεφάλου. Μιμούνται την δομή του εγκεφάλου η οποία επιτρέπει την παράλληλη επεξεργασία δεδομένων και τη δυνατότητα συνεχούς μάθησης μέσω της αλληλεπίδρασης επιτρέποντας την συνεχή εξέλιξη και μάθηση από το περιβάλλον του (Grossberg, 1988).

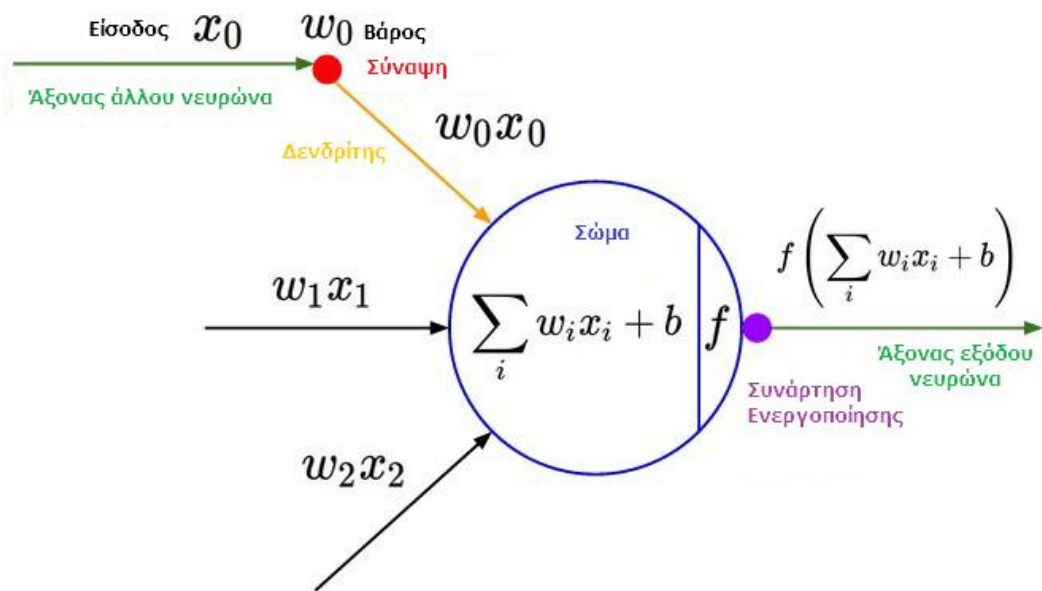
2.2.2 Βιολογικός και Τεχνητός Νευρώνας

Ο βιολογικός νευρώνας αποτελεί την θεμελιώδη κυτταρική δομική μονάδα του ανθρώπινου εγκεφάλου. Ως στοιχείο επεξεργασίας, δέχεται και συνδυάζει (επεξεργάζεται) σήματα από άλλους νευρώνες μέσα εισόδους που ονομάζονται δενδρίτες (Καμπουρλάζος και Παπακώστας, 2015; Τζελεπίδης, 2001). Αν το συνδυασμένο σήμα εισόδου είναι αρκετά ισχυρό προκαλεί ενεργοποίηση του νευρώνα, που παράγει με τη σειρά του ένα σήμα εξόδου. Το σήμα διαδίδεται μέσω του άξονα, ο οποίος συνδέεται με τους δενδρίτες άλλων νευρώνων. Κάθε σήμα περνάει μέσα από μια σύναψη ή συναπτικό σύνδεσμο (synaptic joint) που είναι γεμάτο με νευροδιαβιβαστικό υγρό, το οποίο είτε επιταχύνει, είτε επιβραδύνει τη ροή των ηλεκτρικών φορτίων (Zhang et al., 1998). Η βασική δράση των νευρώνων είναι χημική και το νευροδιαβιβαστικό υγρό παράγει ηλεκτρικά σήματα που πηγαίνουν στον πυρήνα του νευρώνα. Η ρύθμιση της αντίστασης ή της αγωγιμότητας του κενού της σύναψης είναι σημαντική, γιατί επηρεάζει τις ικανότητες μάθησης και απομνημόνευσης. Όσο οι συναπτικές δυνάμεις του νευρώνα προσαρμόζονται, ο εγκέφαλος μαθαίνει και αποθηκεύει πληροφορίες. Η βασική διαφορά στην επεξεργασία της πληροφορίας (σημάτων) είναι ότι στους εγκεφαλικούς νευρώνες και τους συναπτικούς δεσμούς γίνεται με τη βοήθεια ηλεκτροχημικών διεργασιών, ενώ στους τεχνητούς νευρώνες στα σύγχρονα ολοκληρωμένα κυκλώματα με ηλεκτρονική επεξεργασία σημάτων (Grossberg, 1988; Haykin, 1994).



Εικόνα 2.1 Μοντέλο ανθρώπινου νευρώνα.

Ο τεχνητός νευρώνας είναι η δομική μονάδα του τεχνητού νευρωνικού δικτύου όπου πραγματοποιείται η επεξεργασία της πληροφορίας. Κάθε νευρώνας επεξεργάζεται την κάθε πληροφορία που δέχεται, και δίνει μία τιμή εξόδου. Οι εισοδοί του είναι είτε οι εξοδοί άλλων νευρώνων, είτε το πρωταρχικό σήμα εισόδου. Υπάρχουν διάφορα είδη νευρώνων. Το είδος που θα επιλεγθεί για να δομηθεί ένα συγκεκριμένο τεχνητό νευρωνικό δίκτυο, εξαρτάται από τη φύση του κάθε προβλήματος (Zhang et al., 1998). Σε πολλές περιπτώσεις χρησιμοποιείται συνδυασμός διαφορετικών ειδών νευρώνα. Στην παρακάτω εικόνα, παρουσιάζεται το βασικό μοντέλο ενός τεχνητού νευρώνα.



Εικόνα 2.2 Απλοποιημένη αναπαράσταση τεχνητού νευρώνα σε αντιστοιχία με τον ανθρώπινου νευρώνα.

Η κατάσταση μίας μονάδας τεχνητού νευρώνα δίνεται από την τιμή ενεργοποίησής της α_i . Η κατάσταση ενός νευρωνικού δικτύου με N μονάδες τη χρονική στιγμή t μπορεί να αποδοθεί από το παρακάτω διάνυσμα που ονομάζεται και μνήμη μικρής διάρκειας (Short Term Memory-STM) (Haykin, 1994):

$$[\alpha_1(t) \alpha_2(t) \dots \alpha_N(t)] \quad (2.1)$$

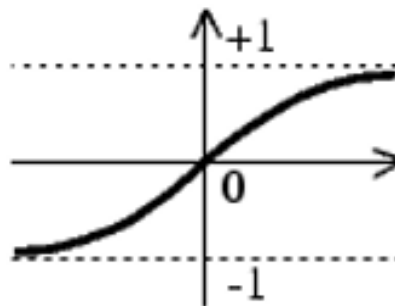
Η συνάρτηση εξόδου χρησιμοποιεί ως όρισμα την τιμή ενεργοποίησης για να υπολογίσει την έξοδο της μονάδας. Η τιμή εξόδου έπειτα, μεταφέρεται στις υπόλοιπες μονάδες του νευρωνικού δικτύου.

2.2.3 Συναρτήσεις Μεταφοράς των Νευρωνικών Δικτύων

Οι συνηθέστερες συναρτήσεις μεταφοράς που χρησιμοποιούνται στα νευρωνικά δίκτυα είναι: η εφαπτομενική – σιγμοειδής συνάρτηση μεταφοράς, η λογιστική – σιγμοειδής συνάρτηση μεταφοράς, η γραμμική συνάρτηση μεταφοράς και η βηματική συνάρτηση, οι οποίες εκφράζονται από τους παρακάτω μαθηματικούς τύπους (Βλαχάβας κ.α., 2006).

Η εφαπτομενική – σιγμοειδής συνάρτηση μεταφοράς (Tan-Sigmoid Transfer Function (TSTF)) εκφράζεται από τον τύπο:

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.2)$$

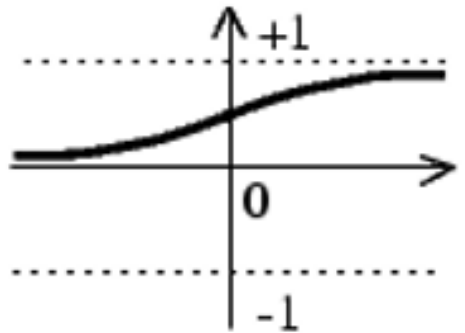


Εικόνα 2.3 Γραφική παράσταση της εφαπτομενικής – σιγμοειδούς συνάρτησης μεταφοράς.

Η λογιστική – σιγμοειδής συνάρτηση μεταφοράς (Log-Sigmoid Transfer Function (LSTF)) εκφράζεται ως:

$$f(x) = \frac{1}{1 + e^{-ax}} \quad (2.2)$$

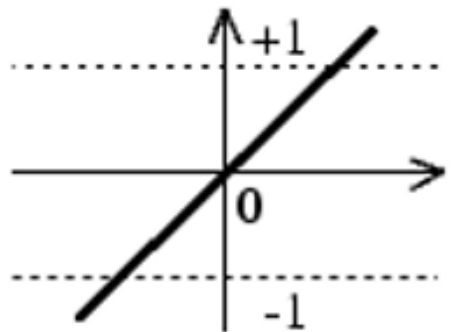
όπου a η παράμετρος κλίσης της συνάρτησης.



Εικόνα 2.4 Γραφική παράσταση της λογιστικής – σιγμοειδούς συνάρτησης μεταφοράς.

Η γραμμική συνάρτηση μεταφοράς (Linear Transfer Function (LTF)) εκφράζεται ως:

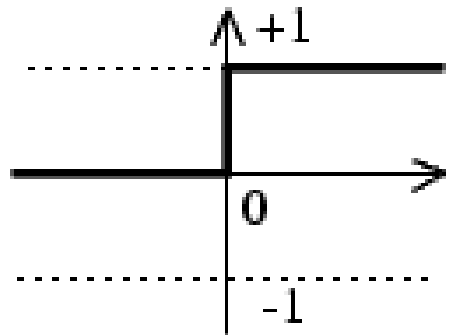
$$f(x) = x \quad (2.3)$$



Εικόνα 2.5 Γραφική παράσταση της γραμμικής συνάρτησης μεταφοράς.

Η βηματική συνάρτηση μεταφοράς εκφράζεται ως:

$$f(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2.4)$$



Εικόνα 2.6 Γραφική παράσταση της βηματικής συνάρτησης.

2.2.4 Δομικά στοιχεία Τεχνητού Νευρωνικού Δικτύου

Τα ουσιώδη συστατικά στοιχεία ενός μοντέλου Τεχνητού Νευρωνικού Δικτύου (ΤΝΔ) είναι τα εξής (Διαμαντάρας, 2007; Zhang et al., 1998):

- Ένα σύνολο μονάδων επεξεργασίας.
- Μια κατάσταση ενεργοποίησης για την κάθε μονάδα.
- Ένα πρότυπο σύνδεσης ανάμεσα στις μονάδες ή την τοπολογία του δικτύου
- Ένας κανόνας διάδοσης, ή μια συνάρτηση συνδυασμού, για να διαδίδει τις ενέργειες των μονάδων μέσω του δικτύου.
- Ένας κανόνας ενεργοποίησης για να αναπροσαρμόζει τις ενέργειες κάθε μονάδας χρησιμοποιώντας την τρέχουσα τιμή ενεργοποίησης και τις εισόδους που έρχονται από άλλες μονάδες
- Ένα εξωτερικό περιβάλλον που παρέχει πληροφορίες στο δίκτυο και/ή αλληλεπιδρά με αυτό.
- Ένας κανόνας μάθησης για να τροποποιεί το πρότυπο των συνδέσεων

χρησιμοποιώντας πληροφορίες από το εξωτερικό περιβάλλον.

- Μια συνάρτηση εξόδου για κάθε μονάδα.

2.2.5 Κατηγορίες Τεχνητών Νευρωνικών Δικτύων

Οι βασικές κατηγορίες των Τεχνητών Νευρωνικών Δικτύων σύμφωνα με την προσέγγιση μάθησης που ακολουθείται είναι (Haykin, 1994):

- Μάθηση με επίβλεψη (supervised learning) ή εποπτευόμενη μάθηση: το εξωτερικό περιβάλλον παρέχει τις επιθυμητές εξόδους για κάθε ένα από τα διανύσματα εισόδου μάθησης και το εξωτερικό περιβάλλον συμπεριφέρεται ως “δάσκαλος”.

- Μάθηση χωρίς επίβλεψη (unsupervised learning) ή μη-εποπτευόμενη μάθηση: το εξωτερικό περιβάλλον δεν παρέχει την επιθυμητή έξοδο του νευρωνικού δικτύου, ούτε την πληροφορία για το εάν αυτή είναι καλή ή κακή.

Η ενισχυτική μάθηση (reinforcement learning) θεωρείται μια ειδική περίπτωση της εποπτευόμενης μάθησης όπου το εξωτερικό περιβάλλον δίνει μόνο την πληροφορία ότι η έξοδος του δικτύου είναι καλή ή κακή, αντί να δίνει τη σωστή έξοδο. Στην ενισχυτική μάθηση, το εξωτερικό περιβάλλον συμπεριφέρεται ως “κριτής”. Κάποιοι ειδικοί κατατάσσουν την ενισχυτική μάθηση σε μία τρίτη κατηγορία μάθησης και δεν την αντιμετωπίζουν ως μία υποπερίπτωση της εποπτευόμενης μάθησης.

Επίσης, ανάλογα με την τοπολογία τους τα Τεχνητά Νευρωνικά Δίκτυα διακρίνονται σε:

- Εμπρόσθιας τροφοδότησης, ή αλλιώς προσο-τροφοδοτούμενα ΤΝΔ ή εμπρόσθιας διάδοσης (Feed-Forward Network – FFNN) νευρωνικό δίκτυο.
- Τεχνητά νευρωνικά δίκτυα ανατροφοδότησης (Feedback ή Recurrent Neural Networks – RNN).

Στην πρώτη κατηγορία οι μονάδες στέλνουν τις εξόδους τους σε μονάδες από τις οποίες δεν λαμβάνουν ανατροφοδότηση αμέσως ή εμμέσως (μέσω άλλων μονάδων) δηλαδή, δεν υπάρχουν βρόγχοι ανάδρασης σε αντίθεση με την δεύτερη κατηγορία.

2.2.6 Κανόνες ανανέωσης βαρών των νευρωνικών δικτύων

2.2.6.1 Ο αρχικός κανόνας του Hebb.

Ο Hebb (1949), διατύπωσε την υπόθεση που αναφέρεται στο βιβλίο του “Η οργάνωση της συμπεριφοράς” (“The organization of Behavior”), που στηρίχθηκε στην φυσική παρατήρηση δύο κοντινών κυττάρων και στο πως η αποδοτικότητα του ενός ‘πυροδοτεί’ το άλλο. ‘Ο νόμος του Hebb μπορεί να εκφραστεί μαθηματικά ως εξής:

$$\Delta w_{ij} = \eta y_i(x) x_j \quad (2.5)$$

Όπου x είναι το διάνυσμα εισόδου και εκφράζεται ως:

$$x = [x_1 \ x_2 \ \dots \ x_p]^T \quad (2.6)$$

και y είναι το διάνυσμα εξόδου που εκφράζεται ως:

$$y = [y_1 \ y_2 \ \dots \ y_p]^T \quad (2.7)$$

Το $\eta > 0$, ονομάζεται ρυθμός μάθησης, ενώ τα p και q είναι αντίστοιχα ο αριθμός των μονάδων εισόδου και εξόδου.

2.2.6.2 Ο τροποποιημένος κανόνας του Hebb

Από την προηγούμενη σχέση που εκφράζει τον αρχικό κανόνα του Hebb (1949), παρατηρούμε ότι, κάτω από ορισμένες συνθήκες το υπό ανανέωση βάρος, μετά

από την εκτέλεση ενός αριθμού επαναλήψεων, λαμβάνει πολύ μεγάλες τιμές με αποτέλεσμα η δράση του νευρώνα οδηγείται σε κορεσμό. Για να αποφευχθεί η κατάσταση αυτή, εισήχθη ένας παράγοντας λήθης α και ο κανόνας ανανέωσης των βαρών εκφράζεται πλέον ως εξής (Haykin, 1994):

$$\Delta w_{ij} = \eta y_i(x)x_j - \alpha y_i(x)w_{ij} \quad (2.8)$$

2.2.6.3 Ο κανόνας Δέλτα

Ο κανόνας Δέλτα είναι ο αλγόριθμος των ελαχίστων μέσων τετραγώνων (Least Mean Square - LMS) και επίσης ο κανόνας διόρθωσης σφάλματος. Προτάθηκε για την εκπαίδευση του δικτύου ADALINE/MADALINE, από τους Widrow και Hoff (1960), και είναι γνωστός και ως κανόνας Widrow - Hoff. Η βασική ιδέα του κανόνα Δέλτα είναι να αλλάζει, τα βάρη του δικτύου μειώνοντας το τετραγωνικό σφάλμα της εξόδου E_{pat} για κάθε πρότυπο εισόδου/επιθυμητής εξόδου από το σετ εκπαίδευσης. Το τετραγωνικό σφάλμα της εξόδου E_{pat} ορίζεται ως:

$$E^{pat} = \sum_{i=1}^q E_i^{pat} = \sum_{i=1}^q \frac{1}{2} [D_i - Y_i]^2 \quad (2.9)$$

Ο κανόνας Δέλτα είναι μια διαδικασία αναζήτησης κατηφορικής κλίσης (gradient-descent search procedure) που εκτελείται σε κάθε επανάληψη. Η αλλαγή του βάρους γίνεται σε κατεύθυνση αντίθετη με αυτήν που έχει η παράγωγος ως προς το βάρος της επιφάνειας του σφάλματος, με αποτέλεσμα την απότομη κατάβαση του σφάλματος προς το ελάχιστο σημείο του. Με τον τρόπο αυτό ελαχιστοποιείται το "μέσο" σφάλμα εξόδου E_{av} :

$$E^{av} = \frac{1}{M} \sum_{pat=1}^M E^{pat} \quad (2.10)$$

$$DW_{IJ} = -\eta \frac{\partial E^{pat}}{\partial W_{ij}} = -\eta \frac{\partial E^{pat}}{\partial Y_I} \frac{\partial Y_I}{\partial W_{IJ}} = \eta [D_i - Y_i] \frac{\partial Y_I}{\partial W_{IJ}} \quad (2.11)$$

Όπου Y η έξοδος του δικτύου και D το διάνυσμα της επιθυμητής εξόδου.

Οι Widrow and Hoff (1960), πρότειναν να χρησιμοποιηθεί, μία γραμμική συνάρτηση ενεργοποίησης ή $Y = W * X + bias$. Έτσι η μάθηση είναι πιο γρήγορη, επειδή αλλάζει τα βάρη ακόμα και όταν η ταξινόμηση της εξόδου είναι σχεδόν σωστή, σε αντίθεση με τον κανόνα perceptron που ανανεώνει τα βάρη όταν υπάρχει ένα μεγάλο σφάλμα ταξινόμησης. Μια άλλη σημαντική διαφορά είναι η χρήση διπολικών εισόδων αντί για δυαδικές εισόδους, επειδή υπήρχε το πρόβλημα ότι στις δυαδικές εισόδους, όταν η είσοδος είναι 0, τα βάρη που συνδέονται με την είσοδο αυτή δεν αλλάζουν. Η απαίτηση για τις συναρτήσεις ενεργοποίησης ενός τεχνητού νευρωνικού δικτύου στο οποίο η ανανέωση των βαρών γίνεται με τη βοήθεια του κανόνα Δέλτα, αφορά το ότι πρέπει να είναι συνεχής και άρα διαφορίσιμες, αλλιώς δεν μπορεί να εφαρμοστεί η τελευταία εξίσωση.

2.2.7 Αλγόριθμος οπισθοδρόμης διάδοσης σφάλματος (backpropagation)

Η ανανέωση των βαρών στον αλγόριθμο οπισθοδρόμης διάδοσης σφάλματος (backpropagation) γίνεται με τον κανόνα Δέλτα των προηγούμενων επιπέδων, αλλά καθώς το σφάλμα διαδίδεται προς τα πίσω ο κανόνας καταλήγει σε πιο σύνθετες μορφές που προκύπτουν από τη χρήση του κανόνα της αλυσίδας στην παραγωγή πεπλεγμένων συναρτήσεων (Haykin, 1994).

Ο αλγόριθμος οπισθοδρόμης διάδοσης σφάλματος παρέχει μια διαδικασία με την οποία αλλάζουν τα βάρη σε ένα δίκτυο back-propagation για να ταξινομηθούν σωστά τα πρότυπα εισόδου. Η βάση για τον αλγόριθμο αυτό αλλαγής των βαρών είναι η μέθοδος κατάβασης βαθμίδας (gradient – descent).

Σε ζεύγος εισόδου - εξόδου, ο αλγόριθμος οπισθοδρόμης διάδοσης σφάλματος,

παρουσιάζει δυο φάσεις μετάδοσης δεδομένων την εμπρόσθια φάση και την φάση οπισθοδιάδοσης (backpropagation). Στην εμπρόσθια φάση, το πρότυπο εισόδου $x(k)$ μεταδίδεται μέσω του δικτύου, παράγοντας την τρέχουσα έξοδο $y(k)$ του δικτύου.

Στη φάση οπισθοδιάδοσης τα σήματα λάθους διαδίδονται προς τα πίσω. Με τη χρήση του κανόνα Δέλτα, τα βάρη των συνδέσεων μεταξύ του κρυμμένου επιπέδου και του επιπέδου εξόδου μεταβάλλονται σύμφωνα με τον τύπο (Haykin, 1994):

$$\Delta W_{iq} = -\eta \frac{\partial E}{\partial W_{iq}} \quad (2.12)$$

Όπου E η συνάρτηση σφάλματος. Για την αλλαγή των βαρών των συνδέσεων μεταξύ του επιπέδου εισόδου και του κρυμμένου επιπέδου, χρησιμοποιείται ξανά ο κανόνας της αλυσίδας, και έχουμε έτσι την αλλαγή του βάρους της σύνδεσης μεταξύ της j μονάδας του επιπέδου εισόδου και της q μονάδας του κρυμμένου επιπέδου:

$$\Delta v_{qj} = -\eta \frac{\partial E}{\partial v_{qj}} = -\eta \frac{\partial E}{\partial net_q} \frac{\partial net_q}{\partial v_{qj}} \quad (2.13)$$

ή

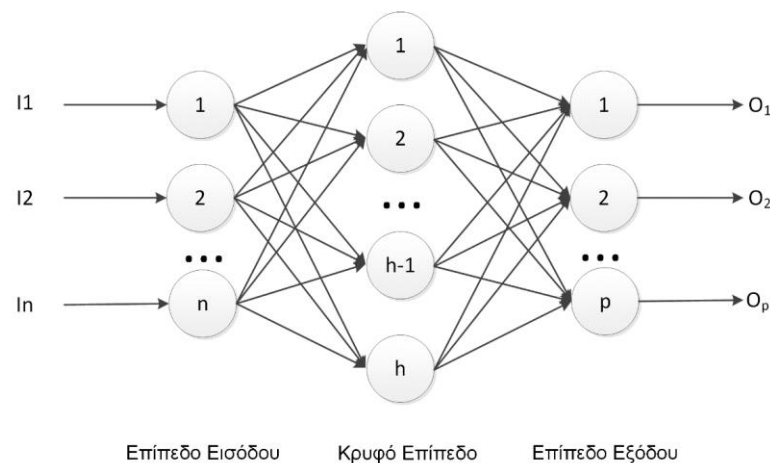
$$\Delta v_{qj} = -\eta \left[\frac{\partial E}{\partial y_i} \right] \left[\frac{\partial y_i}{\partial z_q} \right] \left[\frac{\partial z_q}{\partial net_q} \right] \left[\frac{\partial net_q}{\partial v_{qj}} \right] \quad (2.14)$$

Το σήμα σφάλματος μιας μονάδας στο επίπεδο αυτό είναι διαφορετικό από το σήμα σφάλματος μιας μονάδας στο εξωτερικό επίπεδο. Εξαιτίας της διαφοράς αυτής, η διαδικασία αλλαγής των βαρών λέγεται γενικευμένος κανόνας μάθησης Δέλτα (Generalized Delta Learning Rule - GDLR).

2.2.8 Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης

Σε ένα πολυεπίπεδο νευρωνικό δίκτυο πολλαπλών στρωμάτων, οι νευρώνες συνδέονται μόνο προς τα εμπρός. Τα επίπεδα του νευρωνικού δικτύου περιέχουν

συνδέσεις με τα επόμενα επίπεδα, αλλά δεν έχουν συνδέσεις προς τα πίσω (Hornik, 1991). Ο Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης (Feedforward Multilayer Perceptron (FFMLP)) εφαρμόστηκε σε αυτή την έρευνα, καθώς αρκετοί ερευνητές έχουν δείξει ότι είναι ο πιο κατάλληλος για προβλήματα που σχετίζονται με την πρόβλεψη χρονοσειρών (Tang and Fishwick, 1993; Zhang, 2001). Η τοπολογία ενός τυπικού νευρωνικού δικτύου εμπρόσθιας διάδοσης απεικονίζεται στην παρακάτω εικόνα, όπου n είναι ο αριθμός των νευρώνων στο επίπεδο εισόδου, h είναι ο αριθμός των νευρώνων στο κρυφό επίπεδο και ο p είναι ο αριθμός των νευρώνων στο επίπεδο εξόδου. I_1, I_2, \dots, I_n είναι οι τιμές εισόδου και O_1, O_2, \dots, O_p είναι οι τιμές (προβλεφθείσες) εξόδου.

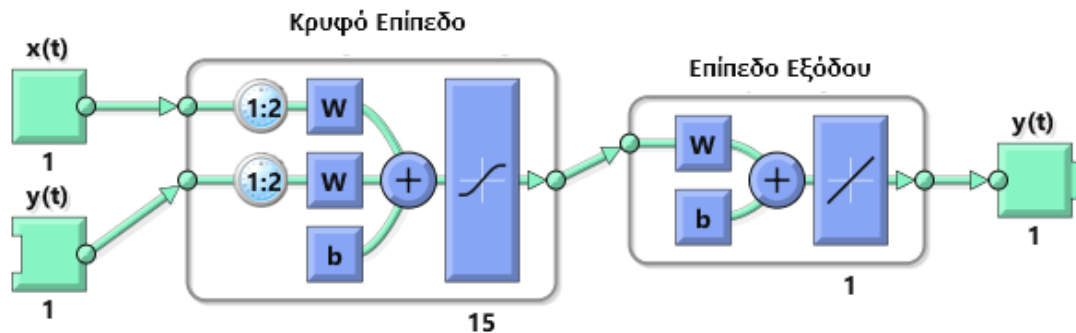


Εικόνα 2.7. Η τοπολογία ενός τυπικού πολυεπίπεδου νευρωνικού δικτύου εμπρόσθιας διάδοσης.

2.2.9 Νευρωνικά Δίκτυα Μη Γραμμικής Αυτοπαλινδρόμησης Εξωγενή Παράγοντα

Το Δίκτυο της Μη Γραμμικής Αυτοπαλινδρόμησης με Εξωγενή Παράγοντα (Nonlinear Autoregressive Network with Exogenous Input – NARX) αποτελεί ένα μοντέλο ανάλυσης χρονοσειρών που επιτρέπει την πρόβλεψη των μελλοντικών τιμών της σχετικής χρονοσειράς $y(t)$, χρησιμοποιώντας σαν εισόδους τις προηγούμενες τιμές της ίδιας χρονοσειράς $y(t)$ καθώς και τις ιστορικές τιμές μία εξωγενούς χρονοσειράς $x(t)$ άλλου παράγοντα εξωγενούς.

Στην επόμενη εικόνα, αναπαρίσταται σε απλοποιημένη μορφή ένα τέτοιο νευρωνικό δίκτυο μη γραμμικής αυτοπαλινδρόμησης με εξωγενή παράγοντα $x(t)$, όπου ο αριθμός καθυστερήσεων (delays) είναι δύο για την χρονοσειρά $y(t)$ και την χρονοσειρά $x(t)$ του εξωγενούς παράγοντα. Τα βάρη συμβολίζονται με w και η παράμετρος bias με b . Στο παράδειγμα αυτό, το νευρωνικό δίκτυο αποτελείται από ένα κρυφό επίπεδο με 15 νευρώνες και σιγμοειδή συνάρτηση μεταφοράς και ένα επίπεδο εξόδου.



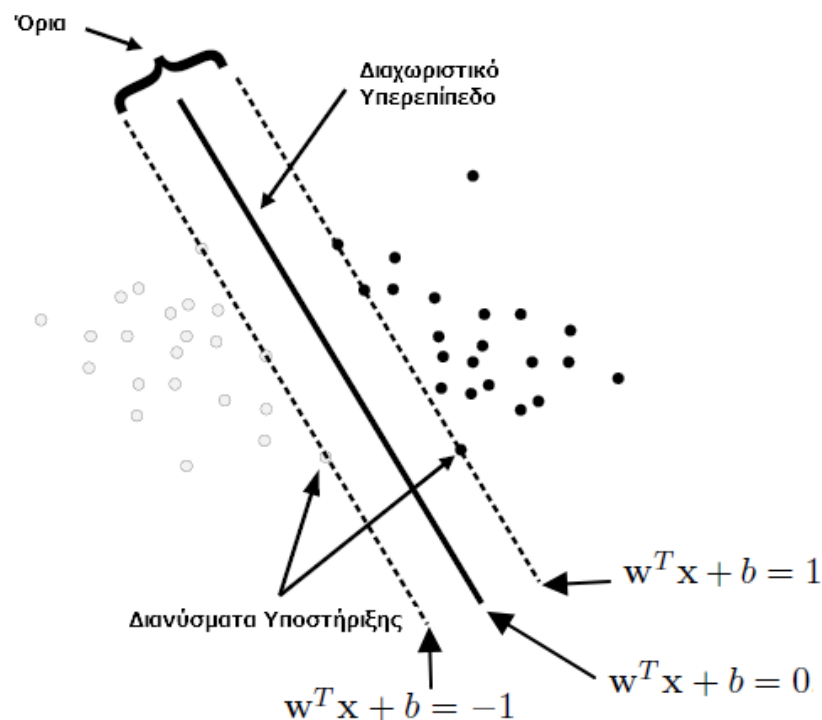
Εικόνα 2.8 Παράδειγμα τοπολογίας νευρωνικού μοντέλου Μη Γραμμικής Αυτοπαλινδρόμησης με Εξωγενή Παράγοντα.

2.3 Μηχανές Διανυσμάτων Υποστήριξης

2.3.1 Γενικά

Οι Μηχανές Διανυσμάτων Υποστήριξης (SVM) ανήκουν στις τεχνικές της Τεχνητής Νοημοσύνης. Είναι ένα σύνολο αλγορίθμων μηχανικής μάθησης που εφαρμόζονται στην ταξινόμηση και την παλινδρόμηση. Η θεωρία του Μηχανών Διανυσμάτων Υποστήριξης προτάθηκε σχετικά πρόσφατα (Vapnik, 1998) και βασίζονται στον αλγόριθμο νευρωνικών δικτύων τύπου Perceptron (Βλαχάβας κ.α., 2006). Οι μηχανές διανυσμάτων υποστήριξης θεωρούνται από τους πιο επιτυχημένους αλγόριθμους ταξινόμησης.

Οι Μηχανές Διανυσμάτων Υποστήριξης χρησιμοποιούνται κυρίως για τον εντοπισμό του βέλτιστου διαχωριστικού περιθωρίου για τον διαχωρισμό κατηγοριών σε ένα σύνολο δεδομένων. Τα δείγματα που βρίσκονται στο μεγιστοποιημένο περιθώριο ονομάζονται διανύσματα υποστήριξης και το επίπεδο διχοτόμησης του περιθωρίου ονομάζεται διαχωριστικό υπερεπίπεδο. Τα υπερεπίπεδα επιλέγονται έτσι ώστε να μεγιστοποιηθεί η απόσταση μεταξύ των τάξεων που θα ταξινομηθούν. Η εικόνα 50 απεικονίζει την γενική δομή των Μηχανών Διανυσμάτων Υποστήριξης.



Εικόνα 2.9 Απλοποιημένη δομική αναπαράσταση Μηχανών Διανυσμάτων Υποστήριξης.

Με την χρήση ενός συνόλου εκπαίδευσης ως εισόδου και του αλγόριθμου SVM, οι Μηχανές Διανυσμάτων Υποστήριξης (SVM) δομούν ένα μοντέλο. Το μοντέλο που κατασκευάστηκε μπορεί να χρησιμοποιηθεί για να προβλέψει αν ένα νέο στιγμιότυπο εισόδου ανήκει σε μια κατηγορία ή σε μία άλλη για να το ταξινομήσει κατάλληλα.

Σκοπός του αλγορίθμου SVM αποτελεί η εύρεση του βέλτιστου διαχωριστικού υπερεπίπεδου (hyperplane) για τον διαχωρισμό των κλάσεων. Η συνάρτηση απόφασης

στον αλγόριθμο Μηχανών Διανυσμάτων Υποστήριξης για μία νέα παρατήρηση x_i υπολογίζεται από την εξίσωση είναι η παρακάτω:

$$f(x_i) = \text{signal}(w^T x_i + b) \quad (2.15)$$

Όπου x είναι το διάνυσμα εισόδου, w^T το διάνυσμα βαρών και η σταθερά b (bias) που εκφράζει την μεροληψία. Η συνάρτηση απόφασης $f(x)$ θα έχει την τιμή $+1$ για όλα τα διανύσματα εισόδου x πάνω από το όριο και την τιμή -1 για όλα τα διανύσματα εισόδου x κάτω από το όριο, όπως αναπαρίσταται στην παραπάνω εικόνα.

Στο διαχωριστικό υπερεπίπεδο θα ισχύει η εξίσωση:

$$w^T x + b = 0 \quad (2.16)$$

Στην ταξινόμηση κλάσεων, οι Μηχανές Διανυσμάτων Υποστήριξης προσπαθούν να βρουν ένα υπερ-επίπεδο που να διαχωρίζει στο χώρο των παραδειγμάτων τα αρνητικά από τα θετικά. Η υπερεπιφάνεια που επιλέγεται πρέπει να απέχει όσο το δυνατόν περισσότερο από τα πλησιέστερα θετικά και αρνητικά παραδείγματα (Βλαχάβας κ.α., 2006).

2.3.2 Πυρήνες Μηχανών Διανυσμάτων Υποστήριξης

Υπάρχουν διαφόρων ειδών πυρήνων που χρησιμοποιούνται στις Μηχανές Διανυσμάτων Υποστήριξης (SVM). Η συνάρτηση του πυρήνα μπορεί να θεωρηθεί ως μέτρο ομοιότητας μεταξύ των δειγμάτων x_i και x_j η οποία εφαρμόζεται στους ταξινομητές SVM για να εκτελούν διαχωρισμούς μεταξύ των κλάσεων ακόμη και όταν υπάρχουν πολύπλοκα όρια (Suykens and Vandewalle, 1999). Οι κυριότεροι πυρήνες Μηχανών Διανυσμάτων Υποστήριξης, είναι: ο γραμμικός πυρήνας, ο πολυωνυμικός, και ο ακτινικός.

Η συνάρτηση για τον γραμμικό πυρήνα (Linear Kernel) εκφράζεται από την σχέση:

$$K(x_i, x_j) = \langle x_i, x_j \rangle \quad (2.17)$$

Η συνάρτηση για τον πολυωνυμικό πυρήνα (Polynomial Kernel) εκφράζεται από την σχέση:

$$K(x_i, x_j) = (\sigma(\gamma) * x_i^T x_j + r)^d, \quad \sigma > 0 \quad (2.18)$$

Όπου d αναπαριστά τον βαθμό του πολυωνύμου (degree), το r την σταθερή παράμετρο, η συνάρτηση σ (σίγμα) εξαρτάται από την παράμετρο γ (γάμμα) που ελέγχει το πλάτος του πυρήνα. Η σχέση του γ με το σ εκφράζεται σύμφωνα με την εξίσωση:

$$\gamma = \frac{1}{\sigma^2} \quad (2.19)$$

Οι ακτινικοί πυρήνες RBF (Radial Basis Function) είναι τύποι πυρήνων όπου ένα μέτρο απόστασης εξομαλύνεται από μια ακτινική (εκθετική συνάρτηση). Αυτός ο πυρήνας μπορεί να εφαρμοστεί στην ταξινόμηση για μη γραμμικά δείγματα υψηλότερων διαστάσεων, σε αντίθεση με τον γραμμικό πυρήνα. Στην ακτινωτή συνάρτηση πυρήνα εφαρμόζεται παρακάτω εξίσωση:

$$K(x_i, x_j) = e^{-\sigma(\gamma)\|x_i - x_j\|^2}, \quad \sigma > 0 \quad (2.20)$$

2.4 Τυχαία Δάση Δέντρων Απόφασης

2.4.1 Γενικά

Μία σχετικά νέα μέθοδος μηχανικής μάθησης για την ταξινόμηση και την παλινδρόμηση είναι τα «Τυχαία Δάση» (Random Forests). Τα Τυχαία Δάση είναι μία συλλογή δέντρων απόφασης και προτάθηκαν σχετικά πρόσφατα από τον Leo Breiman

(1996). Η τεχνική Bagging (Ενσάκισης) Δέντρων χρησιμοποιείται στα Τυχαία Δάση και θα αναπτυχθεί στην επόμενη ενότητα.

Η τεχνική μηχανικής μάθησης των Τυχαίων Δασών είναι μια ομάδα μεγάλου αριθμού ταξινομητών - δέντρων αποφάσεων. Κάθε δέντρο αποφάσεως αναπτύσσεται ανάλογα με την τυχαία δειγματοληψία με αντικατάσταση των αρχικών δεδομένων εκπαίδευσης. Η κλάση μιας παρατήρησης που θα προβλεφθεί στην ταξινόμηση, καθορίζεται από την πλειοψηφία των ψήφων από το σύνολο των δέντρων αποφάσεων που αναπτύσσονται στο τυχαίο δάσος (Ghosh and Joshi, 2014).

Ο αλγόριθμος των τυχαίων δασών αποτελείται από δύο στάδια. Συγκεκριμένα, στο πρώτο στάδιο ακολουθούνται τα παρακάτω βήματα, για n δείγματα, m χαρακτηριστικών ή μεταβλητών για την δημιουργία τυχαίου δάσους, ενώ στο δεύτερο στάδιο πραγματοποιούνται οι προβλέψεις βάσει του δημιουργηθέντος μοντέλου (Liaw and Wiener, 2002). Συνοπτικά τα στάδια του αλγορίθμου περιγράφονται παρακάτω:

1ο στάδιο

1. Επέλεξε τυχαία χαρακτηριστικά K από τα συνολικά χαρακτηριστικά m όπου $k < m$
2. Μεταξύ των χαρακτηριστικών K , υπολόγισε τον κόμβο d χρησιμοποιώντας το καλύτερο σημείο διαίρεσης
3. Διαχώρισε τον κόμβο σε θυγατρικούς κόμβους χρησιμοποιώντας τον καλύτερο χωρισμό.
4. Επαναλάβετε τα βήματα 1 έως 3 έως ότου ο αριθμός των κόμβων γίνει "1"
5. Δημιούργησε το δάσος επαναλαμβάνοντας τα βήματα 1 έως 4 για " n " φορές για την δημιουργία " n " αριθμού δέντρων.

2ο στάδιο

1. Υπολόγισε τις ψήφους για κάθε στόχο που προβλέφθηκε.
2. Θεώρησε τον ψήφο της πλειοψηφίας (majority voting) του στόχου που προβλέφθηκε ως την τελική πρόβλεψη από τον αλγόριθμο τυχαίων δασών.

2.4.2 Τεχνική Bagging στα Τυχαία Δάση

Η τεχνική bagging προτάθηκε από τον Leo Breiman το 1996 για να βελτιώσει την ταξινόμηση συνδυάζοντας τις ταξινομήσεις τυχαία παραγόμενων σετ εκπαίδευσης (Breiman 1996). Συγκαταλέγεται στις μεθόδους μηχανικής μάθησης (machine learning) που αποτελεί κλάδο της τεχνητής νοημοσύνης.

Η τεχνική bagging (ενσάκιση) ενός ταξινομητή χρησιμοποιείται για τη μείωση της διακύμανσης. Η τεχνική Bagging δημιουργεί ένα σύνολο ταξινομητών με επαναδειγματοληψία. Η ονομασία Bagging προήλθε από την συνένωση των όρων Bootstrap και Aggregating. Η μεθοδολογία του Bootstrap Aggregating βελτιώνει την ικανότητα πρόβλεψης των δέντρων απόφασης. Η Bootstrap μεθοδολογία δημιουργεί παρατηρήσεις, με την εκπαίδευση του κάθε ταξινομητή σε τυχαία ανακατανομή των μετρήσεων του συνόλου εκπαίδευσης (training set). Κατά συνέπεια ενσωματώνει τα οφέλη τόσο της bootstrap όσο και της aggregating προσέγγισης.

Ένα bootstrap δείγμα παράγεται με δειγματοληψία με επανάθεση. Στη δειγματοληψία με επανατοποθέτηση (Bootstrap) το αρχικό σύνολο δεδομένων S δειγματοληπτείται τόσες φορές όσες είναι και οι μετρήσεις του. Σε κάθε δειγματοληψία επιλέγεται ένα δείγμα που επανατοποθετείται στο σύνολο εκπαίδευσης (training set), χωρίς να αφαιρείται από το αρχικό σύνολο δεδομένων. Έτσι παράγεται ένα σύνολο εκπαίδευσης S_{training} , με έναν αριθμό δειγμάτων που είναι ίσο με το αρχικό σύνολο περιέχοντας πολλαπλά αντίτυπα ανακατανεμημένα του ίδιου δείγματος. Το σύνολο ελέγχου (validation set) αποτελεί η διαφορά των δύο συνόλων $S - S_{\text{training}}$ (Pino-Mejías et al., 2008 Prasad et al., 2006).

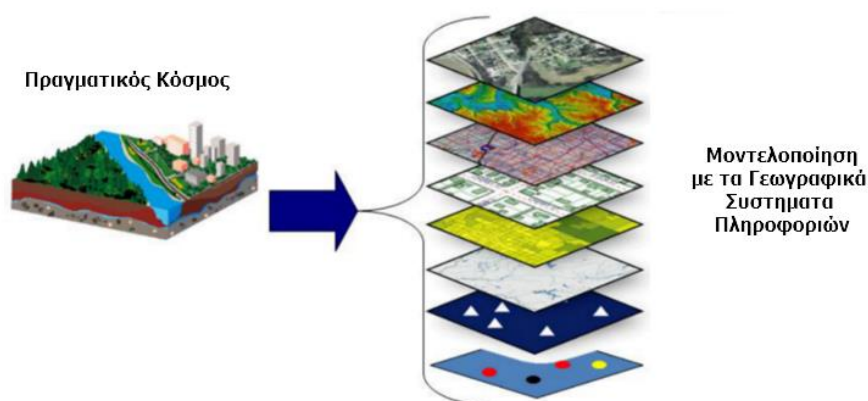
Η προσέγγιση aggregating αποτελεί τον συνδυασμό πολλαπλών μοντέλων έτσι ώστε για κάθε bootstrap σετ, προσαρμόζεται ένα μοντέλο που παράγει μια πρόβλεψη. Η μέθοδος bagging μπορεί να βελτιώσει τη μεταβλητότητα σχετικά ασταθών διαδικασιών όπως τα νευρωνικά δίκτυα και τα δέντρα αποφάσεων, έτσι ώστε να παράγεται μια βελτιωμένη πρόβλεψη, καθόσον ο μέσος όρος μειώνει τη μεταβλητότητα (Breiman 1996).

2.5 Γεωγραφικά Συστήματα Πληροφοριών

Τα Γεωγραφικά Πληροφοριακά Συστήματα (GIS) έχουν ραγδαία ανάπτυξη τις τελευταίες δεκαετίες, και κυρίως από τη δεκαετία του '80. Τα Γεωγραφικά Πληροφοριακά Συστήματα έχουν την ικανότητα να οπτικοποιούν, να επεξεργάζονται και να αναλύουν μεγάλες ποσότητες χωρικών δεδομένων (Αστάρας κ.α., 2011; Ευελπίδου και Αντωνίου, 2015; Κάβουρας κ.α., 2016; Χαλκιάς και Γκούσια, 2015; Κουτσόπουλος 2005; Κουτσόπουλος και Ανδρουλακάκης, 2012). Μία από τις βασικές τους δυνατότητες, είναι αφενός η δημιουργία μοντέλων, δομώντας ρεαλιστικά τον πραγματικό κόσμο, αφετέρου να προσφέρουν πολλές εναλλακτικές λύσεις σε πολλά προβλημάτων διαφόρων ειδών που σχετίζονται με χωρικά προβλήματα και χωρικούς αλγορίθμους (Antenucci et al., 1991).

Τα Γεωγραφικά Συστήματα Πληροφοριών (GIS) είναι μια οργανωμένη συλλογή μηχανικών υπολογιστικών συστημάτων (hardware), λογισμικού (software), χωρικών δεδομένων με σκοπό τη συλλογή, καταχώρηση, επεξεργασία, διαχείριση, ανάλυση και απόδοση κάθε μορφής χωρικής πληροφορίας που αφορά το γεωγραφικό περιβάλλον (Burrough, 1992; Elwood, 2011).

Τα Γεωγραφικά συστήματα πληροφοριών συνεισφέρουν στην μοντελοποίηση του πραγματικού κόσμου σε διάφορα επίπεδα (layers): εικόνων (imagery), ανύψωσης (elevation), μεταφορών (transportation), συνόρων (boundaries), διευθύνσεων (addresses), καθώς και χωρικά χαρακτηριστικά υδάτων (water features), κλπ. (Εικ. 1).



Εικόνα 2.10 Μοντελοποίηση του πραγματικού κόσμου σε επίπεδα στα GIS.

Η αλματώδης ανάπτυξη των Γεωγραφικών Συστημάτων Πληροφοριών (ΓΣΠ) έχει επηρεάσει σημαντικά τη δημόσια διοίκηση και τον χωροταξικό σχεδιασμό. Έχουν μελετηθεί από ερευνητές διάφορων ειδών συστήματα βασισμένα στα Γεωγραφικά Συστήματα Πληροφοριών (ΓΣΠ) με τη χρήση διαφόρων τεχνολογιών, για να χρησιμοποιηθούν στη δημόσια διοίκηση και τον χωρικό σχεδιασμό (Goddard et al., 2003; Kouziokas, 2016a; Ventura, 1995).

2.6 Τηλεπισκόπηση

Η τηλεπισκόπηση είναι μία επιστήμη, ένα σύνολο θεωρητικών και εφαρμοσμένων μεθόδων με σκοπό την ανάλυση της επιφάνειας της γης. Η βασική αρχή της τηλεπισκόπησης βασίζεται στην αποτύπωση αντικειμένων της επιφάνειας της γης με την καταγραφή της ηλεκτρομαγνητικής ακτινοβολίας σε διάφορα μήκη κύματος (Περάκης κ.α., 2015).

Τηλεπισκόπηση είναι η επιστήμη της καταγραφής, ανάλυσης και ερμηνείας πληροφοριών από απόσταση, με τη χρήση αισθητήρων (Περάκης, 1999). Ο κύριος στόχος της τηλεπισκόπησης είναι η εξαγωγή χρήσιμων δεδομένων για τα αντικείμενα που μελετώνται και το περιβάλλον τους και περιλαμβάνει τις παρατηρήσεις με δορυφορικές εικόνες και αεροφωτογραφίες. Η δορυφορική τηλεπισκόπηση αφορά την καταγραφή πληροφορίας από μετεωρολογικούς και περιβαλλοντικούς δορυφόρους, οι οποίοι φέρουν αισθητήρες που συλλέγουν δεδομένα σε διαφορετικά τμήματα του ηλεκτρομαγνητικού φάσματος (Καϊμάρης, 2014).

Η ταξινόμηση δορυφορικών εικόνων, που θα μελετηθεί στην παρούσα έρευνα, έχει χρησιμοποιηθεί σε διάφορες έρευνες στη δημόσια διοίκηση με σκοπό την διαχείριση των χρήσεων γης, των δασικών εκτάσεων, της βλάστησης, αλλά και την διαχείριση των υδάτων (Cheema and Bastiaanssen, 2010; He et al., 1998; Stefanov et al., 2001).

ΚΕΦΑΛΑΙΟ 3: Μεθοδολογία

3.1 Μεθοδολογία προσέγγισης

Η μεθοδολογία για την προσέγγιση του ερευνητικού προβλήματος, εφαρμόστηκε μελετώντας καταρχάς το θεωρητικό πλαίσιο του ερευνητικού προβλήματος και την επισκόπηση της σχετικής βιβλιογραφίας που αφορά την τεχνητή νοημοσύνη, τις μεθόδους πρόβλεψης παραγόντων σημαντικών στην δημόσια διοίκηση αλλά και μεθόδων χωρικής πρόβλεψης και ταξινόμησης δορυφορικών εικόνων βασισμένων σε μεθόδους τεχνητής νοημοσύνης.

Τα γεωγραφικά συστήματα πληροφοριών και η τηλεπισκόπηση μελετήθηκαν σε σχέση με την εφαρμογή της τεχνητής νοημοσύνης σε μεθόδους χωρικής ανάλυσης, χωρικής πρόβλεψης, χωρικής ομαδοποίησης και ταξινόμησης δορυφορικών εικόνων.

Μελετήθηκε επίσης, η ανάλυση δεδομένων εγκληματικότητας καθώς και η συμβολή τους στη λήψη αποφάσεων με τεχνικές μηχανικής μάθησης για την ποσοτική και χωρική πρόβλεψη της εγκληματικότητας. Από τις μεθόδους τεχνητής νοημοσύνης μελετήθηκαν τα τεχνητά νευρωνικά δίκτυα και οι εφαρμογές τους στην πρόβλεψη της εγκληματικότητας, καθώς και οι Μηχανές Διαनुσμάτων Υποστήριξης και τα Τυχαία Δάση Δέντρων Απόφασης.

Η μεθοδολογία προσέγγισης του ερευνητικού προβλήματος περιγράφεται από τα παρακάτω βασικά σημεία.

- Μελέτη του θεωρητικού πλαισίου του ερευνητικού προβλήματος και επισκόπηση της σχετικής βιβλιογραφίας που αφορά μεθόδους τεχνητής νοημοσύνης, τα γεωγραφικά συστήματα πληροφοριών και την τηλεπισκόπηση σε σχέση με

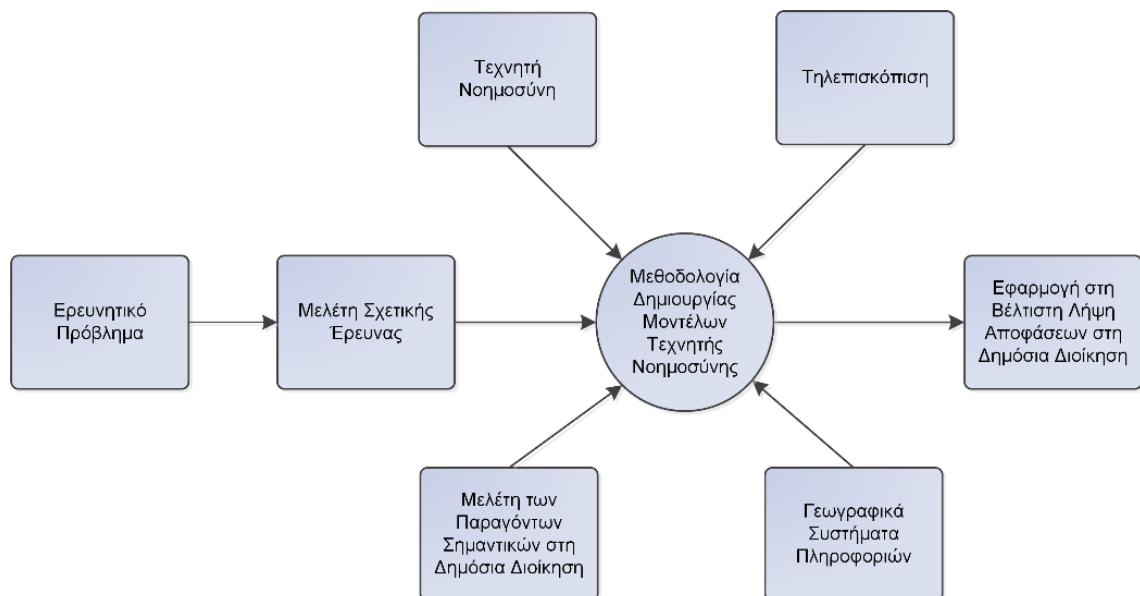
μεθόδους χωρικής ανάλυσης και ομαδοποίησης (clustering) και ανάλυσης δεδομένων εγκληματικότητας και η συμβολή τους στη λήψη αποφάσεων.

- Συστηματική διερεύνηση του προβλήματος, ο ρόλος της χρήσης της τεχνητής νοημοσύνης και της χωρικής πληροφορίας και στη λήψη αποφάσεων στη δημόσια και δικαστική διοίκηση.
- Μελέτη των σημαντικών παραγόντων που επηρεάζουν την λήψη αποφάσεων στην δημόσια διοίκηση, και μπορούν να αποτελέσουν πεδίο έρευνας με μεθοδολογίες τεχνητής νοημοσύνης, όπως η εγκληματικότητα (Αστυνομικές και Δικαστικές Αρχές, ασφάλεια συγκοινωνιών, ασφάλεια εκπαιδευτικών μονάδων, ασφάλεια τουριστικών περιοχών), παράγοντες υποβάθμισης και μόλυνσης του περιβάλλοντος (Περιβαλλοντική διοίκηση) και κοινωνικο-οικονομικοί παράγοντες όπως η ανεργία.
- Εφαρμογή και σύγκριση μεθόδων τεχνητής νοημοσύνης όπως τα τεχνητά νευρωνικά δίκτυα, οι Μηχανές Διαनुσμάτων Υποστήριξης, τα Τυχαία Δάση, σε πραγματικά δεδομένα και η εφαρμογές τους στην πρόβλεψη της εγκληματικότητας, και όλων των υπόλοιπων παραγόντων που αναφέρθηκαν παραπάνω, στην διοίκηση των δικαστικών και δημοσίων υπηρεσιών.
- Μελέτη της ανάπτυξης νέας μεθοδολογίας χωρικής πρόβλεψης της εγκληματικότητας βασισμένης στην τεχνητή νοημοσύνη και τα γεωγραφικά συστήματα πληροφοριών, και εφαρμογή της σε πραγματικά δεδομένα και σύγκριση της με άλλες έρευνες, δεδομένου ότι η χωρική πρόβλεψη θεωρείται πολύ σημαντική στη λήψη αποφάσεων σε πολλούς τομείς της δημόσιας διοίκησης. Έρευνα της χρήσης του δείκτη R (Clark και Evans) στην χωρική πρόβλεψη της εγκληματικότητας. Εφαρμογή των Τεχνητών Νευρωνικών Δικτύων στην πρόβλεψη της εγκληματικότητας για την εύρεση των περιοχών υψηλής συγκέντρωσης (crime hotspots) περιστατικών εγκλημάτων, που θα δώσει την κατευθυντήριες γραμμές στην διοίκηση των εμπλεκόμενων υπηρεσιών για την λήψη διαφόρων ειδών αποφάσεων.
- Μελέτη της ανάπτυξης νέας υβριδικής μεθοδολογίας στην ταξινόμηση δορυφορικών εικόνων, εφαρμογή της σε πραγματικά δεδομένα και σύγκριση των αποτελεσμάτων της με άλλες υπάρχουσες μεθόδους τεχνητής νοημοσύνης στην

τηλεπισκόπηση, δεδομένου ότι η ταξινόμηση δορυφορικών εικόνων θεωρείται σημαντική στη μελέτη των χρήσεων γης σε πολλούς τομείς της δημόσιας διοίκησης.

- Αποτίμηση της προτεινόμενης μεθοδολογίας λήψης αποφάσεων βάσει πολλαπλών τεχνικών τεχνητής νοημοσύνης και σύγκρισή της με την υπάρχουσα βιβλιογραφία.

Μία επισκόπηση της γενικής μεθοδολογίας προσέγγισης του προβλήματος που ακολουθήθηκε αναπαρίσταται στην επόμενη εικόνα.



Εικόνα 3.1 Επισκόπηση της γενικής μεθοδολογίας προσέγγισης του ερευνητικού προβλήματος.

3.2 Σχετική έρευνα

Διάφοροι ερευνητές χρησιμοποίησαν τα Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ) για να προβλέψουν τα ποσοστά της εγκληματικότητας (Alwee et al., 2013; Bertiz Armenico et al., 2015; Bogomolov et al., 2014). Σε μία έρευνα οι Alwee et al. (2013), χρησιμοποίησαν τα Τεχνητά Νευρωνικά Δίκτυα σε συνδυασμό με μεθόδους παλινδρόμησης για την

πρόβλεψη των ποσοστών εγκληματικότητας, λαμβάνοντας υπόψη διάφορους κοινωνικοοικονομικούς παράγοντες όπως το ακαθάριστο εγχώριο προϊόν, το ποσοστό ανεργίας και το δείκτη τιμών καταναλωτή. Τα αποτελέσματα έδειξαν ότι το προτεινόμενο μοντέλο μπορεί να προβλέψει τα ποσοστά εγκληματικότητας ιδιοκτησίας με καλή ακρίβεια. Σε μια άλλη μελέτη οι Bertiz Armenico et al. (2015), ανέπτυξαν ένα μαθηματικό μοντέλο που βασίζεται στην ανάλυση χρονοσειρών χρησιμοποιώντας Τεχνητά Νευρωνικά Δίκτυα και τεχνικές εξόρυξης δεδομένων για την πρόβλεψη των ποσοστών της εγκληματικότητας. Τα αποτελέσματα έδειξαν μεγάλη ακρίβεια προβλέψεων. Επίσης, σε μια άλλη έρευνα οι Bogomolov et al. (2014) χρησιμοποίησαν τεχνητά νευρωνικά δίκτυα σε συνδυασμό με δημογραφικά δεδομένα για να προβλέψουν τα ποσοστά εγκληματικότητας στην πόλη του Λονδίνου. Τα αποτελέσματα έδειξαν πολύ καλή ακρίβεια πρόβλεψης.

Επίσης, διάφοροι ερευνητές χρησιμοποίησαν τα Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ) στο περιβαλλοντικό μάνατζμεντ για να προβλέψουν διάφορους περιβαλλοντικούς παράγοντες όπως παράδειγμα τις συγκεντρώσεις των ατμοσφαιρικών ρύπων: όζον και διοξείδιο του αζώτου, που μπορούν να διαδραματίσουν σημαντικό ρόλο στη δημόσια διοίκηση, αλλά και στην περιβαλλοντική διαχείριση και τον σχεδιασμό, προωθώντας βιώσιμες στρατηγικές λήψης αποφάσεων στη διαχείριση του περιβάλλοντος (Cortés et al., 2000; Graf-Jaccottet and Jaunin, 1998; Kanevski et al., 2004). Μια σχετική έρευνα (Graf-Jaccottet and Jaunin, 1998), μελέτησε την εφαρμογή προβλεπτικών μεθόδων βασισμένων σε μοντέλα αυτοπαλινδρόμησης για την πρόβλεψη των τιμών του όζοντος και του διοξειδίου του αζώτου για χρονικό διάστημα πέντε ετών σε ημερήσια βάση. Τα αποτελέσματα έδειξαν καλή ακρίβεια στην πρόβλεψη των τιμών. Μια άλλη σχετική έρευνα (Cortés et al., 2000), έδειξε ότι η τεχνητή νοημοσύνη μπορεί να εφαρμοστεί στα συστήματα υποστήριξης περιβαλλοντικών αποφάσεων για την επίλυση διαφόρων ειδών περιβαλλοντικών ζητημάτων. Η έρευνα έδειξε ότι η ενσωμάτωση διαφόρων μεθόδων τεχνητής νοημοσύνης με στατιστικά μοντέλα σε ένα ενιαίο σύστημα παρέχει μεγαλύτερη ακρίβεια και αξιοπιστία. Μια άλλη έρευνα (Kanevski et al., 2004), χρησιμοποίησε αλγορίθμους τεχνητών νευρωνικών δικτύων και μεθόδους γεωστατιστικής

επικεντρώθηκε στα ζητήματα που έχουν σχέση με την λήψη αποφάσεων και μελέτησε την εξόρυξη και την μοντελοποίηση δεδομένων.

Επίσης, η ταξινόμηση δορυφορικών εικόνων έχει χρησιμοποιηθεί σε διάφορες έρευνες στη δημόσια διοίκηση με σκοπό την διαχείριση των χρήσεων γης, των δασικών εκτάσεων, αλλά και την διαχείριση των υδάτων (Cheema and Bastiaanssen, 2010; He et al., 1998; Stefanov et al., 2001). Οι Stefanov et al. (2001), εφάρμοσαν τεχνικές ταξινόμησης καλύψεων-χρήσεων γης για την παρακολούθηση των ξηρών αστικών καλύψεων γης. Επίσης, οι He et al. (1998), μελέτησαν την εφαρμογή ταξινομημένων δορυφορικών εικόνων και των γεωγραφικών συστημάτων πληροφοριών στην αξιολόγηση αστικών περιοχών κάλυψης πρασίνου. Οι δορυφορικές εικόνες ταξινομήθηκαν ώστε να μοντελοποιηθούν οι περιοχές με δασικές εκτάσεις. Επίσης, οι Cheema and Bastiaanssen (2010), ερεύνησαν την ταξινόμηση χρήσεων και καλύψεων γης χρησιμοποιώντας δορυφορικά δεδομένα για την υποστήριξη της διαχείρισης των υδάτων.

3.3 Η παρούσα έρευνα

Στην παρούσα έρευνα, ερευνήθηκε η δημιουργία τεχνητών νευρώνων με την κατάλληλη τοπολογία ώστε να είναι δυνατή η πρόβλεψη της εγκληματικότητας και των περιοχών υψηλής συγκέντρωσης της (hotspots). Για τον συγκεκριμένο σκοπό διερευνήθηκε η δημιουργία μοντέλων νευρωνικών δικτύων με την επιλογή των κατάλληλων παραμέτρων χρησιμοποιώντας τον αλγόριθμο αναζήτησης πλέγματος (Grid Search) και τον αλγόριθμο Βελτιστοποίησης Σμήνους Σωματιδίων (Particle Swarm Optimization, PSO). Οι παράμετροι που ερευνήθηκαν ήταν: τα συναπτικά βάρη του νευρωνικού δικτύου, η τοπολογία των νευρώνων, οι συναρτήσεις ενεργοποίησης (activation functions), η αρχικοποίηση των βαρών, οι αλγόριθμοι εκπαίδευσης (training algorithms), η δομή τους (πολυεπίπεδοι αισθητήρες (multilayer perceptrons), κλπ), τα σφάλματα νευρωνικών δικτύων και ο υπολογισμός τους και η ελαχιστοποίησή τους για την βελτιστοποίηση της προβλεψιμότητας των νευρωνικών δικτύων, κλπ. Η έρευνα επικεντρώθηκε στην κατασκευή ενός πολυεπίπεδου αισθητήρα εμπρόσθιας διάδοσης

(Feedforward Multilayer Perceptron), ο οποίος βρέθηκε ότι θα έχει καλύτερα αποτελέσματα σε προβλέψεις χρονικές που αφορούν την εγκληματικότητα, ο οποίος συγκρίθηκε με άλλες μεθόδους τεχνητής νοημοσύνης όπως οι Μηχανές Διανυσμάτων Υποστήριξης.

Η Τεχνητή Νοημοσύνη (TN) χρησιμοποιήθηκε σε αυτή την έρευνα για την δημιουργία και σύγκριση μοντέλων πρόβλεψης τεχνητών νευρωνικών δικτύων σχετικά με την πρόβλεψη διαφόρων παραγόντων, σημαντικών στην λήψη αποφάσεων σε δημόσιες και δικαστικές υπηρεσίες, όπως: των ποσοστών της εγκληματικότητας, των τιμών περιβαλλοντικών ρύπων, κοινωνικοοικονομικών παραγόντων (πχ ανεργία, κατανάλωση ενέργειας), κλπ. Η έρευνα της εφαρμογής της τεχνητής νοημοσύνης (τεχνητά νευρωνικά δίκτυα) στην πρόβλεψη της εγκληματικότητας θα βοηθήσει σημαντικά στην λήψη αποφάσεων για την καταπολέμησή της. Επίσης, μελετήθηκε η έρευνα της εφαρμογής της τεχνητής νοημοσύνης στην πρόβλεψη περιβαλλοντικών και άλλης φύσεως δεδομένων πολύ σημαντικών στην λήψη αποφάσεων στην δημόσια διοίκηση. Στην παρούσα έρευνα επιχειρήθηκε η βελτίωση της ακρίβειας προβλέψεων σε σχέση με τις άλλες έρευνες με την χρησιμοποίηση διαφορετικών παραγόντων εισόδου στο νευρωνικό μοντέλο πρόβλεψης, αλλά και η έρευνα διαφόρων τοπολογιών νευρώνων και συναρτήσεων μεταφοράς αλλά και αλγορίθμων εκπαίδευσης. Τα αποτελέσματα και η συζήτηση παρουσιάζονται στα επόμενα κεφάλαια.

Επίσης, στην παρούσα διατριβή, τα γεωγραφικά συστήματα πληροφοριών χρησιμοποιήθηκαν, εφαρμόζοντας μεθόδους χωρικής ανάλυσης και ομαδοποίησης (clustering) και ανάλυσης δεδομένων εγκληματικότητας. Επιπλέον, διερευνήθηκε η συμβολή τους στη λήψη αποφάσεων σε δημόσιες και δικαστικές υπηρεσίες σε θέματα χωρικής φύσεως σε συνδυασμό με μεθοδολογίες τεχνητής νοημοσύνης. Η ενσωμάτωση των γεωγραφικών συστημάτων πληροφοριών σε συστήματα λήψης αποφάσεων θεωρείται πολύ σημαντική στην λήψη αποφάσεων για πολύπλοκα ζητήματα χωρικής φύσεως στην δημόσια και δικαστική διοίκηση. Το πλεονέκτημα της μεθοδολογίας που αναπτύχθηκε σε σύγκριση με άλλες έρευνες (Caplan and Kennedy, 2016; Chainey and Ratcliffe, 2013; Gorr and Lee, 2015; Ki, 2013; Ku and Leroy, 2014; Leipnik and Albert, 2003; Palocsay et al., 2000), είναι ότι οι δημόσιοι φορείς μπορούν να λάβουν υπόψη και

να αξιοποιήσουν τις χωρικές προβλέψεις σε μεγαλύτερο βαθμό στην δημόσια και δικαστική διοίκηση, αλλά και τον αστικό χωρικό σχεδιασμό.

Επιπλέον, στην παρούσα διατριβή, ταξινομήθηκαν δεδομένα τηλεπισκόπησης και συγκεκριμένα, δεδομένα εικονοστοιχείων δορυφορικής εικόνας με την χρήση μεθόδων τεχνητής νοημοσύνης: Τεχνητά Νευρωνικά Δίκτυα, Μηχανές Διανυσμάτων Υποστήριξης και προτάθηκαν δύο νέες υβριδικές μέθοδοι ταξινόμησης: ένας νέος υβριδικός πυρήνας Μηχανών Διανυσμάτων Υποστήριξης και Τυχαίων Δασών καθώς και ένα νέο υβριδικό μοντέλο Νευρωνικών Δικτύων και Τυχαίων Δασών. Τα νέα υβριδικά μοντέλα έδειξαν καλύτερα αποτελέσματα ταξινόμησης έναντι των συμβατικών μοντέλων τεχνητής νοημοσύνης.

Τα δεδομένα επεξεργάστηκαν και προετοιμάστηκαν για να τροφοδοτήσουν τα μοντέλα νευρωνικών δικτύων και άλλων τεχνικών τεχνητής νοημοσύνης, ώστε ερευνώντας διάφορες τοπολογίες νευρώνων, κρυφών επιπέδων, συναρτήσεων ενεργοποίησης (νευρωνικά δίκτυα), συναρτήσεων πυρήνα (Μηχανές Διανυσμάτων Υποστήριξης), να δομηθεί το βέλτιστο μοντέλο πρόβλεψης των επιθυμητών παραγόντων βασισμένο σε τεχνικές τεχνητής νοημοσύνης. Στην παρούσα έρευνα, για την ταξινόμηση της πολυφασματικής δορυφορικής εικόνας, επιλέχθηκαν οι αλγόριθμοι των Τεχνητών Νευρωνικών Δικτύων, των Μηχανών Διανυσμάτων Υποστήριξης και των Τυχαίων δασών δεδομένου ότι οι πιο συνηθισμένες μεθοδολογίες όπως ο ταξινομητής της μέγιστης πιθανοφάνειας (Maximum Likelihood Classifier) έχουν τον περιορισμό ότι πρέπει να ακολουθούν την κανονική κατανομή (Strahler, 1980), που είναι δύσκολο σε προβλήματα τηλεπισκόπησης, ιδίως στην ταξινόμηση πολυφασματικών δορυφορικών εικόνων. Επίσης, αναπτύχθηκε μία νέα υβριδική μεθοδολογία ταξινόμησης. Επιπλέον ένας άλλος παράγοντας επιλογής τους ήταν το ότι σύμφωνα με την βιβλιογραφία στους αλγόριθμους αυτούς επιδρά ελάχιστα έως καθόλου το φαινόμενο Hughes, που αλλιώς θα μείωνε την ακρίβεια ταξινόμησης λόγω της ύπαρξης πολλών διαστάσεων (Alonso et al., 2001; Hughes, 1968; Watanachaturaporn et al., 2004).

Μέχρι στιγμής σύμφωνα με την βιβλιογραφία δεν υπάρχει σχετική έρευνα που να συνδυάζει σε ένα σύστημα λήψης αποφάσεων στη δημόσια διοίκηση τεχνητή

νοημοσύνη, γεωγραφικά συστήματα πληροφοριών και τηλεπισκόπηση και αυτό αποτελεί άλλον έναν παράγοντα πρωτοτυπίας της παρούσας έρευνας.

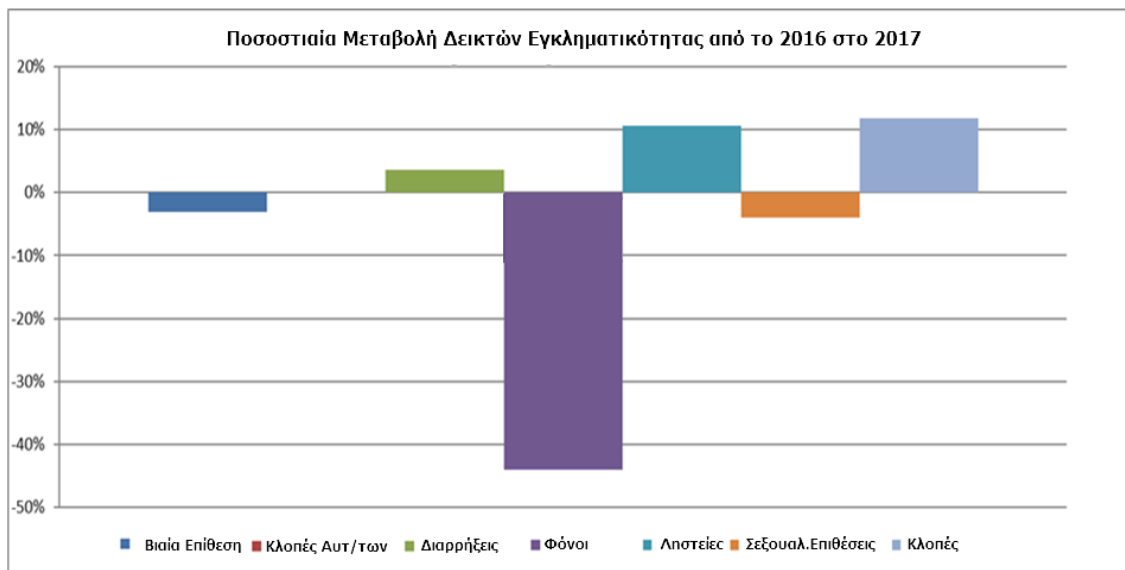
3.4 Παράγοντες στη λήψη αποφάσεων στη δημόσια διοίκηση

Στην ενότητα αυτή, θα ερευνηθεί ο εντοπισμός των σημαντικότερων παραγόντων στη λήψη αποφάσεων σε δημόσιους οργανισμούς και δικαστικές υπηρεσίες σύμφωνα με την βιβλιογραφία. Ο εντοπισμός των παραγόντων αυτών είναι σημαντικός, ώστε να μπορούν να εφαρμοστούν στους παράγοντες αυτούς τα μοντέλα της τεχνητής νοημοσύνης.

3.4.1 Παράγοντες εγκληματικότητας στην δικαστική διοίκηση

Η εγκληματικότητα είναι ένας από τους κυριότερους παράγοντες που επηρεάζουν την δικαστική διοίκηση. Οι δικαστικές αρχές προΐστανται των διωκτικών Αρχών (Αστυνομίας, Λιμενικού, κλπ) και στα πλαίσια του ρόλου τους καταστρώνουν στρατηγικές για την καταπολέμηση της εγκληματικότητας.

Οι δείκτες εγκληματικότητας θεωρούνται ως μια συστηματική προσέγγιση για τη μέτρηση και την αξιολόγηση της έντασης της εγκληματικότητας αποσκοπώντας στην πιο αξιόπιστη εκτίμηση των τάσεων της εγκληματικότητας και στην εφαρμογή βέλτιστων και βιώσιμων πολιτικών στη διοίκηση των πόλεων (Adams, et al., 2014). Η ταξινόμηση του εγκλήματος σε κύριες κατηγορίες σε ένα ενοποιημένο σύστημα ταξινόμησης έχει διερευνηθεί από τις αρχές της Ευρωπαϊκής Ένωσης, ώστε να δημιουργηθεί ένα σύστημα ταξινόμησης για την προσαγωγή σε επίπεδο ΕΕ με τον καθορισμό κατηγοριών παραβάσεων (De Bondt, and Vermeulen, 2009). Ένα παράδειγμα ορισμού δεικτών εγκληματικότητας που προτείνονται από την Αστυνομία του Τορόντο παρουσιάζονται στην επόμενη εικόνα.



Εικόνα 3.2 Ποσοστιαία μεταβολή των δεικτών εγκληματικότητας από το έτος 2016 στο έτος 2017 που προτείνονται από την Αστυνομία του Τορόντο².

3.4.2 Παράγοντες σημαντικοί στην περιβαλλοντική διοίκηση

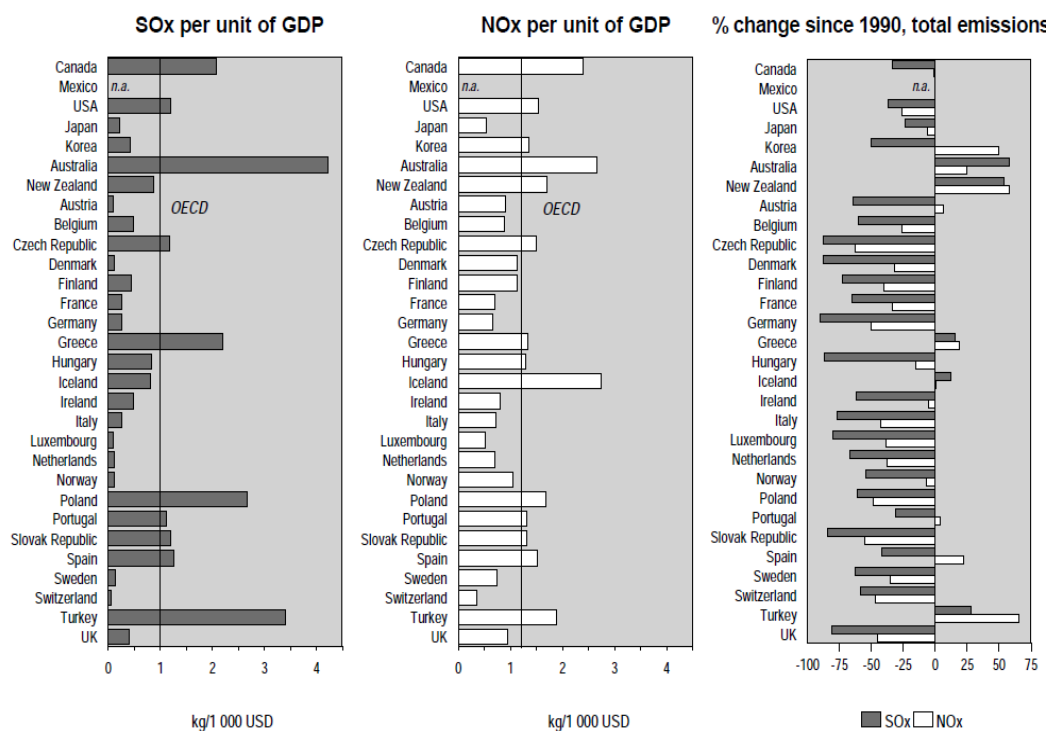
Η σημασία των περιβαλλοντικών δεδομένων είναι υψηλή στη δημόσια διοίκηση, αλλά και στην περιβαλλοντική διαχείριση και το σχεδιασμό για διάφορους λόγους. Τα περιβαλλοντικά δεδομένα και ειδικότερα οι πληροφορίες σχετικά με τους παράγοντες ρύπανσης μπορούν να μελετηθούν και να αναλυθούν προκειμένου να υιοθετηθούν οι κατάλληλες περιβαλλοντικές στρατηγικές για την προστασία της ανθρώπινης υγείας καθώς και του περιβάλλοντος και των φυσικών πόρων (Council Directive, 1992; Evans et al., 2013).

Για παράδειγμα, δύο από τους σημαντικότερους παράγοντες ρύπανσης του αέρα στο αστικό περιβάλλον είναι το διοξείδιο του αζώτου (NO₂) και το όζον (O₃). Αυτοί οι παράγοντες έχουν αρνητικές επιπτώσεις στην ανθρώπινη υγεία, όπως έδειξαν οι έρευνες. Σύμφωνα με διάφορες έρευνες, το όζον και το διοξείδιο του αζώτου έχουν συνδεθεί με

² Πηγή: https://www.torontopolice.on.ca/statistics/crime_indicators.php

την καθημερινή θνησιμότητα σε διάφορα μέρη του κόσμου, όπως το Λονδίνο (Williams et al., 2014) και το Δέλτα του ποταμού Περγλ της νότιας Κίνας (Tao et al., 2012). Επιπλέον, μια άλλη έρευνα έδειξε την επίδραση του ατμοσφαιρικού ρύπου NO_2 σε μικροοργανισμούς και στην αλλεργιογένεση του ανθρώπου (Cuinica et al., 2014).

Στην επόμενη εικόνα, αναπαρίστανται η ένταση των εκπομπών των οξειδίων του θείου SO_x και των οξειδίων του αζώτου NO_x μεταξύ των χωρών του ΟΟΣΑ το 2008 και η επί τοις εκατό αλλαγή από το 1990 στις συνολικές εκπομπές των SO_x και NO_x .



Εικόνα 3.3 Ένταση εκπομπών SO_x και NO_x μεταξύ των χωρών του ΟΟΣΑ το 2008 και η % αλλαγή από το 1990 στις συνολικές εκπομπές SO_x και NO_x (Πηγή: Directorate, 2008).

Το βασικό σετ περιβαλλοντικών δεικτών – κλειδιών, που προτείνεται από τον Ευρωπαϊκό Οργανισμό Περιβάλλοντος (European Environment Agency), υιοθετείται στην παρούσα έρευνα και παρουσιάζεται στον επόμενο πίνακα.

Πίνακας 3.1 Βασικό σετ περιβαλλοντικών δεικτών που προτείνεται από τον Ευρωπαϊκό Οργανισμό Περιβάλλοντος (European Environment Agency) (EEA, 2005).

Ζητήματα Ρύπανσης	Περιβαλλοντικοί δείκτες – κλειδιά
Κλιματική αλλαγή	Ένταση εκπομπών CO ₂
Στοιβάδα του όζοντος	Δείκτες κατανάλωσης ουσιών που καταστρέφουν τη στιβάδα του όζοντος (Ozone Depleting Substances – ODS)
Ποιότητα του αέρα	Εκπομπές SO _x , NO _x , PM ₁₀
Απόρριψη αποβλήτων	Ένταση αστικών αποβλήτων
Ποιότητα πόσιμου νερού	Ποσοστά επεξεργασίας νερού λυμάτων

3.4.3 Παράγοντες κοινωνικοοικονομικοί σημαντικοί στη Δημόσια διοίκηση

Σύμφωνα με τη βιβλιογραφία, πολλοί κοινωνικοοικονομικοί παράγοντες επηρεάζουν τη λήψη αποφάσεων στη δημόσια διοίκηση, όπως για παράδειγμα τα επίπεδα των ποσοστών ανεργίας, η αύξηση του ΑΕΠ και η αύξηση των εξαγωγών αγαθών και υπηρεσιών (Bayar 2014; Ozughalu and Ogwumike 2013). Το σύνολο δεικτών – κοινωνικο-οικονομικών παραγόντων που προτείνεται στην παρούσα διατριβή μετά από επισκόπηση της σχετικής διεθνούς βιβλιογραφίας (Diaz-Chavez, 2014; Metaxas, 2010) παρουσιάζεται στον επόμενο πίνακα.

Πίνακας 3.2 Σετ προτεινόμενων δεικτών – παραγόντων δημόσιας διοίκησης.

Κατηγορία	Δείκτης – παράγοντας
Δημογραφία και υγεία	Ποσοστό γεννήσεων
	Θνησιμότητα
	Πληθυσμός
Εκπαίδευση	Ποσοστό αναλφαριθμητισμού
	Μέσος χρόνος εκπαίδευσης
Εργασία	Ποσοστό ανεργίας
	Μέσος μισθός
Εισόδημα και φτώχεια	ΑΕΠ κατά κεφαλήν
	Μέσο οικογενειακό εισόδημα
Δημόσια Ασφάλεια	Ποσοστό εγκληματικότητας

3.5 Προτεινόμενη μεθοδολογία λήψης αποφάσεων - Προβλεπτική Διοίκηση

Σύμφωνα με τα παραπάνω πραγματοποιήθηκε μία προσπάθεια χαρτογράφησης των κυριότερων παραγόντων – δεικτών που επηρεάζουν τα κυριότερα προβλήματα λήψης αποφάσεων στην δημόσια και δικαστική διοίκηση.

Συνοψίζοντας σύμφωνα με τα παραπάνω, για την αντιμετώπιση ενός προβλήματος λήψης απόφασης στη δημόσια και δικαστική διοίκηση, η προτεινόμενη μεθοδολογία με την εφαρμογή της τεχνητής νοημοσύνης, των γεωγραφικών συστημάτων πληροφοριών και της τηλεπισκόπησης συνοψίζεται στα κάτωθι στάδια:

Στο πρώτο στάδιο, πραγματοποιείται, ο καθορισμός του είδους του προβλήματος για την λήψη απόφασης στη δημόσια και δικαστική διοίκηση, που θα αφορά συγκεκριμένη εξαρτημένη μεταβλητή ή μεταβλητές. Στον καθορισμό του προβλήματος προσδιορίζεται επίσης αν η φύση του είναι χωρική ή μη ώστε να αναζητηθούν δεδομένα με χωρικά χαρακτηριστικά. Επίσης, προσδιορίζεται ο χρονικός ορίζοντας της λήψης απόφασης για την αντιμετώπιση του προβλήματος, δηλαδή αν η απόφαση που πρέπει να ληφθεί είναι βραχυπρόθεσμη, μεσοπρόθεσμη ή μακροπρόθεσμη. Για παράδειγμα αν η λήψη απόφασης αφορά την επόμενη ημέρα, τον επόμενο μήνα ή τον επόμενο χρόνο. Στα επόμενα κεφάλαια, εξετάζονται οι βασικότεροι χρονικοί ορίζοντες λήψης απόφασης με την χρήση της τεχνητής νοημοσύνης.

Στο δεύτερο στάδιο, πραγματοποιείται ο καθορισμός των παραγόντων/δεικτών που επηρεάζουν την εξαρτημένη μεταβλητή στη λήψη αποφάσεων στη δημόσια διοίκηση στο συγκεκριμένο πρόβλημα. Η φύση του προβλήματος επηρεάζει τον τύπο των παραγόντων που θα ληφθούν υπόψη για να εφαρμοστούν τα μοντέλα βασισμένα στην τεχνητή νοημοσύνη. Για παράδειγμα αν το πρόβλημα λήψης απόφασης είναι χωρικής φύσεως θα χρησιμοποιηθούν και τα γεωγραφικά συστήματα πληροφοριών για να επιχειρηθεί η χωρική πρόβλεψη του παράγοντα. Αν η φύση του προβλήματος αφορά εκτίμηση

παραγόντων όπως η αποτίμηση του πρασίνου μία πόλης ή μίας περιοχής μπορούν να εφαρμοστούν τεχνικές τηλεπισκόπησης με την χρήση δεικτών όπως ο Δείκτης Βλάστησης Κανονικοποιημένης Διαφοράς (NDVI: Normalized Difference Vegetation Index) και με εφαρμογή τεχνικών ταξινόμησης δορυφορικών εικόνων βασισμένων στην τεχνητή νοημοσύνη για τον υπολογισμό της βλάστησης στην κάλυψη γης.

Στο τρίτο στάδιο, πραγματοποιείται η εύρεση των κατάλληλων δεδομένων για τους παράγοντες που επηρεάζουν την εξαρτημένη μεταβλητή, που να μπορούν να χρησιμοποιηθούν στο διάνυσμα εισόδου των μοντέλων τεχνητής νοημοσύνης.

Στο τέταρτο στάδιο, πραγματοποιείται ο προσδιορισμός του διανύσματος εισόδου των n παραγόντων για τους οποίους συλλέχθηκαν δεδομένα, που επηρεάζουν σημαντικά την υπό αξιολόγηση εξαρτημένη μεταβλητή και προεπεξεργασία και κανονικοποίηση των σχετικών μεταβλητών. Το διάνυσμα εισόδου X_i την χρονική στιγμή k , για την i -οστή τιμή του δείγματος θα ορίζεται ως:

$$X_i = (f_1(k), f_2(k), \dots, f_{n-1}(k), f_n(k)) \quad (3.1)$$

Όπου οι n παράγοντες (Factors) εισόδου την χρονική στιγμή k συμβολίζονται ως $F_1(k)$, $F_2(k)$, ..., $F_n(k)$.

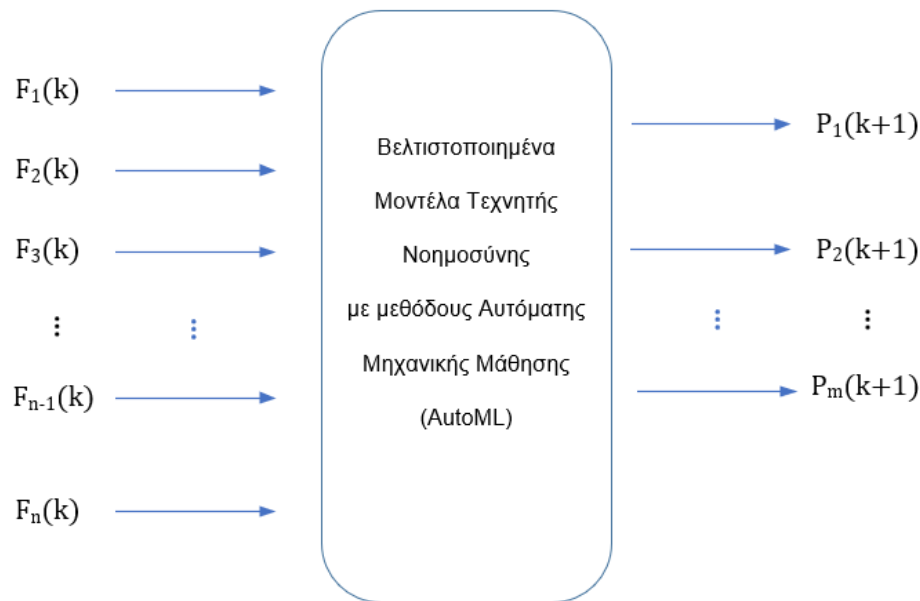
Ενώ το διάνυσμα εξόδου Y_i , που αποτελείται από τις τιμές των m παραμέτρων εξόδου την χρονική στιγμή $k+1$ θα ορίζεται ως:

$$Y_i = (P_1(k+1), P_2(k+1), \dots, P_{n-1}(k+1), P_m(k+1)) \quad (3.2)$$

Όπου οι παράμετροι (Parameters) εξόδου (εξαρτημένες μεταβλητές) της επόμενης χρονικής στιγμής $k+1$ συμβολίζονται ως $P_1(k+1)$, $P_2(k+1)$, ..., $P_n(k+1)$.

Στο πέμπτο στάδιο, επιλέγονται οι κατάλληλες μέθοδοι τεχνητής νοημοσύνης για την πρόβλεψη χρονοσειρών, σε προβλήματα παλινδρόμησης, με χρονική καθυστέρηση μίας χρονικής στιγμής στην πρόβλεψη της εξαρτημένης μεταβλητής, όπως: τεχνητά νευρωνικά δίκτυα εμπρόσθιας διάδοσης, Μηχανές Διανυσμάτων Υποστήριξης, Τυχαία Δάση Δέντρων Απόφασης, Νοημοσύνη Σμήνους Σωματιδίων, κλπ. Δηλαδή, χρήση

διανύσματος εισόδου με μεταβλητές της προηγούμενης χρονικής στιγμής t_k για να προβλεφθεί η επόμενη χρονική στιγμή t_{k+1} .



Εικόνα 3.4 Το προτεινόμενο μοντέλο προβλεπτικής διοίκησης με την χρήση βελτιστοποιημένων μοντέλων Τεχνητής Νοημοσύνης με μεθόδους Αυτόματης Μηχανικής Μάθησης (AutoML). Οι n παράγοντες (Factors) εισόδου την χρονική στιγμή k συμβολίζονται ως $F_1(k)$, $F_2(k)$, ... $F_n(k)$, ενώ οι m παράμετροι (Parameters) εξόδου (εξαρτημένες μεταβλητές) της επόμενης χρονικής στιγμής $k+1$ συμβολίζονται ως $P_1(k+1)$, $P_2(k+1)$, ... $P_m(k+1)$.

Στο έκτο στάδιο, πραγματοποιείται η ανάπτυξη των μοντέλων, και η βελτιστοποίηση των μοντέλων τεχνητής νοημοσύνης με αλγόριθμους αυτοματοποιημένης μηχανικής μάθησης AutoML (Automated Machine Learning) (Οι μηχανές μαθαίνουν – βελτιστοποιούν αυτόματα από μόνες τους), (Feurer et al., 2015; Guyon et al., 2016; Mendoza et al., 2016) όπως ο αλγόριθμος βελτιστοποίησης Σμήνους Σωματιδίων (PSO), η βελτιστοποίηση Bayes, ανάλογα με το μοντέλο τεχνητής νοημοσύνης. Επίσης, ερευνάται και η βέλτιστη τοπολογία των μοντέλων τεχνητής νοημοσύνης με χρήση άλλων τεχνικών, όπως για παράδειγμα η μέθοδος αναζήτησης πλέγματος, που θα αναπτυχθεί περισσότερο στο επόμενο κεφάλαιο, για την διερεύνηση παραγόντων όπως: ο αριθμός των νευρώνων στο κρυφό επίπεδο, η φύση των

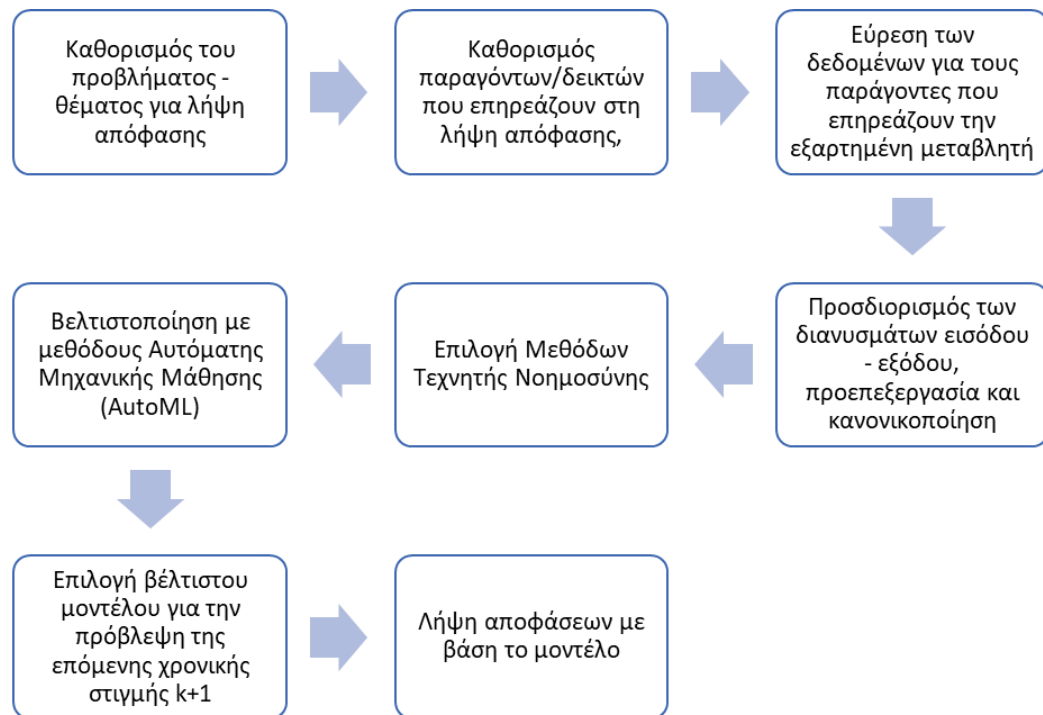
συναρτήσεων ενεργοποίησης, ο αριθμός των νευρώνων στην είσοδο και στην έξοδο, ο αριθμός των κρυφών επιπέδων, ο τύπος ενός νευρωνικού δικτύου. Επίσης, με τις μεθόδους αυτούς πραγματοποιείται η διερεύνηση της βέλτιστης τοπολογίας στην περίπτωση χρήσης των Μηχανών Διανυσμάτων Υποστήριξης, όπως για παράδειγμα ο τύπος του πυρήνα, η συνάρτηση γάμμα και η παράμετρος έψιλον. Επίσης, διερευνάται η βέλτιστη τοπολογία στην χρήση των Τυχαίων Δασών Δέντρων Απόφασης, όπως για παράδειγμα ο αριθμός των δέντρων των τυχαίων δασών. Επιπλέον, διερευνάται η εφαρμογή των νέων μεθοδολογιών βασισμένων στην τεχνητή νοημοσύνη που προτείνονται στην παρούσα διατριβή και που αφορούν την χωρική πρόβλεψη παραγόντων και την ταξινόμηση δορυφορικών εικόνων, που θα αναλυθούν στα επόμενα κεφάλαια.

Στο έβδομο στάδιο, πραγματοποιείται η επιλογή του καλύτερου μοντέλου πρόβλεψης και χρήση του μοντέλου με δεδομένα της προηγούμενης χρονικής στιγμής k ώστε το μοντέλο να προβλέψει την τιμή των υπό εξέταση εξαρτημένων μεταβλητών της επόμενης χρονικής στιγμής $k + 1$, που είναι πολύ σημαντική πληροφορία στην δημόσια διοίκηση.

Στο όγδοο στάδιο, πραγματοποιείται η λήψη αποφάσεων σύμφωνα με τα αποτελέσματα των μοντέλων τεχνητής νοημοσύνης που εφαρμόστηκαν. Η λήψη αποφάσεων στην προβλεπτική διοίκηση σχετίζεται με την αναπροσαρμογή των παραγόντων εισόδου που επηρεάζουν την εξαρτημένη μεταβλητή, και εξετάζονται ως σενάρια σε μία χρονική στιγμή k ώστε να επιτευχθεί η επιθυμητή τιμή της εξαρτημένης μεταβλητής της επόμενης χρονικής στιγμής $k+1$. Με την εφαρμογή του βέλτιστου μοντέλου τεχνητής νοημοσύνης που αναπτύχθηκε, αν αυτές οι αλλαγές θα επιδράσουν θετικά στην τιμή της εξαρτημένης μεταβλητής της επόμενης χρονικής στιγμής $k+1$. Για παράδειγμα αποφάσεις για την μείωση των τιμών των παραμέτρων που επηρεάζουν την τιμή του όζοντος της επόμενης χρονικής στιγμής (Π.χ. της επόμενης ημέρας), μπορούν να πραγματοποιηθούν με την χρήση του μοντέλου και μπορεί να περιλαμβάνουν για παράδειγμα, μείωση εκπομπών από εργοστάσια στην περιοχή που επηρεάζουν την

ρύπανση της ατμόσφαιρας ή απαγόρευση της κυκλοφορίας αυτοκινήτων συγκεκριμένων πινακίδων (όπως μονά – ζυγά) που έχουν εφαρμοστεί κατά καιρούς, κλπ.

Επιπλέον, η λήψη αποφάσεων σε μεσοπρόθεσμο ή μακροπρόθεσμο ορίζοντα μπορεί να αφορά για παράδειγμα την χωροθέτηση εκπαιδευτικών μονάδων, ώστε να βρίσκονται όσο το δυνατόν πιο μακριά από τις περιοχές αυξημένης εγκληματικότητας που προβλέφθηκαν με τα μοντέλα τεχνητής νοημοσύνης άλλων ειδών αποφάσεων που σχετίζονται με την εγκληματικότητα (Caplan and Kennedy, 2016; Chainey and Ratcliffe, 2013; Gorr and Lee, 2015; Leipnik and Albert, 2003; Palocsay et al., 2000; Paulsen and Robinson, 2004; Worrall, 2014). Επίσης, η πρόβλεψη της εγκληματικότητας μπορεί να σχετίζεται με την λήψη αποφάσεων για την βέλτιστη κατανομή των αστυνομικών δυνάμεων για την καταπολέμηση της εγκληματικότητας, στις συγκεκριμένες περιοχές αυξημένης έντασης περιστατικών που προβλέφθηκαν από τα μοντέλα (Camacho-Collados and Liberatore, 2015; Junior et al., 2017; Tambe et al., 2016). Η πρόβλεψη της εγκληματικότητας μπορεί να αποβεί χρήσιμη σε πολλές περιπτώσεις, όπως για παράδειγμα στην πρόβλεψη των λιγότερο ασφαλών περιοχών στο συγκοινωνιακό δίκτυο, (Jenkins, 2001; Kouziokas, 2017c) ή την πρόβλεψη των λιγότερο ασφαλών περιοχών τουριστικού ενδιαφέροντος (Ferreira and Harmse, 2000; Montolio and Planells-Struse, 2016). Επίσης, η πρόβλεψη περιβαλλοντικών παραγόντων μπορούν να είναι αρκετά χρήσιμες σε υπηρεσίες που έχουν να αντιμετωπίσουν την λήψη αποφάσεων σε περιβαλλοντικά θέματα, όπως για παράδειγμα την λήψη μέτρων για την μείωση των αέριων ρύπων. Ένα άλλο πεδίο εφαρμογής αφορά την πρόβλεψη κοινωνικοοικονομικών παραγόντων που μπορεί να είναι αρκετά χρήσιμη σε δημόσιους οργανισμούς που σχετίζονται με τέτοιους παράγοντες όπως για παράδειγμα η ανεργία. Στην επόμενη εικόνα, απεικονίζεται συνοπτικά η προτεινόμενη μεθοδολογία λήψης αποφάσεων.



Εικόνα 3.5 Συνοπτική παρουσίαση των σταδίων της προτεινόμενης μεθοδολογίας για την λήψη αποφάσεων στη δημόσια και δικαστική διοίκηση με την εφαρμογή της Τεχνητής Νοημοσύνης.

ΚΕΦΑΛΑΙΟ 4: Τεχνητή Νοημοσύνη στην Πρόβλεψη της Εγκληματικότητας στην Δημόσια και Δικαστική Διοίκηση

4.1 Συνοπτική Περιγραφή της Έρευνας

Οι εφαρμογές τεχνητής νοημοσύνης έχουν αυξηθεί σημαντικά σε διάφορους επιστημονικούς τομείς, τις τελευταίες δεκαετίες, με την ανάπτυξη νέων τεχνικών μηχανικής μάθησης (Machine Learning) και αλγορίθμων και νέων εργαλείων τεχνητής νοημοσύνης για την ανάπτυξη μοντέλων νευρωνικών δικτύων. Στην παρούσα έρευνα, εφαρμόστηκε ο αλγόριθμος Levenberg Marquardt σε μοντέλα νευρωνικών δικτύων εμπρόσθιας διάδοσης (Feedforward Neural Network Models) για την πρόβλεψη δεδομένων εγκληματικότητας. Ο αλγόριθμος Levenberg Marquardt είναι ένας συνδυασμός του Αλγορίθμου Ταχύτερης Καθόδου (ATK) και του αλγορίθμου Gauss-Newton, ο οποίος χρησιμοποιείται για την επίλυση μη γραμμικών προβλημάτων ελάχιστων τετραγώνων. Ο αλγόριθμος Levenberg Marquardt συνδυάζει τα πλεονεκτήματα ελαχιστοποίησης του αλγορίθμου ταχύτερης καθόδου με το τετραγωνικό μοντέλο (quadratic model) της μεθόδου Gauss-Newton προκειμένου να αυξηθεί η ταχύτητα της συνολικής διαδικασίας εύρεσης του ελάχιστου μιας συνάρτησης. Ο αλγόριθμος Levenberg Marquardt χρησιμοποιήθηκε καθώς θεωρείται ως ο πλέον κατάλληλος για προβλέψεις χρονολογικών σειρών, μεταξύ πολλών άλλων αλγορίθμων μάθησης. Επίσης, εφαρμόστηκαν μοντέλα πρόβλεψης βασισμένα στις Μηχανές Διανυσμάτων Υποστήριξης τα οποία συγκρίθηκαν με αυτά των νευρωνικών δικτύων. Η πρόβλεψη των ποσοστών της εγκληματικότητας μπορεί να διαδραματίσει σημαντικό

ρόλο στον πολεοδομικό σχεδιασμό και τη δημόσια διοίκηση διευκολύνοντας τη λήψη αποφάσεων και την υιοθέτηση των πλέον κατάλληλων στρατηγικών πρόληψης του εγκλήματος και του σχεδιασμού διαχείρισης της δημόσιας ασφάλειας.

4.2 Εισαγωγή – Σχετική Έρευνα

Διάφορες έρευνες ασχολήθηκαν με προβλήματα που σχετίζονται με την εγκληματικότητα και τις αρνητικές επιπτώσεις της στην κοινωνία και στην ποιότητα της ανθρώπινης ζωής (Bursik, 1988; Potter and Kappeler, 2006). Αρκετές έρευνες έχουν μελετήσει την επίδραση διαφόρων κοινωνικοοικονομικών παραγόντων στα ποσοστά εγκληματικότητας, όπως η ανεργία και η φτώχεια (Hooghe et al., 2010; Gillani et al., 2009).

Συγκεκριμένα, διάφορες έρευνες έδειξαν ότι υπάρχει σχέση μεταξύ των ποσοστών των εγκλημάτων ιδιοκτησίας και της ανεργίας (Altindag, 2012; Melick, 2003; Yildiz et al., 2013). Ο Melick (2003), έδειξε ότι υπάρχει μια σημαντική συσχέτιση μεταξύ του ποσοστού ανεργίας και των ποσοστών κλοπής αυτοκινήτων μελετώντας στοιχεία είκοσι δύο ετών μεταξύ του 1979 και του 2000 σε αρκετές πολιτείες των ΗΠΑ. Σε αυτή τη μελέτη κατασκευάστηκαν εμπειρικά μοντέλα με την χρήση δεδομένων χρονοσειρών από διάφορες πολιτείες των ΗΠΑ.

Σε μια άλλη έρευνα, ο Altindag (2012) ερευνήσε την επίδραση της ανεργίας στην εγκληματικότητα χρησιμοποιώντας σύνολα δεδομένων σε διάφορες ευρωπαϊκές χώρες. Τα αποτελέσματα έδειξαν ότι η ανεργία έχει θετική επίδραση στην αύξηση των ποσοστών εγκλημάτων ιδιοκτησίας. Διάφοροι ερευνητές χρησιμοποίησαν τα Τεχνητά Νευρωνικά Δίκτυα (TNN) για να προβλέψουν τα ποσοστά εγκληματικότητας λαμβάνοντας υπόψη διάφορους κοινωνικοοικονομικούς παράγοντες (Alwee et al., 2013; Bertiz et al., 2015; Bogomolov et al., 2014; Kouziokas 2017b, 2017c, 2017d; Kouziokas et al., 2017a). Οι Alwee et al. (2013), χρησιμοποίησαν τα Τεχνητά Νευρωνικά Δίκτυα σε συνδυασμό με τεχνικές παλινδρόμησης για την πρόβλεψη των ποσοστών εγκληματικότητας για αδικήματα ιδιοκτησίας, λαμβάνοντας υπόψη πολλούς οικονομικούς παράγοντες όπως το ακαθάριστο εγχώριο προϊόν, το ποσοστό ανεργίας και το δείκτη τιμών καταναλωτή. Τα πειραματικά αποτελέσματα έδειξαν ότι το προτεινόμενο

μοντέλο μπορεί να προβλέψει τα ποσοστά εγκληματικότητας ιδιοκτησίας με καλύτερη ακρίβεια από άλλα μοντέλα. Οι Bertiz et al. (2015), σε μια άλλη μελέτη, ανέπτυξαν ένα μαθηματικό μοντέλο που βασίζεται στην ανάλυση χρονικών σειρών χρησιμοποιώντας Τεχνητά Νευρωνικά Δίκτυα και τεχνικές εξόρυξης δεδομένων για την πρόβλεψη των μηνιαίων ποσοστών εγκληματικότητας. Τα αποτελέσματα έδειξαν μεγάλη ακρίβεια προβλέψεων των περιστατικών εγκληματικότητας. Οι Bogomolov et al. (2014), ερεύνησαν την εφαρμογή των τεχνητών νευρωνικών δικτύων σε συνδυασμό με δημογραφικά δεδομένα για να προβλέψουν τα ποσοστά της εγκληματικότητας στην πόλη του Λονδίνου. Τα αποτελέσματα έδειξαν πολύ καλή ακρίβεια πρόβλεψης.

Στην παρούσα έρευνα, η τεχνητή νοημοσύνη χρησιμοποιήθηκε ως εργαλείο για την πρόβλεψη των ποσοστών εγκληματικότητας σε αστικές περιοχές καθώς μπορεί να μοντελοποιήσει και μη γραμμικές σχέσεις μεταξύ εισόδου – εξόδου (Svozil et al., 1997). Τα μοντέλα πρόβλεψης που αναπτύχθηκαν χρησιμοποιήθηκαν για την δημιουργία του βέλτιστου μοντέλου πρόβλεψης της εγκληματικότητας. Η μεθοδολογία της έρευνας, τα αποτελέσματα και η συζήτηση περιγράφονται στις ακόλουθες ενότητες.

4.3 Θεωρητικό πλαίσιο

4.3.1 Τεχνητά Νευρωνικά Δίκτυα

Η Τεχνητή Νοημοσύνη (TN) χρησιμοποιήθηκε σε αυτή την έρευνα για την δημιουργία και σύγκριση μοντέλων πρόβλεψης τεχνητών νευρωνικών δικτύων για την πρόβλεψη των ποσοστών της εγκληματικότητας. Τα Τεχνητά Νευρωνικά Δίκτυα είναι υπολογιστικά συστήματα που μιμούνται τη δομή των νευρωνικών δικτύων του ανθρώπινου εγκεφάλου. Τα τεχνητά νευρωνικά δίκτυα επεξεργάζονται δεδομένα από τις τιμές εισόδου. Οι πληροφορίες διαχέονται μέσω των συνδέσεων του νευρωνικού δικτύου προκειμένου να παράγουν την έξοδο σύμφωνα με τις τιμές εισόδου.

Το πλεονέκτημα τους είναι ότι μπορούν να μοντελοποιήσουν και μη γραμμικές σχέσεις μεταξύ εισόδου και εξόδου (Grossberg, 1988; Svozil et al., 1997; Zhang et al.,

1998). Στην παρούσα μελέτη χρησιμοποιούνται τεχνητά νευρωνικά δίκτυα για την πρόβλεψη δεδομένων εγκληματικότητας.

4.3.2 Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης

Ο Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης (Feedforward Multilayer Perceptron (FFMLP)) εφαρμόστηκε προκειμένου να δημιουργηθούν τα μοντέλα πρόβλεψης τεχνητών νευρωνικών δικτύων. Ο Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης εφαρμόστηκε σε αυτή την έρευνα, καθώς αρκετοί ερευνητές έχουν δείξει ότι είναι ο πιο κατάλληλος για προβλήματα που σχετίζονται με την πρόβλεψη χρονοσειρών (Tang and Fishwick, 1993; Zhang, 2001).

Σε ένα πολυεπίπεδο νευρωνικό δίκτυο πολλαπλών στρωμάτων, οι νευρώνες συνδέονται μόνο προς τα εμπρός. Στην τοπολογία ενός τυπικού νευρωνικού δικτύου εμπρόσθιας διάδοσης τα επίπεδα του νευρωνικού δικτύου περιέχουν συνδέσεις με τα επόμενα επίπεδα, αλλά δεν έχουν συνδέσεις προς τα πίσω (Blum and Li, 1991; Hornik, 1991).

4.3.3 Αλγόριθμος Levenberg Marquardt

Ο αλγόριθμος Levenberg Marquardt επιλέχθηκε ως αλγόριθμος εκπαίδευσης για το νευρωνικό δίκτυο εμπρόσθιας διάδοσης που χρησιμοποιήθηκε, αφού θεωρείται ένας από τους γρηγορότερους αλγορίθμους μάθησης σε σύγκριση με άλλους (Marquardt, 1963). Επίσης, ο αλγόριθμος Levenberg Marquardt παρήγαγε καλύτερα αποτελέσματα από τους πιο συνηθισμένους αλγορίθμους που δοκιμάστηκαν: Ελαστικής Οπισθοδιάδοσης (Resilient Backpropagation), BFGS Quasi-Newton, Κλιμακούμενων Συζυγών Κλίσεων (Scaled Conjugate Gradient), χρησιμοποιώντας διαφορετικές τοπολογίες σχετικά με τον αριθμό των νευρώνων στα κρυφά επίπεδα, τον αριθμό των κρυφών επιπέδων και τις πιο κοινές συναρτήσεις μεταφοράς (transfer functions).

Ο αλγόριθμος Levenberg Marquardt είναι ένας συνδυασμός του αλγορίθμου ταχύτερης καθόδου και του αλγορίθμου Gauss-Newton ο οποίος χρησιμοποιείται για την

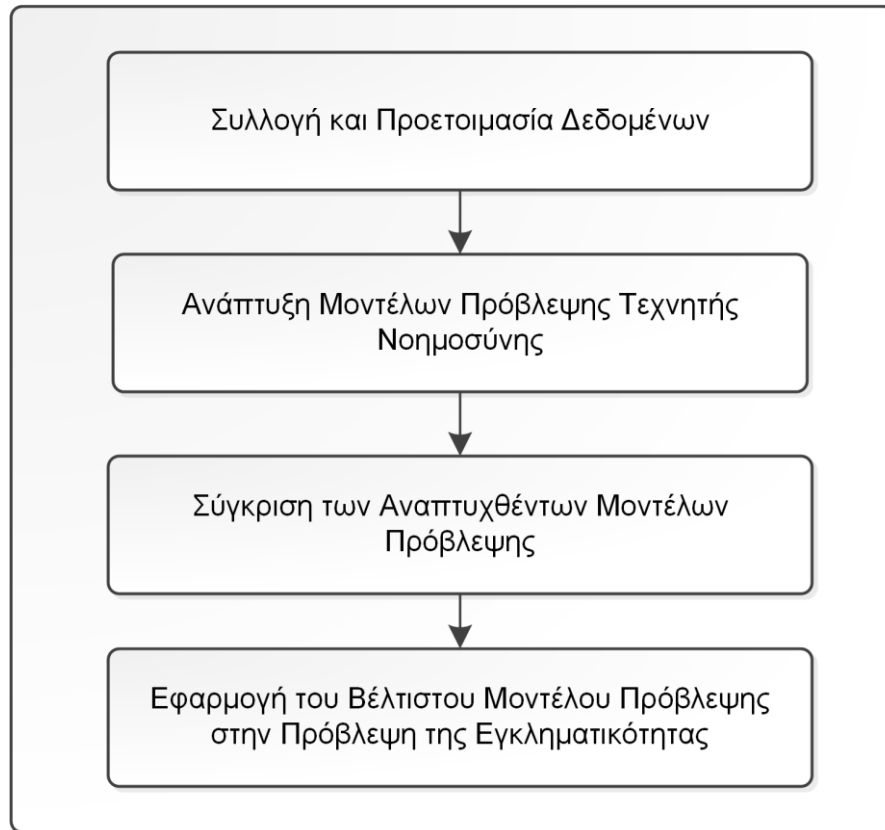
επίλυση μη γραμμικών προβλημάτων ελάχιστων τετραγώνων (Lourakis, 2005). Ο αλγόριθμος Levenberg Marquardt συνδυάζει τα πλεονεκτήματα ελαχιστοποίησης του αλγόριθμου ταχύτερης καθόδου με το τετραγωνικό μοντέλο (quadratic model) της μεθόδου Gauss-Newton προκειμένου να αυξηθεί η ταχύτητα της συνολικής διαδικασίας εύρεσης του ελάχιστου μιας συνάρτησης (Liu, 2010; Moré, 1978).

4.4 Μεθοδολογία

Η μεθοδολογία της έρευνας αποτελείται συνοπτικά από τέσσερα στάδια: συλλογή και προετοιμασία δεδομένων, ανάπτυξη μοντέλων πρόβλεψης, σύγκριση των μοντέλων πρόβλεψης με σκοπό την επιλογή του βέλτιστου μοντέλου σύμφωνα με τις επιδόσεις και την εφαρμογή του βέλτιστου μοντέλου για την πρόβλεψη των ποσοστών εγκληματικότητας. Για την ανάπτυξη των μοντέλων τεχνητής νοημοσύνης χρησιμοποιήθηκε το λογισμικό MATLAB.

Στο πρώτο στάδιο, συλλέχθηκαν και προετοιμάστηκαν δεδομένα σχετικά με την εγκληματικότητα και την ανεργία προκειμένου να τροφοδοτηθούν τα μοντέλα νευρωνικών δικτύων. Στο δεύτερο στάδιο, αναπτύχθηκαν μοντέλα πρόβλεψης τεχνητών νευρωνικών δικτύων με δοκιμές διαφορετικών τοπολογιών προκειμένου να κατασκευαστεί το βέλτιστο νευρωνικό μοντέλο πρόβλεψης. Επίσης, αναπτύχθηκαν μοντέλα πρόβλεψης βασισμένα στις Μηχανές Διανυσμάτων Υποστήριξης. Στο επόμενο στάδιο, πραγματοποιήθηκε σύγκριση των αναπτυχθέντων μοντέλων νευρωνικών δικτύων, προκειμένου να επιλεγεί το βέλτιστο μοντέλο ανάλογα με τις επιδόσεις του. Για την βελτιστοποίηση του νευρωνικού δικτύου χρησιμοποιήθηκαν ο αλγόριθμος αναζήτησης πλέγματος και ο αλγόριθμος βελτιστοποίησης Σμήνους Σωματιδίων. Για την βελτιστοποίηση των υπερπαραμέτρων των Μηχανών Διανυσμάτων Υποστήριξης χρησιμοποιήθηκε ο αλγόριθμος βελτιστοποίησης Bayes. Στην τελευταία φάση, το βέλτιστο μοντέλο πρόβλεψης με κάθε μέθοδο τεχνητής νοημοσύνης χρησιμοποιήθηκαν για να προβλέψουν τα ποσοστά της εγκληματικότητας και τέλος συγκρίθηκαν μεταξύ

τους ως προς την ακρίβεια πρόβλεψης. Στην παρακάτω εικόνα, παρουσιάζεται μια επισκόπηση των σταδίων της μεθοδολογίας της έρευνας.



Εικόνα 4.1. Επισκόπηση των σταδίων της εφαρμοσθείσας μεθοδολογίας για την πρόβλεψη των ποσοστών εγκληματικότητας με την εφαρμογή μεθόδων τεχνητής νοημοσύνης.

4.5 Αποτελέσματα

4.5.1 Συλλογή Δεδομένων – Περιοχή Μελέτης

Τα εγκλήματα ιδιοκτησίας (διάρρηξη, κλοπή αυτοκινήτων, ληστεία και κλοπή) επελέγησαν για να ερευνηθεί η πρόβλεψη του ποσοστού τους σε αυτήν την μελέτη. Προκειμένου να βελτιωθεί η ακρίβεια των αποτελεσμάτων, η ανεργία επελέγη ως μεταβλητή εισόδου στα μοντέλα νευρωνικών δικτύων που αναπτύχθηκαν, λαμβάνοντας υπόψη ότι οι προηγούμενες έρευνες δείχνουν μια θετική σχέση μεταξύ των ποσοστών

ανεργίας και των ποσοστών εγκληματικότητας που αφορούν αδικήματα ιδιοκτησίας (Levitt, 2001). Επίσης, η έρευνα αυτή επικεντρώθηκε στην πρόβλεψη του εγκλήματος σε επίπεδο πολιτείας στις ΗΠΑ και όχι σε επίπεδο χώρας. Ο πρώτος λόγος είναι ότι κάθε πολιτεία έχει τα δικά της χαρακτηριστικά που την καθιστούν μοναδική σε σχέση με τις άλλες. Ο δεύτερος λόγος είναι ότι στο εσωτερικό μιας πολιτείας, τα πρότυπα της ανεργίας και των εγκλημάτων είναι συνήθως ομογενή. Επιλέχθηκε η περιοχή αυτή, λόγω της διαθεσιμότητας μεγαλύτερου αριθμού δεδομένων σε βάθος χρόνου των παραγόντων εισόδου: εγκληματικότητας και ανεργίας σε μηνιαία βάση, σε αντίθεση με την Ελλάδα, ώστε να είναι ικανά να εκπαιδεύσουν επαρκώς μοντέλα τεχνητής νοημοσύνης.

Τα στοιχεία σχετικά με τα εγκλήματα ιδιοκτησίας (διάρρηξη, κλοπή αυτοκινήτων, ληστεία, κλοπή) ανακτήθηκαν από τον ιστότοπο του Στατιστικού Γραφείου Δικαιοσύνης των ΗΠΑ (U.S. Bureau of Justice Statistics 2018). Τα στοιχεία σχετικά με την ανεργία του Ιλλινόις ανακτήθηκαν από τον επίσημο ιστότοπο του Γραφείου Στατιστικής Εργασίας του Τμήματος Εργασίας των Ηνωμένων Πολιτειών (U.S. Bureau of Labor Statistics 2018). Τα δεδομένα συλλέχθηκαν για τη χρονική περίοδο από το 2004 έως το 2015 σε μηνιαία βάση, έτσι ώστε τα ιστορικά δεδομένα της περιοχής της μελέτης να αρκούν για να τροφοδοτήσουν και να εκπαιδεύσουν επαρκώς τα τεχνητά νευρωνικά δίκτυα. Τα δεδομένα προεπεξεργάστηκαν και προετοιμάστηκαν ελέγχοντας πιθανές ασυνέπειες και διπλότυπα, ώστε να χρησιμοποιηθούν ως μεταβλητές εισόδου στα μοντέλα τεχνητών νευρωνικών δικτύων.

Τα δεδομένα κανονικοποιήθηκαν χρησιμοποιώντας την κανονικοποίηση min-max (ελαχίστου – μεγίστου) στο διάστημα $[0,1]$, η οποία μετασχηματίζει γραμμικά τα δεδομένα στο διάστημα αυτό. Η κανονικοποίηση min-max θεωρείται η καλύτερη και εξάγει ακριβέστερα αποτελέσματα από άλλες μεθόδους κανονικοποίησης, όπως η κανονικοποίηση z-σκορ, και εκφράζεται από την ακόλουθη εξίσωση (Al Shalabi et al., 2006):

$$n_i = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (4.1)$$

Όπου:

n_i εκφράζει την i-οστή κανονικοποιημένη τιμή του δείγματος

x_i εκφράζει την i -οστή τιμή του δείγματος

$\min(x)$ εκφράζει την ελάχιστη τιμή του δείγματος

$\max(x)$ εκφράζει την μέγιστη τιμή του δείγματος

4.5.2 Μοντέλα Πρόβλεψης Τεχνητών Νευρωνικών Δικτύων

Τα μοντέλα πρόβλεψης τεχνητών νευρωνικών δικτύων αναπτύχθηκαν χρησιμοποιώντας ως παραμέτρους εισόδου τα ιστορικά δεδομένα των εγκλημάτων ιδιοκτησίας και τα ποσοστά ανεργίας της περιοχής μελέτης.

Το διάνυσμα εισόδου θα αποτελείται από την τιμή της εγκληματικότητας CR_{n-1} που αφορά εγκλήματα ιδιοκτησίας και του ποσοστού της ανεργίας U_{n-1} (Unemployment) του προηγούμενου μήνα:

$$X_i = (CR_{n-1}, U_{n-1}) \quad (4.2)$$

Οι τιμές εξόδου Y_i αφορούν τις τιμές που προβλέπει το κάθε μοντέλο σχετικά με την εγκληματικότητα του επόμενου μήνα:

$$Y_i = (CR_n) \quad (4.3)$$

Τα δεδομένα που συλλέχθηκαν χωρίστηκαν σε τρία διαφορετικά μέρη. Το 60% των δεδομένων χρησιμοποιήθηκε ως σύνολο εκπαίδευσης, το 20% των δεδομένων για το σύνολο επικύρωσης και το 20% των δεδομένων για το σύνολο δοκιμών. Το σύνολο δεδομένων εκπαίδευσης χρησιμοποιήθηκε στη διαδικασία εκπαίδευσης των τεχνητών νευρωνικών δικτύων χρησιμοποιώντας τα συλλεχθέντα ιστορικά δεδομένα. Το σύνολο επικύρωσης (validation set) χρησιμοποιήθηκε για την αξιολόγηση των μοντέλων πρόβλεψης τεχνητών νευρωνικών δικτύων. Ο αλγόριθμος Levenberg Marquardt (LM) βρέθηκε ως βέλτιστος αλγόριθμος εκπαίδευσης.

Για την βελτιστοποίηση της τοπολογίας του νευρωνικού δικτύου εφαρμόστηκε η μέθοδος αναζήτησης πλέγματος (Grid Search) και για την βελτιστοποίηση των βαρών

του νευρωνικού δικτύου εφαρμόστηκε η μέθοδος Βελτιστοποίησης Σμήνους Σωματιδίων (Particle Swarm Optimization, PSO) (Eberhart and Kennedy, 1995).

4.5.3 Βελτιστοποίηση με Αναζήτηση Πλέγματος

Για την βελτιστοποίηση της τοπολογίας του νευρωνικού δικτύου εφαρμόστηκε η μέθοδος αναζήτησης πλέγματος (Grid Search) (Basnayake et al., 2017). Το πλεονέκτημα της μεθόδου είναι ότι αυτή είναι πλήρης, δηλαδή ερευνά κάθε δυνατό συνδυασμό για να βρεθεί ο βέλτιστος. Όταν ο αριθμός όλων των πιθανών συνδυασμών των παραμέτρων δεν είναι τόσο μεγάλος και μπορεί ελεγχθεί υπολογιστικά θεωρείται η πιο πλήρης και αποδοτική μέθοδος (Basnayake et al., 2017).

Επίσης, ένα άλλο πλεονέκτημα είναι ότι μπορεί να βελτιστοποιήσει υπερπαραμέτρους που αποτελούν μίγμα κατηγορικών και μη κατηγορικών υπερπαραμέτρων σε αντίθεση με άλλες βελτιστοποιήσεις αυτόματης μηχανικής μάθησης (AutoML) όπως η βελτιστοποίηση Bayes, οι οποίες αντιμετωπίζουν προβλήματα όταν υπάρχει συνδυασμός τέτοιου είδους υπερπαραμέτρων (Feurer et al., 2015; Guyon et al., 2016; Mendoza et al., 2016). Όπως για παράδειγμα, οι κατηγορικές μεταβλητές στο παρόν πρόβλημα θα είναι: τύπος συνάρτησης μεταφοράς, τύπος αλγορίθμου εκπαίδευσης που πρέπει να βελτιστοποιηθούν σε συνδυασμό με μη κατηγορικές μεταβλητές όπως: ο αριθμός νευρώνων στο κρυφό επίπεδο.

Το βέλτιστο μοντέλο νευρωνικού δικτύου κατασκευάστηκε σύμφωνα με τη ν-διάστατη απεικόνιση σε πλέγμα των συνδυασμών των συντελεστών της τοπολογίας του νευρωνικού δικτύου καθώς και σύγκριση των αποτελεσμάτων ακρίβειας πρόβλεψης για την επιλογή του καλύτερου συνδυασμού παραμέτρων.

Οι παράμετροι που συνδυάστηκαν αφορούν τις πιο συνηθισμένες συναρτήσεις μεταφοράς που δοκιμάστηκαν για τα κρυφά επίπεδα είναι οι εξής: Η εφαπτομενική – σιγμοειδής συνάρτηση μεταφοράς (Tan-Sigmoid Transfer Function (TSTF)), η λογιστική – σιγμοειδής συνάρτηση μεταφοράς (Log-Sigmoid Transfer Function (LSTF)), η γραμμική συνάρτηση μεταφοράς (Linear Transfer Function (LTF)), η Elliot σιγμοειδής

συνάρτηση μεταφοράς (Elliot Sigmoid Transfer Function (ESTF)) και η θετική γραμμική συνάρτηση μεταφοράς (Positive Linear Transfer Function (PLTF)).

Επίσης, δοκιμάστηκαν οι συνδυασμοί των πιο συνηθισμένων αλγορίθμων μηχανικής μάθησης δοκιμάστηκαν για την εκπαίδευση κάθε νευρωνικού μοντέλου και συγκεκριμένα οι: αλγόριθμος Κλιμακούμενων Συζυγών Κλίσεων (Scaled Conjugate Gradient – SCG), BFGS Quasi-Newton (BFGS-QN), Ελαστικής Οπισθοδιάδοσης (Resilient Backpropagation – RB) και Levenberg Marquardt (LM). Επίσης, εξετάστηκαν τοπολογίες νευρωνικών δικτύων με αριθμό κρυφών επιπέδων (hidden layers) από ένα έως δύο, και αριθμό νευρώνων κάθε κρυφού επιπέδου από έναν έως πενήντα.

Η Ρίζα του Μέσου Τετραγωνικού Σφάλματος (Root Mean Square Error - RMSE) χρησιμοποιήθηκε για την εκτίμηση του σφάλματος πρόβλεψης για κάθε νευρωνικό μοντέλο που αναπτύχθηκε. Ο υπολογισμός της Ρίζας του Μέσου Τετραγωνικού Σφάλματος (RMSE) δίνεται από την εξίσωση:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_{p_i} - y_{r_i})^2} \quad (4.4)$$

Όπου y_{p_i} η τιμή που προβλέφθηκε, y_{r_i} η πραγματική τιμή και N ο αριθμός των τιμών στην έξοδο.

Στον παρακάτω πίνακα, απεικονίζονται τα βέλτιστα νευρωνικά μοντέλα που βρέθηκαν για κάθε αλγόριθμο που δοκιμάστηκε χρησιμοποιώντας διαφορετικές τοπολογίες και συναρτήσεις ενεργοποίησης. Όπως φαίνεται από τον πίνακα το βέλτιστο μοντέλο με το μικρότερο σφάλμα επικύρωσης είναι το ANN4.

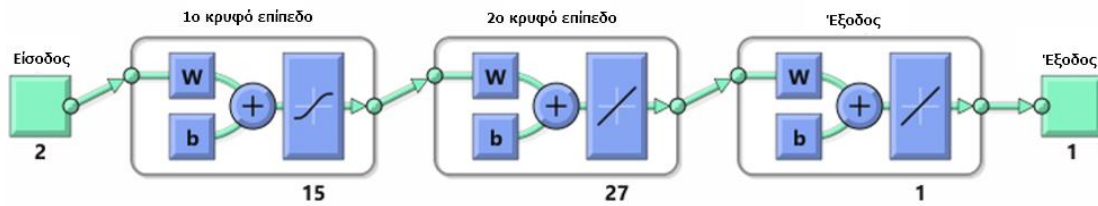
Πίνακας 4.1. Τα βέλτιστα νευρωνικά μοντέλα που βρέθηκαν για κάθε αλγόριθμο που δοκιμάστηκε χρησιμοποιώντας διαφορετικές τοπολογίες και συναρτήσεις ενεργοποίησης.

Μοντέλο	Αλγόριθμος Εκπαίδευσης	Αριθμός Κρυφών Επιπέδων	Αριθμός Νευρώνων στο 1 ^ο Κρυφό Επίπεδο	Αριθμός Νευρώνων στο 2 ^ο Κρυφό Επίπεδο	Συναρτήσεις Μεταφοράς (<i>Transfer Functions</i>)	RMSE Σφάλμα Εκπαίδευσης (<i>Training Error</i>)	RMSE Σφάλμα Επικύρωσης (<i>Validation Error</i>)
ANN1	SCG	2	16	29	LTF- LTF	18.9887	1.6327
ANN2	BFGS- QN	2	15	27	LTF- LTF	735.5156	186.0214
ANN3	RB	2	15	18	LTF- LTF	30.2417	13.5655
ANN4	LM	2	15	27	TSAF- LTF	0.1943	0.2548

Στο βέλτιστο μοντέλο νευρωνικού δικτύου, ο αλγόριθμος εκπαίδευσης ήταν ο αλγόριθμος Levenberg Marquardt. Η τοπολογία του βέλτιστου μοντέλου νευρωνικού δικτύου αποτελείται από 15 νευρώνες στο πρώτο κρυφό επίπεδο και την εφαπτομενική – σιγμοειδής συνάρτηση μεταφοράς (Tan-Sigmoid Transfer Function, TSTF) ως συνάρτηση μεταφοράς και 27 νευρώνες στο δεύτερο κρυφό επίπεδο και την γραμμική συνάρτηση μεταφοράς (Linear Transfer Function, LTF) ως συνάρτηση μεταφοράς.

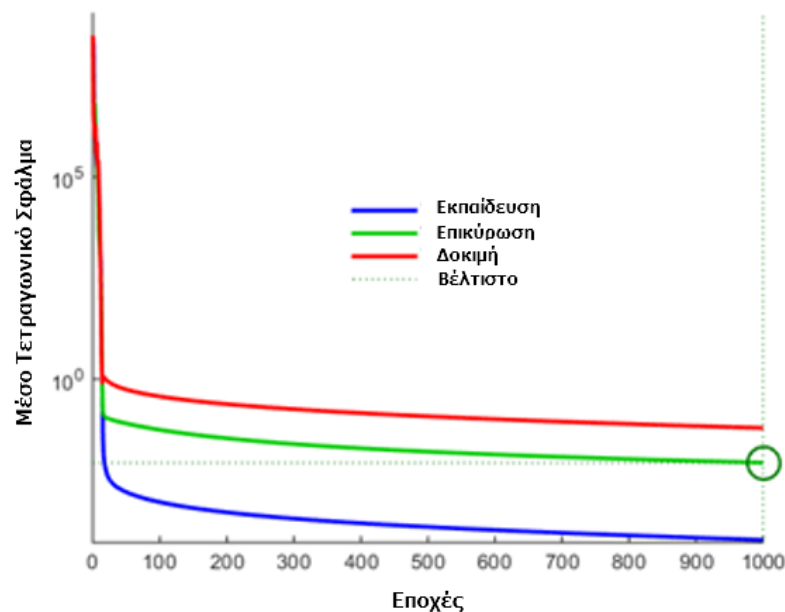
Το βέλτιστο μοντέλο αξιολογήθηκε με τη Ρίζα του Μέσου Τετραγωνικού Σφάλματος (Root Mean Squared Error (RMSE)), σύμφωνα με την σύγκριση που πραγματοποιήθηκε μεταξύ όλων των άλλων αναπτυχθέντων μοντέλων νευρωνικών δικτύων.

Η Ρίζα του Μέσου Τετραγωνικού Σφάλματος (RMSE) του συνόλου επικύρωσης του βέλτιστου μοντέλου βρέθηκε να είναι 0.2548 στην εποχή (epoch) 1000. Στην παρακάτω εικόνα, απεικονίζεται η τοπολογία του του βέλτιστου μοντέλου τεχνητού νευρωνικού δικτύου.



Εικόνα 4.2. Το σχήμα αναπαριστά την τοπολογία του βέλτιστου μοντέλου τεχνητού νευρωνικού δικτύου με δύο κρυφά επίπεδα (hidden layers). Το πρώτο κρυφό επίπεδο αποτελείται από 15 νευρώνες και χρησιμοποιεί την εφαπτομενική – σιγμοειδής συνάρτηση μεταφοράς, ενώ το δεύτερο επίπεδο αποτελείται από 27 νευρώνες και χρησιμοποιεί την γραμμική συνάρτηση μεταφοράς.

Το διάγραμμα απόδοσης του βέλτιστου μοντέλου τεχνητού νευρωνικού δικτύου παρουσιάζεται στην παρακάτω εικόνα, το οποίο αξιολογείται σύμφωνα με το Μέσο Τετραγωνικό Σφάλμα (Mean Squared Error – MSE) του συνόλου επικύρωσης.



Εικόνα 4.3. Η γραφική παράσταση απόδοσης του βέλτιστου μοντέλου τεχνητού νευρωνικού δικτύου, σύμφωνα με το ελάχιστο Μέσο Τετραγωνικό Σφάλμα (Mean Squared Error) συναρτήσει των εποχών.

4.5.4 Βελτιστοποίηση Σμήνους Σωματιδίων (Particle Swarm Optimization)

Για την βελτιστοποίηση των βαρών του νευρωνικού δικτύου εφαρμόστηκε η μέθοδος Βελτιστοποίησης Σμήνους Σωματιδίων (Particle Swarm Optimization, PSO). Η βελτιστοποίηση Σμήνους Σωματιδίων προτάθηκε από τους Eberhart and Kennedy (1995). Επιλέχθηκε ο αλγόριθμος Βελτιστοποίησης Σμήνους Σωματιδίων, διότι σε σύγκριση με άλλες μεθόδους βελτιστοποίησης νευρωνικών δικτύων, όπως ο γενετικός αλγόριθμος πλεονεκτεί ως προς διάφορα σημεία όπως η ευκολία εφαρμογής και η γρήγορη ρύθμιση λιγότερων παραμέτρων (Ou and Lin, 2006).

Έχουν αναπτυχθεί διάφοροι αλγόριθμοι τεχνητής νοημοσύνης που μιμούμενοι διαδικασίες της φύσης έχουν ως βασικό στόχο την επίλυση διαφόρων ειδών υπολογιστικών προβλημάτων και προβλημάτων βελτιστοποίησης. Έχουν προταθεί αλγόριθμοι που μιμούνται πλήθη ζώων ή και εντόμων και προσομοιώνουν την κοινωνική τους συμπεριφορά και τον τρόπο που ανταλλάσσουν πληροφορίες μεταξύ τους και οργανώνονται σε αποικίες. Τέτοιου είδους αλγόριθμοι που κατατάσσονται στην κατηγορία της Νοημοσύνης Σμηνών (Swarm Intelligence) είναι: η Βελτιστοποίηση Σμήνους Σωματιδίων (Particle Swarm Optimization), ο αλγόριθμος Αποικίας Μυρμηγκιών (Ant Colony Algorithm – ACA) και ο αλγόριθμος Αποικίας Μελισσών (Bee Colony Algorithm – BCA) (Blum and Li, 2008; Bonabeau et al., 1999).

Όλοι αυτοί οι αλγόριθμοι αναπτύχθηκαν προσομοιώνοντας τους τρόπους επικοινωνίας των σμηνών της φύσης, όπως τα σμήνη πουλιών, μυρμηγκιών, ώστε να καταφέρνουν με υπολογιστικούς τρόπους και να δώσουν ικανοποιητικά αποτελέσματα σε διαφόρων ειδών προβλήματα βελτιστοποίησης στην τεχνητή νοημοσύνη καθώς και σε άλλες επιστήμες.

Η Βελτιστοποίηση Σμήνους Σωματιδίων ανήκει στους εξελικτικούς αλγόριθμους, οι οποίοι προτιμώνται έναντι των κλασικών μεθόδων βελτιστοποίησης, όπως για παράδειγμα ο αλγόριθμος κατάβασης βαθμίδας και αλγόριθμος του Νεύτωνα (Coleman and Li, 1994), επειδή η φύση τους είναι μετα-ευριστική. Δηλαδή, ο αλγόριθμος έχει ελάχιστη ή και καθόλου επίγνωση του προβλήματος βελτιστοποίησης που πρέπει να επιλύσει και δεν χρειάζεται να γνωρίζει εκ των προτέρων πολύπλοκες παραμέτρους, ούτε την πρώτη ή δεύτερη παράγωγο της αντικειμενικής συνάρτησης (Ou and Lin, 2006).

Η Βελτιστοποίηση Σμήνους Σωματιδίων εξετάζει πληθυσμούς λύσεων για την εύρεση του βέλτιστου μίας συνάρτησης ή ενός προβλήματος.

Ο αλγόριθμος Βελτιστοποίησης Σμήνους Σωματιδίων PSO βασίζεται σε μια στοχαστική τεχνική που εμπνεύστηκε από τη διερεύνηση της κοινωνικής συμπεριφοράς και της νοημοσύνης σμήνους στη φύση. Στον αλγόριθμο PSO, τα σωματίδια του σμήνους διασχίζουν το χώρο αναζήτησης για να αξιολογήσουν την αντικειμενική συνάρτηση (objective function). Κάθε σωματίδιο σμήνους χαρακτηρίζεται από την τρέχουσα θέση, την προηγούμενη καλύτερη θέση και την ταχύτητα.

Η ταχύτητα κάθε σωματιδίου ενημερώνεται χρησιμοποιώντας την εξίσωση (Clerc and Kennedy, 2002; Eberhart and Kennedy, 1995; Eberhart et al. 2001; Fourie and Groenwold, 2002):

$$v_i(t+1) = w v_i(t) + c_1 r_1 [\hat{x}_i(t) - x_i(t)] + c_2 r_2 [g(t) - x_i(t)] \quad (4.5)$$

όπου

$v_i(t)$ είναι η ταχύτητα του σωματιδίου στο χρόνο t

$x_i(t)$ είναι η θέση του σωματιδίου στο χρόνο t

$\hat{x}_i(t)$ είναι η καλύτερη λύση του μεμονωμένου σωματιδίου στο χρόνο t

r_1, r_2 είναι τυχαίοι αριθμοί ομοιόμορφα κατανοημένοι στο διάστημα $[0,1]$

$g(t)$ είναι η καλύτερη λύση του σμήνους στο χρόνο t

c_1 είναι ο συντελεστής προσωπικής μάθησης (personal learning coefficient)

c_2 είναι ο ολικός συντελεστής μάθησης (global learning coefficient)

w είναι ο συντελεστής βάρους αδράνειας (inertia weight coefficient)

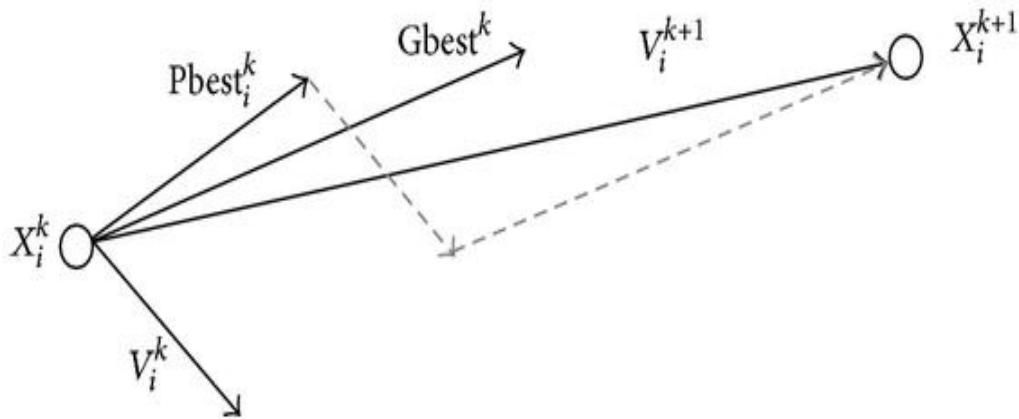
Οι θέσεις των σωματιδίων ενημερώνονται εφαρμόζοντας την ακόλουθη εξίσωση:

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (4.6)$$

Ο αλγόριθμος βελτιστοποίησης σμήνους σωματιδίων (PSO) αποτελείται από τα παρακάτω βασικά στάδια (Eberhart and Kennedy, 1995).:

1. Αρχικοποίησε το σμήνος σωματιδίων, με την ρύθμιση του μεγέθους του πληθυσμού m , των παραμέτρων ταχύτητας σωματιδίων και θέσης, του συντελεστή βάρους αδρανείας w , των συντελεστών μάθησης c_1 και c_2 , των τυχαίων αριθμών στο διάστημα $[0,1]$ των παραμέτρων επιτάχυνσης των σωματιδίων του σμήνους r_1, r_2 , της συνάρτησης καταλληλότητας (fitness function) και το κριτήριο τερματισμού.
2. Χρησιμοποίησε την συνάρτηση καταλληλότητας και υπολόγισε την καταλληλότητα κάθε σωματιδίου και συγκρίνοντας την ατομική βέλτιστη τιμή $Pbest_{(i)}$ (Particle best) με την ολική βέλτιστη τιμή του σμήνους $Gbest_{(i)}$ (Global best), αν η τρέχουσα θέση είναι καλύτερη, τότε αντικαταστήστε, αλλιώς αμετάβλητη.
3. Ενημέρωσε την ταχύτητα και τη θέση των σωματιδίων του σμήνους σύμφωνα με τις εξισώσεις (4.5) και (4.6).
4. Εάν επιτευχθεί το κριτήριο τερματισμού, τερματίζεται ο αλγόριθμος. Αν δεν επιτευχθεί το κριτήριο τερματισμού, πηγαίνει στο βήμα 2.

Στο παρακάτω σχήμα, παρουσιάζεται μία απλοποιημένη σχηματική αναπαράσταση της κίνησης του σωματιδίου από την θέση $x_i(k)$ στην νέα $x_i(k+1)$, όπου k η χρονική στιγμή. Στο σχήμα αναπαρίστανται τα διανύσματα που επηρεάζουν την κίνηση του σωματιδίου του σμήνους. Αυτά είναι: το διάνυσμα της ταχύτητας του σωματιδίου $V_i(k)$ στο χρόνο k , η απόσταση από το βέλτιστο του σμήνους που αναπαρίσταται ως $Gbest$ (Global best) στη χρονική στιγμή k (Κοινωνική Συνιστώσα), η απόσταση από την καλύτερη λύση του μεμονωμένου σωματιδίου που αναπαρίσταται ως $Pbest$ (Particle best) ή αλλιώς στη χρονική στιγμή k (Προσωπική Συνιστώσα). Με την άθροιση των διανυσμάτων αυτών προκύπτει το διάνυσμα της νέας ταχύτητας του σωματιδίου του σμήνους και $V_i(k+1)$ στη χρονική στιγμή $k+1$ και η νέα θέση του σωματιδίου $x_i(k+1)$.



Εικόνα 4.4. Η γραφική παράσταση της κίνησης ενός σωματιδίου στο χώρο αναζήτησης, στον αλγόριθμο Βελτιστοποίησης Σμήνους Σωματιδίων. Η καλύτερη λύση του μεμονωμένου σωματιδίου αναπαρίσταται ως $Pbest$ (Particle best), η καλύτερη λύση του σμήνους αναπαρίσταται ως $Gbest$ (Global best) στη χρονική στιγμή k .

4.5.5 Αλγόριθμος Βελτιστοποίησης Νευρωνικού Δικτύου με Σμήνος Σωματιδίων

Ο προτεινόμενος αλγόριθμος βελτιστοποίησης των βαρών του νευρωνικού δικτύου με την βελτιστοποίηση σμήνους σωματιδίων (PSO) αποτελείται από τα παρακάτω βασικά στάδια:

1. Αρχικοποίησε την αρχιτεκτονική του νευρωνικού δικτύου που βρέθηκε με την μέθοδο αναζήτησης πλέγματος και αφορά: τον αριθμό των κρυφών επιπέδων, τον αριθμό των νευρώνων σε κάθε κρυφό επίπεδο, τις συναρτήσεις μεταφοράς των κρυφών επιπέδων και τον αλγόριθμο εκπαίδευσης.
2. Διάβασε τα δεδομένα από το διάνυσμα εισόδου.
3. Αρχικοποίησε το σμήνος σωματιδίων, με την ρύθμιση του μεγέθους του πληθυσμού m , των παραμέτρων ταχύτητας σωματιδίων και θέσης, του συντελεστή βάρους αδρανείας w , των συντελεστών μάθησης c_1 και c_2 , των τυχαίων αριθμών στο διάστημα $[0,1]$ των παραμέτρων επιτάχυνσης των σωματιδίων του σμήνους r_1, r_2 , της συνάρτησης καταλληλότητας (fitness function) και το κριτήριο τερματισμού.
4. Χρησιμοποίησε την συνάρτηση καταλληλότητας για την ελαχιστοποίηση του σφάλματος RMSE του νευρωνικού δικτύου και υπολόγισε την καταλληλότητα κάθε

σωματιδίου και συγκρίνοντας την ατομική βέλτιστη τιμή $Pbest_{(i)}$ (Particle best) με την ολική βέλτιστη τιμή του σμήνους $Gbest_{(i)}$ (Global best), αν η τρέχουσα θέση είναι καλύτερη, τότε αντικατάστησε, αλλιώς παραμένει αμετάβλητη.

5. Ενημέρωσε την ταχύτητα και τη θέση των σωματιδίων του σμήνους σύμφωνα με τις εξισώσεις (4.5) και (4.6).

6. Εάν επιτευχθεί το κριτήριο τερματισμού, τερματίζεται ο αλγόριθμος. Αν δεν επιτευχθεί το κριτήριο τερματισμού, πήγαινε στο βήμα 2.

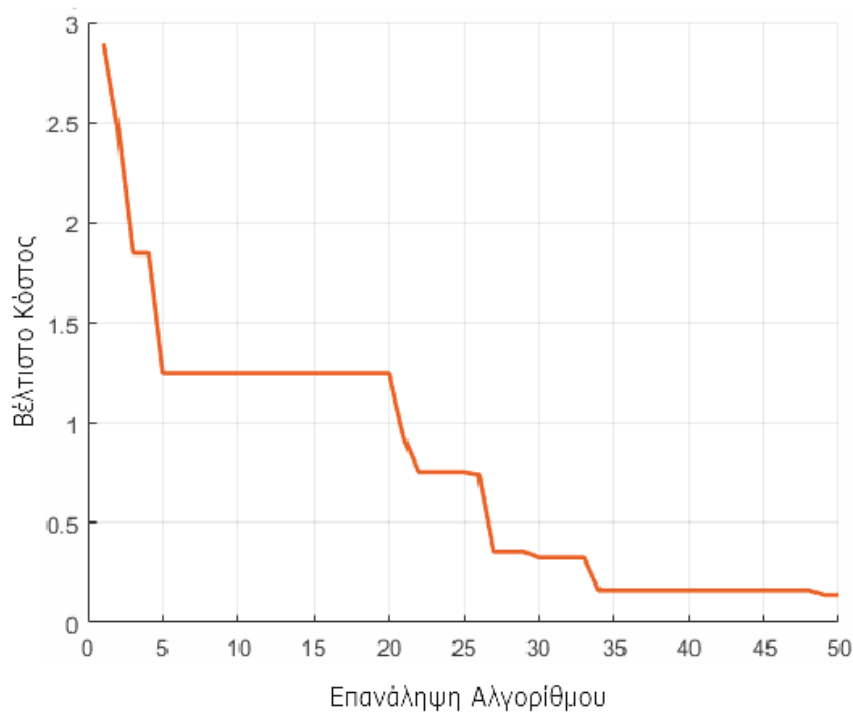
7. Υπολόγισε τα τελικά βέλτιστα βάρη και το τελικό σφάλμα RMSE.

Για την βελτιστοποίηση των παραμέτρων του Σμήνους Σωματιδίων, εφαρμόστηκε η μέθοδος αναζήτησης πλέγματος (Grid Search), που περιγράφηκε παραπάνω, για τις παραμέτρους που αφορούν: τον πληθυσμό των σωματιδίων του σμήνους, τους συντελεστές μάθησης c_1 , c_2 , τον συντελεστή βάρους αδράνειας (inertia weight coefficient). Οι πληθυσμοί των σωματιδίων σμήνους που εξετάστηκαν ήταν 20, 30, 40, 50, που είναι οι πιο συνηθισμένες τιμές σύμφωνα με τη βιβλιογραφία (Fourie and Groenwold, 2002). Οι συντελεστές μάθησης c_1 , c_2 , ορίστηκαν στην τιμή 2.0, όπως προτείνεται από τους Kennedy and Eberhart (1995). Οι τιμές του συντελεστή βάρους αδράνειας (inertia weight coefficient) εξετάστηκαν στο διάστημα [0.8, 1.4] με βήμα 0.1, όπως προτείνεται από τους Shi and Eberhart, (1998).

Οι παράμετροι που παρήγαγαν τα βέλτιστα αποτελέσματα ήταν: μέγεθος πληθυσμού: $P = 40$, βάρος αδράνειας $w = 0.90$, συντελεστής προσωπικής μάθησης $c_1 = 2.0$, ολικός συντελεστής εκμάθησης $c_2 = 2.0$. Ο πίνακας 4.2, παρουσιάζει τα πέντε καλύτερα μοντέλα του αλγορίθμου βελτιστοποίησης PSO, που βρέθηκαν χρησιμοποιώντας διαφορετικές τοπολογίες παραμέτρων. Το καλύτερο μοντέλο ήταν το Μοντέλο 3. Η παρακάτω εικόνα, αναπαριστά τη διαδικασία βελτιστοποίησης με την εφαρμογή του Σμήνους Σωματιδίων απεικονίζοντας την μεταβολή του καλύτερου κόστους συναρτήσεως επαναλήψεων (iterations) του αλγορίθμου PSO για το βέλτιστο μοντέλο, το Μοντέλο 3. Η Ρίζα του Μέσου Τετραγωνικού Σφάλματος (RMSE) του βέλτιστου νευρωνικού δικτύου που εκπαιδεύτηκε με την χρήση του βέλτιστου Σμήνους Σωματιδίων βρέθηκε 0.1159.

Πίνακας 4.2. Τα πέντε καλύτερα νευρωνικά μοντέλα του αλγορίθμου βελτιστοποίησης PSO που βρέθηκαν χρησιμοποιώντας διαφορετικές τοπολογίες παραμέτρων για την βελτιστοποίηση των βαρών του νευρωνικού δικτύου.

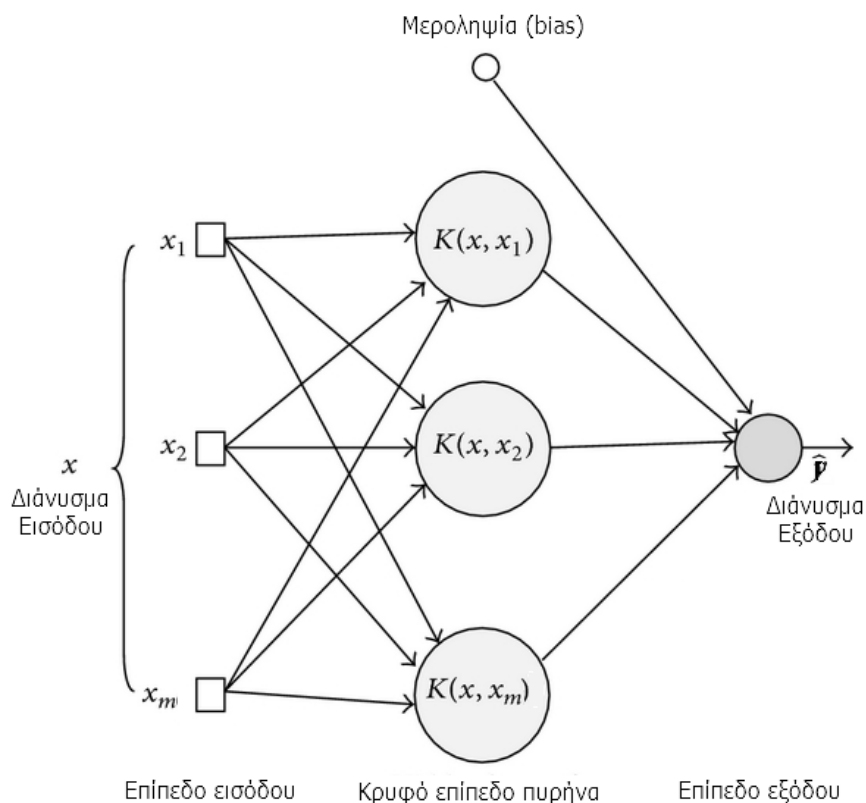
Μοντέλο	Πληθυσμός Σμήνους P	c_1	c_2	w	Βέλτιστο Κόστος (Best Cost)
<i>Μοντέλο 1</i>	20	2.0	2.0	0.80	0.3158
<i>Μοντέλο 2</i>	30	2.0	2.0	1.10	0.2811
<i>Μοντέλο 3</i>	40	2.0	2.0	0.90	0.1159
<i>Μοντέλο 4</i>	20	2.0	2.0	1.10	0.2125
<i>Μοντέλο 5</i>	40	2.0	2.0	1.00	0.2275



Εικόνα 4.5. Η γραφική παράσταση της μεταβολής του καλύτερου κόστους συναρτήσει των επαναλήψεων (iterations) του αλγορίθμου Βελτιστοποίησης Σμήνους Σωματιδίων κατά την βελτιστοποίηση των βαρών του νευρωνικού δικτύου.

4.5.6 Μοντέλα Μηχανών Διανυσμάτων Υποστήριξης στην Παλινδρόμηση

Σε αυτή τη μελέτη, χρησιμοποιήθηκαν επίσης οι Μηχανές Διανυσμάτων Υποστήριξης για την πρόβλεψη της εγκληματικότητας. Οι Μηχανές Διανυσμάτων Υποστήριξης έχουν προταθεί από τον Vapnik (1998), για την επίλυση των προβλημάτων παλινδρόμησης. Το σύνολο δεδομένων που χρησιμοποιήθηκε ήταν τα ίδιο με αυτό των μοντέλων νευρωνικών δικτύων. Για την βελτιστοποίηση της τοπολογίας των μοντέλων Μηχανών Διανυσμάτων Υποστήριξης που αναπτύχθηκαν εφαρμόστηκε ο αλγόριθμος βελτιστοποίησης Bayes (Martinez-Cantin, 2014, Mockus and Mockus, 1991). Αναπτύχθηκαν διάφορα μοντέλα SVM με την χρήση του λογισμικού MATLAB, με εφαρμογή διαφόρων ειδών πυρήνων και την επιλογή των βέλτιστων υπερ-παραμέτρων με τον αλγόριθμο Bayes, για την βελτιστοποίηση των αποτελεσμάτων.



Εικόνα 4.6. Η γραφική παράσταση της αρχιτεκτονικής των Μηχανών Διανυσμάτων υποστήριξης, όπου αναπαρίστανται το διάνυσμα εισόδου, το διάνυσμα εξόδου, τα επίπεδα εισόδου, εξόδου καθώς και το κρυφό επίπεδο.

Σε αυτή τη μελέτη, υλοποιήθηκαν οι Μηχανές Διανυσμάτων Υποστήριξης προσαρμοσμένες για την παλινδρόμηση, όπως προτάθηκε από τον Vapnik (1995), ο οποίος για την επίλυση των προβλημάτων παλινδρόμησης πρότεινε μια συνάρτηση απώλειας (loss function) που καθορίζει τον σωλήνα ε-ευαισθησίας (ε-insensitive tube). Η γενική συνάρτηση εκτίμησης στις Μηχανές Διανυσμάτων Υποστήριξης στην παλινδρόμηση ή αλλιώς επονομαζόμενες Μηχανές Διανυσμάτων Παλινδρόμησης (Support Vector Regression Machines) ορίζεται από την ακόλουθη εξίσωση:

$$f(x) = (w \cdot \Phi(x)) + b \quad (4.7)$$

Όπου $w \in R^n$, $b \in R$ και η Φ αναπαριστά έναν μη γραμμικό μετασχηματισμό από τον R^n σε χώρο υψηλής διαστάσεως. Το w αναπαριστά τον πίνακα των συντελεστών βάρους και το b τον συντελεστή πόλωσης (bias). Ο τελικός στόχος είναι να βρεθούν οι τιμές των w και b παραγόντων έτσι ώστε οι τιμές του x να μπορούν να καθοριστούν με την ελαχιστοποίηση του κινδύνου παλινδρόμησης, που εκφράζεται από την παρακάτω εξίσωση (Vapnik, 1995; Müller et al., 1999):

$$R_{reg}(f) = C \sum_{i=0}^1 \Gamma(f(x_i) - y_i) + \frac{1}{2} \|w\|^2 \quad (4.8)$$

Όπου το $\Gamma(\cdot)$ αναπαριστά την συνάρτηση κόστους και το C μία σταθερά, ενώ το διάνυσμα του βάρους w εκφράζεται από την σχέση:

$$w = \sum_{i=1}^{\lambda} (\alpha_i - \alpha_i^*) \Phi(x_i) \quad (4.9)$$

Αντικαθιστώντας την εξίσωση (4.9) στην εξίσωση (4.7), η γενική συνάρτηση εκτίμησης μπορεί να γραφεί ως:

$$f(x) = \sum_{i=1}^1 (\alpha_i - \alpha_i^*) (\Phi(x_i) \cdot \Phi(x)) + b \quad (4.10)$$

Στην εξίσωση (4.10) αν αντικαταστήσουμε το εσωτερικό γινόμενο με την συνάρτηση $k(x_i, x)$, που είναι η συνάρτηση πυρήνα, τότε θα έχουμε:

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) k(x_i, x) + b \quad (4.11)$$

Η συνάρτηση απώλειας ε -intensive είναι η πιο χρησιμοποιούμενη στην παλινδρόμηση με Μηχανές Διανυσμάτων Υποστήριξης και εμφανίζεται στην συνάρτηση κόστους $\Gamma(\cdot)$, σύμφωνα με την παρακάτω εξίσωση:

$$\Gamma(f(x) - y) = \begin{cases} |f(x) - y| - \varepsilon, & \text{για } |f(x) - y| \geq \varepsilon \\ 0 & \text{αλλιώς} \end{cases} \quad (4.12)$$

4.5.7 Βελτιστοποίηση Bayes στα Μοντέλα Μηχανών Διανυσμάτων Υποστήριξης

Ο αλγόριθμος βελτιστοποίησης Bayes εφαρμόστηκε για να επιτευχθεί η βελτιστοποίηση της επιλογής των υπερ-παραμέτρων (hyperparameters) των μοντέλων Μηχανών Διανυσμάτων Υποστήριξης που αναπτύχθηκαν (Martinez-Cantin, 2014, Mockus and Mockus, 1991). Η βελτιστοποίηση Bayes προτιμάται σε σχέση με την αναζήτηση πλέγματος ή άλλες μεθόδους, όταν η διάσταση των συνδυασμών των παραμέτρων είναι υψηλή, όπως συμβαίνει στην παρούσα περίπτωση. Ο αλγόριθμος μπορεί να έχει υψηλό υπολογιστικό κόστος, όμως με την εφαρμογή του σε μικρό σχετικά αριθμό επαναλήψεων εντοπίζει τον βέλτιστο συνδυασμό υπερ-παραμέτρων (Snoek et al., 2012).

Η βελτιστοποίηση Bayes είναι μία ακολουθιακή διαδικασία, που χρησιμοποιεί το θεώρημα Bayes και μέσω της εκ των προτέρων (prior) γνώσης των κατανομών πιθανοτήτων, υπολογίζουν τις εκ των υστέρων (posterior) πιθανότητες για την υπό βελτιστοποίηση συνάρτηση. Επίσης, χρησιμοποιήθηκε η διασταυρωμένη επικύρωση με 10 φύλλα (10-fold Cross-Validation) για την εύρεση του καλύτερου μοντέλου και την

αποφυγή της υπερεκπαίδευσης (Kohavi, 1995). Σύμφωνα με την μέθοδο διασταυρωμένης επικύρωσης με 10 φύλλων, το σύνολο εκπαίδευσης χωρίστηκε σε 10 υποσύνολα. Για καθένα από τα 10 υποσύνολα χρησιμοποιήθηκαν τα υπόλοιπα 9 για την εκπαίδευση του μοντέλου και το ένα για την δοκιμή του μοντέλου.

Σύμφωνα με τον αλγόριθμο βελτιστοποίησης Bayes, εφαρμόζεται μια Γκαουσιανή διαδικασία βελτιστοποίησης, υπολογίζοντας την αναμενόμενη πιθανότητα την οποία θα βελτιώσει η νέα επανάληψη του αλγορίθμου, λαμβάνοντας υπόψη τις τρέχουσες παρατηρήσεις. Η αναμενόμενη βελτίωση της συνάρτησης απόκτησης (acquisition function) υπολογίζει την αναμενόμενη τιμή της βελτίωσης της αντικειμενικής συνάρτησης (objective function) και ταυτόχρονα απορρίπτει τις τιμές που υπολογίζονται ότι θα καταλήγουν σε υψηλή τιμή της αντικειμενικής συνάρτησης (Snoek et al., 2012).

Η Bayesian βελτιστοποίηση ανήκει στους αλγορίθμους βελτιστοποίησης οι οποίοι ονομάζονται αλγόριθμοι ακολουθιακής βελτιστοποίησης βάσει μοντέλου (Sequential Model-Based Optimization, SMBO). Αυτή η κατηγορία αλγορίθμων εφαρμόζει τις προηγούμενες παρατηρήσεις δεδομένων της συνάρτησης απώλειας (loss function) προκειμένου να προσδιοριστεί το επόμενο σημείο που υπολογίζεται ως το βέλτιστο (Bergstra et al., 2011).

Η συνάρτηση απόκτησης (acquisition function) χρησιμοποιείται για τον υπολογισμό των δυνητικά υψηλών τιμών της αντικειμενικής συνάρτησης. Η μεγιστοποίηση της συνάρτησης απόκτησης χρησιμοποιείται για την επιλογή του επόμενου καλύτερου σημείου στο οποίο θα αξιολογηθεί η βελτίωση της αντικειμενικής συνάρτησης. Υπάρχουν διάφορων ειδών συναρτήσεις απόκτησης, όπως: η βασισμένη στη βελτίωση (improvement-based) συνάρτηση απόκτησης που θεωρείται από τις πιο κατάλληλες, η αναζήτηση της εντροπίας (Entropy search) και η πιθανότητα βελτίωσης (probability of improvement) (Snoek et al., 2012).

Η βασισμένη στη βελτίωση (improvement-based) συνάρτηση απόκτησης $I(\mathbf{x})$ που χρησιμοποιήθηκε στην παρούσα έρευνα, εκφράζεται με την ακόλουθη εξίσωση (Martinez-Cantin, 2014; Snoek et al., 2012):

$$I(\mathbf{x}) = \max\{0, f_{t+1}(\mathbf{x}) - f(\mathbf{x}^+)\} \quad (4.13)$$

όπου

$$\mathbf{x}^+ = \operatorname{argmax}_{\mathbf{x}_i \in \mathbf{x}_{1:t}} f(\mathbf{x}_i) \quad (4.14)$$

Η πιθανότητα βελτίωσης όσον αφορά μια κανονική εκ των υστέρων κατανομή (normal posterior distribution) μπορεί να εκτιμηθεί με τη συνήθη συνάρτηση πυκνότητας:

$$f(I) = \frac{1}{\sqrt{2\pi}\sigma(x)} \exp\left(-\frac{(\mu(x)-f(\mathbf{x}^+)-I)^2}{2\sigma^2(x)}\right) \quad (4.15)$$

και η αναμενόμενη βελτίωση $EI(\mathbf{x})$ (Expected Improvement) υπολογίζεται από το ολοκλήρωμα της συνάρτησης πυκνότητας (Jones et al., 1998; Mockus, 1994):

$$EI(\mathbf{x}) = \int_{I=0}^{I=\infty} I f(I) dI \quad (4.16)$$

$$EI(\mathbf{x}) = \int_{I=0}^{I=\infty} I \frac{1}{\sqrt{2\pi}\sigma(x)} \exp\left(-\frac{(\mu(x)-f(\mathbf{x}^+)-I)^2}{2\sigma^2(x)}\right) dI \quad (4.17)$$

$$EI(\mathbf{x}) = \sigma(x) \left[\frac{\mu(x)-f(\mathbf{x}^+)}{\sigma(x)} \Phi\left(\frac{\mu(x)-f(\mathbf{x}^+)}{\sigma(x)}\right) + \varphi\left(\frac{\mu(x)-f(\mathbf{x}^+)}{\sigma(x)}\right) \right] \quad (4.18)$$

Επομένως, η αναμενόμενη βελτίωση $EI(x)$ (Expected Improvement) εκφράζεται από την ακόλουθη εξίσωση:

$$EI(x) = \begin{cases} (\mu(x) - f(x^+))\Phi(Z) + \sigma(x)\varphi(Z) & \text{if } \sigma(x) > 0 \\ 0 & \text{if } \sigma(x) = 0 \end{cases} \quad (4.19)$$

όπου

$$Z = \frac{(\mu(x) - f(x^+))}{\sigma(x)} \quad (4.20)$$

Όπου τα $\varphi(\cdot)$ και $\Phi(\cdot)$ αναπαριστούν την συνάρτηση πυκνότητας πιθανότητας (Probability Density Function) και την αθροιστική συνάρτηση κατανομής (Cumulative Distribution Function) αντίστοιχα, της κανονικής κατανομής.

Ο αλγόριθμος βελτιστοποίησης Bayes περιγράφεται παρακάτω (Brochu et al., 2010; Shahriari et al., 2016):

1. Επανέλαβε για $n = 1, 2, \dots$
2. Επέλεξε νέο x_{n+1} βελτιστοποιώντας την συνάρτηση απόκτησης f
$$x_{n+1} = \operatorname{argmax}_{x_i} f(x; D_n)$$
3. Εφάρμοσε την αντικειμενική συνάρτηση για τον υπολογισμό του y_{n+1}
4. Αύξησε τα δεδομένα $D_{n+1} = \{D_n, (x_{n+1}, y_{n+1})\}$
5. Ενημέρωσε το στατιστικό μοντέλο
6. Τέλος επανάληψης

4.5.8 Βελτιστοποιημένα Μοντέλα Μηχανών Διανυσμάτων Υποστήριξης

Εφαρμόζοντας τον αλγόριθμο βελτιστοποίησης Bayes, όπως περιγράφηκε στην προηγούμενη ενότητα, υπολογίστηκαν οι υπερ-παράμετροι για την εύρεση των βέλτιστων μοντέλων Μηχανών Διανυσμάτων Υποστήριξης (ΜΔΥ) για καθένα από τους κάτωθι βασικούς πυρήνες, οι οποίοι θεωρούνται ως οι πιο συνήθεις στην υλοποίηση των

Μηχανών Διανυσμάτων Υποστήριξης στην παλινδρόμηση: γραμμικός, πολυωνυμικός, ακτινικός και γκαουσιανός χρησιμοποιώντας το λογισμικό MATLAB. Οι συναρτήσεις πυρήνων που εφαρμόστηκαν στον αλγόριθμο βελτιστοποίησης Bayes για την παλινδρόμηση, περιγράφονται παρακάτω.

Η συνάρτηση για τον γραμμικό πυρήνα (Linear Kernel) εκφράζεται από την σχέση:

$$K(x_i, x_j) = \langle x_i, x_j \rangle \quad (4.21)$$

Η συνάρτηση για τον πολυωνυμικό πυρήνα (Polynomial Kernel) που θα χρησιμοποιηθεί εκφράζεται από την σχέση:

$$K(x_i, x_j) = (\sigma(\gamma) * x_i^T x_j + 1)^d, \quad \sigma > 0 \quad (4.22)$$

Όπου το d αναπαριστά τον βαθμό του πολυωνύμου, η συνάρτηση σ (σίγμα) εξαρτάται από την παράμετρο γ (γάμμα) που ελέγχει το πλάτος του πυρήνα. Η σχέση του γ με το σ εκφράζεται σύμφωνα με την εξίσωση:

$$\gamma = \frac{1}{\sigma^2} \quad (4.23)$$

Οι πυρήνες RBF (Radial Basis Function) είναι τύποι πυρήνων όπου ένα μέτρο απόστασης εξομαλύνεται από μια ακτινική (εκθετική συνάρτηση). Η συνάρτηση ακτινικού πυρήνα RBF εκφράζεται από με την εξίσωση:

$$K(x_i, x_j) = e^{-\sigma(\gamma)\|x_i - x_j\|^2}, \quad \sigma > 0 \quad (2.24)$$

Η συνάρτηση του γκαουσιανού πυρήνα που αποτελεί μία παραλλαγή του ακτινικού πυρήνα εκφράζεται από με την εξίσωση:

$$K(x_i, x_j) = e^{-\frac{1}{2\sigma(\gamma)^2}\|x_i - x_j\|^2}, \quad \sigma > 0 \quad (2.25)$$

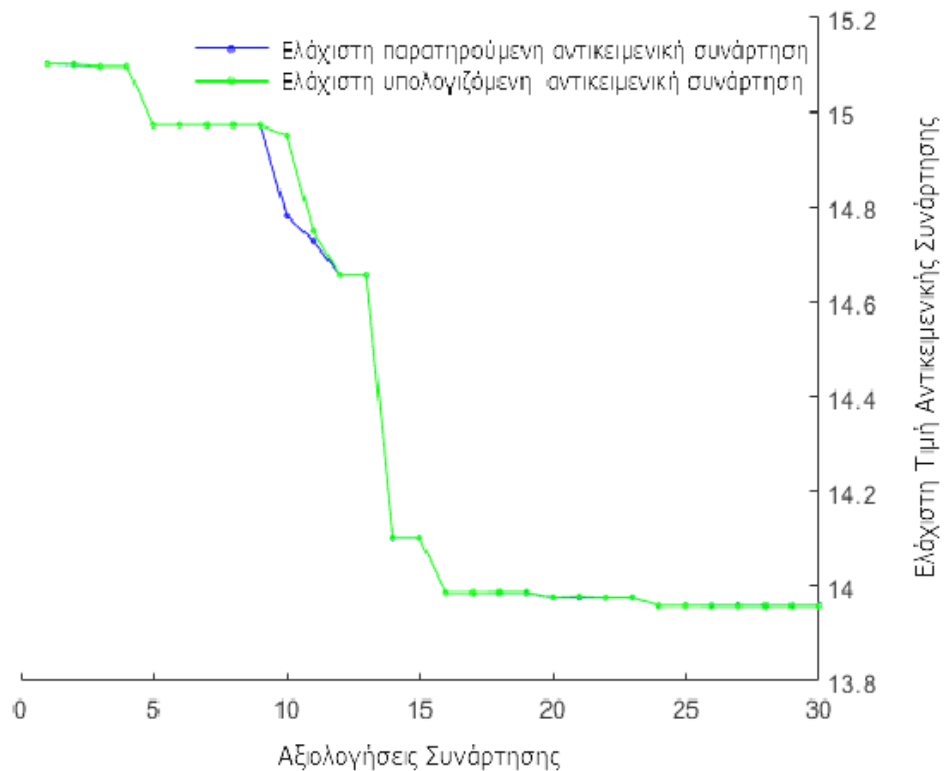
Για την αξιολόγηση των μοντέλων χρησιμοποιήθηκε η ελάχιστη τιμή της αντικειμενικής συνάρτησης. Με τον αλγόριθμο βελτιστοποίησης Bayes βελτιστοποιήθηκαν οι υπερ-παράμετροι των μοντέλων Μηχανών Διανυσμάτων Υποστήριξης: Η τιμή κόστους (C), η παράμετρος epsilon (ϵ), η παράμετρος γάμμα (γ) των πυρήνων εκτός του γραμμικού, ο βαθμός του πολυωνύμου για τον πολυωνυμικό πυρήνα. Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον Γκαουσιανό πυρήνα παρουσιάζονται στον παρακάτω πίνακα.

Πίνακας 4.3. Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον Γκαουσιανό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμενική Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζ.	C	γ	ϵ
1	15.104	15.104	15.104	109.940	0.006	46.862
2	15.102	15.102	15.103	0.011	0.004	7386.700
3	15.098	15.098	15.098	21.651	20.659	7.444
4	15.103	15.098	15.098	0.005	0.115	5.586
5	14.975	14.975	14.975	38.471	283.460	6.508
6	15.073	14.975	14.975	6.150	740.720	955.110
7	15.102	14.975	14.975	0.902	221.230	26289.000
8	15.103	14.975	14.975	0.003	323.010	2.922
9	15.102	14.975	14.975	46.288	340.010	10064.000
10	14.783	14.783	14.953	108.670	306.640	8.325
11	14.730	14.730	14.751	135.310	314.250	8.391
12	14.658	14.658	14.658	178.040	310.200	6.999
13	15.109	14.658	14.658	341.190	0.005	6.359
14	14.103	14.103	14.103	519.110	581.660	1.983
15	14.292	14.103	14.103	971.320	221.250	1.918

Πίνακας 4.3. (συνέχεια) Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον Γκαουσιανό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμενική ή Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογίζ.	C	γ	ε
16	13.987	13.987	13.987	960.070	986.880	5.360
17	15.102	13.987	13.987	910.970	0.695	89372.000
18	13.989	13.987	13.988	965.390	978.220	2.031
19	14.027	13.987	13.988	999.740	720.440	4.443
20	13.977	13.977	13.978	819.420	980.310	341.430
21	14.014	13.977	13.979	710.000	957.170	14.520
22	15.102	13.977	13.978	968.790	984.310	20529.000
23	13.978	13.977	13.978	972.100	985.190	94.840
24	13.960	13.960	13.960	995.230	989.640	204.290
25	14.022	13.960	13.960	994.820	762.680	246.870
26	14.005	13.960	13.959	644.890	981.850	159.810
27	13.978	13.960	13.959	984.680	993.870	21.569
28	15.102	13.960	13.959	0.001	2.187	180390.000
29	15.102	13.960	13.959	720.610	0.001	181610.000
30	15.103	13.960	13.959	0.001	0.001	2.683



Εικόνα 4.7. Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον Γκαουσιανό πυρήνα (Gaussian Kernel) Μηχανών Διανυσμάτων Υποστήριξης.

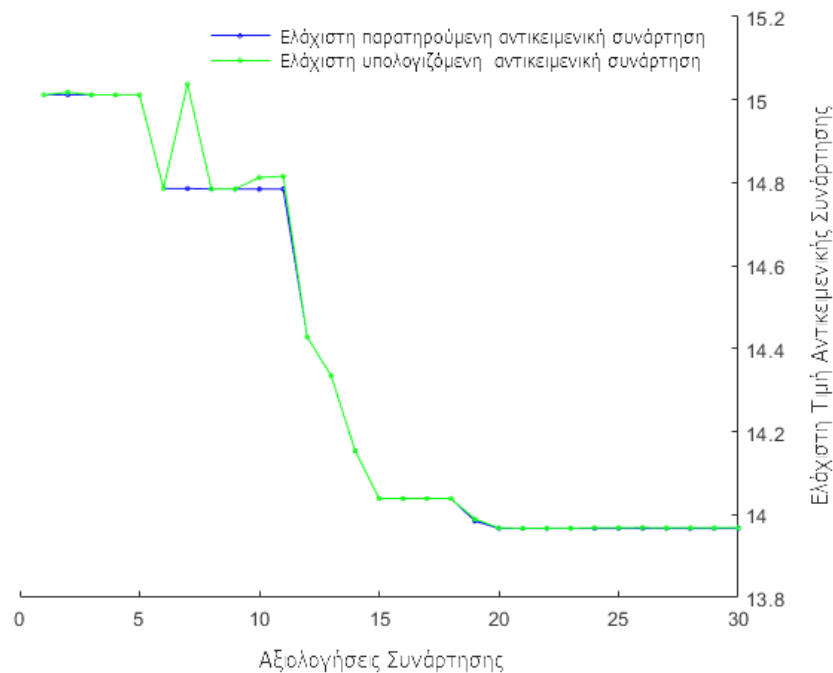
Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον ακτινικό πυρήνα, παρουσιάζονται στον παρακάτω πίνακα.

Πίνακας 4.4. Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον ακτινικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμενική Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζ.	C	γ	ϵ
1	15.012	15.012	15.012	209.060	37.288	12.816
2	15.118	15.012	15.018	0.016	34.659	4183.100
3	15.096	15.012	15.012	0.002	72.259	3.488
4	15.096	15.012	15.012	0.001	0.002	5.111

Πίνακας 4.4. (συνέχεια) Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον ακτινικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμ. Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζ.	C	γ	ε
5	15.089	15.012	15.012	998.530	7.262	59419.000
6	14.786	14.786	14.786	966.860	59.575	1.968
7	15.067	14.786	15.038	203.220	12.492	1.912
8	14.785	14.785	14.785	966.170	58.947	86.291
9	14.836	14.785	14.785	986.730	45.098	2.362
10	15.089	14.785	14.812	930.840	64.293	8034.800
11	15.103	14.785	14.815	968.770	0.001	19.203
12	14.428	14.428	14.428	930.220	144.730	16.843
13	14.334	14.334	14.334	925.450	201.710	14.719
14	14.154	14.154	14.154	950.520	415.690	13.670
15	14.039	14.039	14.039	961.850	919.610	15.756
16	14.047	14.039	14.039	701.930	993.820	6.139
17	14.123	14.039	14.039	298.100	996.450	16.282
18	15.089	14.039	14.039	0.003	0.001	183130.000
19	13.984	13.984	13.989	909.170	985.340	298.200
20	13.967	13.967	13.968	997.210	997.560	419.700
21	15.089	13.967	13.967	966.630	986.800	8026.500
22	14.491	13.967	13.967	97.748	972.930	447.210
23	14.574	13.967	13.967	76.735	997.710	1.972
24	13.987	13.967	13.968	970.900	914.180	359.130
25	15.089	13.967	13.968	912.100	0.001	159210.000
26	14.050	13.967	13.969	887.860	820.820	54.974
27	13.979	13.967	13.968	927.310	969.990	413.640
28	15.098	13.967	13.968	0.478	0.001	1087.200
29	15.089	13.967	13.968	0.001	589.440	188080.000
30	14.035	13.967	13.968	967.340	958.420	3.711



Εικόνα 4.8. Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον ακτινικό πυρήνα (RBF Kernel) Μηχανών Διανυσμάτων Υποστήριξης.

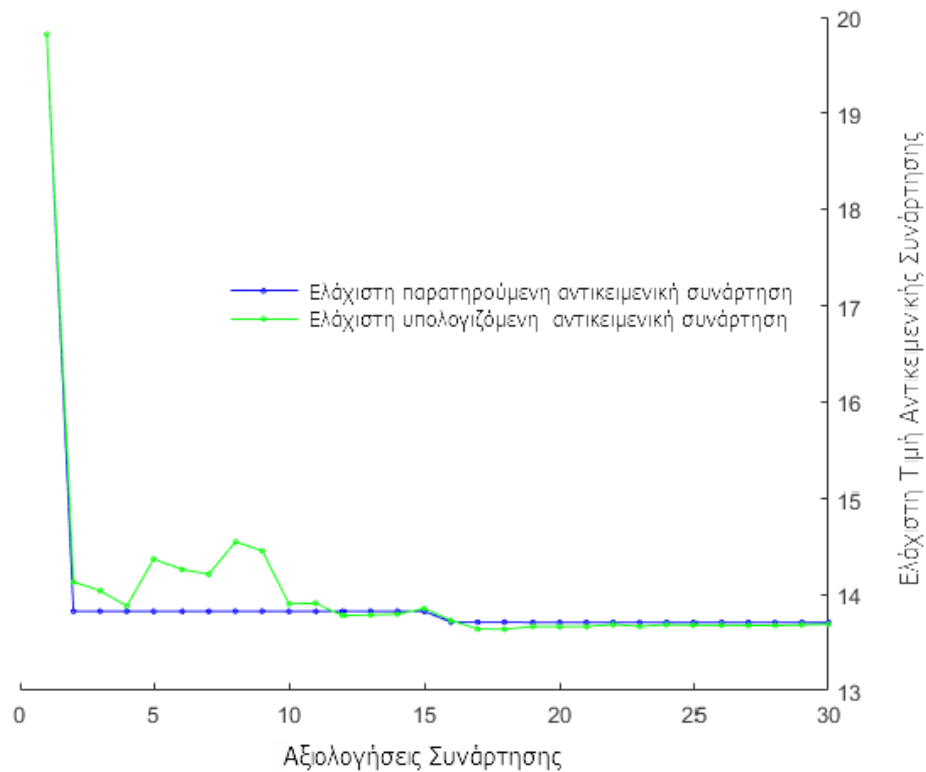
Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον πολωνυμικό πυρήνα παρουσιάζονται στον παρακάτω πίνακα.

Πίνακας 4.5. Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον πολωνυμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης. Όπου d , ο βαθμός του πολωνύμου.

Επανάληψη Αλγορίθμου	Αντικειμ. Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζ.	C	γ	ϵ	d
1	19.821	19.821	19.821	0.006	71.788	8.881	2
2	13.828	13.828	14.135	396.360	969.020	104.480	4
3	15.086	13.828	14.044	0.330	0.998	12414.000	2

Πίνακας 4.5. (συνέχεια) Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον πολυωνυμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης. Όπου d, ο βαθμός του πολυωνύμου.

Επανάληψη Αλγορίθμου	Αντικειμ. Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζ.	C	γ	ε	d
4	15.086	13.828	13.882	478.210	0.044	51.504	2
5	15.086	13.828	14.374	290.190	17.230	299.080	3
6	14.204	13.828	14.264	992.550	974.990	511.970	3
7	15.086	13.828	14.216	3.678	976.530	110210.000	6
8	15.857	13.828	14.555	957.010	954.510	2.937	4
9	15.086	13.828	14.461	992.960	7.024	14803.000	5
10	14.502	13.828	13.907	997.100	877.730	571.370	3
11	15.066	13.828	13.912	994.100	942.400	60.309	2
12	13.833	13.828	13.784	527.810	996.510	1530.700	7
13	13.832	13.828	13.791	84.370	997.510	1766.900	5
14	15.086	13.828	13.798	8.165	0.013	157720.000	7
15	15.086	13.828	13.858	239.220	972.300	185120.000	5
16	13.715	13.715	13.733	35.674	990.110	32.231	2
17	13.714	13.714	13.643	131.410	970.900	146.720	6
18	15.086	13.714	13.641	6.656	3.413	41.108	2
19	13.711	13.711	13.669	79.470	998.000	175.980	3
20	15.086	13.711	13.668	913.040	0.001	183370.000	3
21	13.724	13.711	13.669	7.304	937.060	2.613	5
22	13.716	13.711	13.690	6.377	965.510	80.270	4
23	13.720	13.711	13.671	11.823	997.240	9.837	2
24	15.086	13.711	13.687	0.222	0.001	131580.000	7
25	13.761	13.711	13.683	208.010	999.550	566.680	3
26	15.086	13.711	13.684	11.001	0.001	3.009	5
27	13.716	13.711	13.681	18.369	996.140	80.077	6
28	13.723	13.711	13.680	3.804	994.510	14.247	2
29	13.725	13.711	13.687	6.608	978.270	13.885	4
30	13.737	13.711	13.692	205.750	996.710	352.060	2



Εικόνα 4.9. Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον πολωνυμικό πυρήνα (Polynomial Kernel) Μηχανών Διανυσμάτων Υποστήριξης.

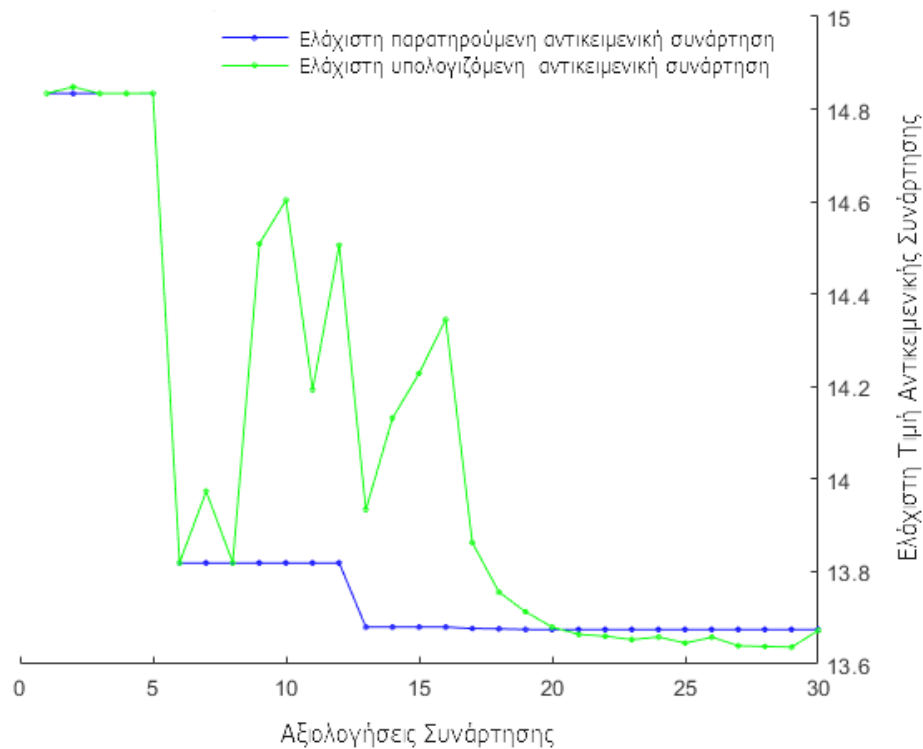
Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον γραμμικό πυρήνα παρουσιάζονται στον παρακάτω πίνακα.

Πίνακας 4.6. Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον γραμμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη	Αντικειμενική Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζόμενη	C	ε
1	14.834	14.834	14.834	0.165	518.420
2	15.090	14.834	14.848	0.001	23136.000
3	15.090	14.834	14.834	505.560	38875.000

Πίνακας 4.6. (συνέχεια) Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον γραμμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμενική Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζόμενη	C	ε
4	18.501	14.834	14.834	0.825	1.973
5	15.090	14.834	14.834	54.131	102980.000
6	13.818	13.818	13.819	15.342	1516.900
7	14.141	13.818	13.973	248.810	2625.800
8	15.066	13.818	13.819	0.094	1669.400
9	15.689	13.818	14.509	277.060	577.490
10	14.878	13.818	14.603	126.450	4276.300
11	14.054	13.818	14.193	1.070	30.918
12	15.090	13.818	14.506	382.610	8717.800
13	13.680	13.680	13.934	96.143	10.536
14	14.975	13.680	14.132	0.440	1.969
15	14.489	13.680	14.228	2.695	28.580
16	15.095	13.680	14.345	0.002	22.085
17	13.677	13.677	13.863	994.180	34.625
18	13.676	13.676	13.756	990.810	42.521
19	13.675	13.675	13.713	996.120	21.669
20	13.674	13.674	13.680	996.600	21.343
21	13.676	13.674	13.664	952.070	157.110
22	13.677	13.674	13.660	989.540	70.873
23	13.677	13.674	13.653	993.730	23.449
24	13.700	13.674	13.659	0.005	497.040
25	13.686	13.674	13.645	44.949	461.360
26	13.678	13.674	13.658	1.878	161.740
27	13.692	13.674	13.640	0.671	557.630
28	13.675	13.674	13.638	987.490	12.320
29	13.769	13.674	13.637	0.001	1278.500
30	13.677	13.674	13.673	991.590	78.961



Εικόνα 4.10. Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον γραμμικό πυρήνα (Linear Kernel) Μηχανών Διανυσμάτων Υποστήριξης.

Όπως παρατηρούμε το καλύτερο μοντέλο των Μηχανών Διανυσμάτων Υποστήριξης (SVM) που βρέθηκε με τον αλγόριθμο Bayes είναι αυτό του γραμμικού πυρήνα με τα χαρακτηριστικά υπερπαραμέτρων που αναπαρίστανται στον παρακάτω πίνακα.

Πίνακας 4.7. Τα βέλτιστα μοντέλα των Μηχανών Διανυσμάτων Υποστήριξης (SVM) που βρέθηκαν με τον αλγόριθμο Bayes για κάθε πυρήνα σύμφωνα με την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Πυρήνας Μοντέλου	Τιμές Αντικειμενικής Συνάρτησης
Γραμμικός	13.674
Πολυωνυμικός	13.711
Γκαουσιανός	13.960
Ακτινικός	13.967

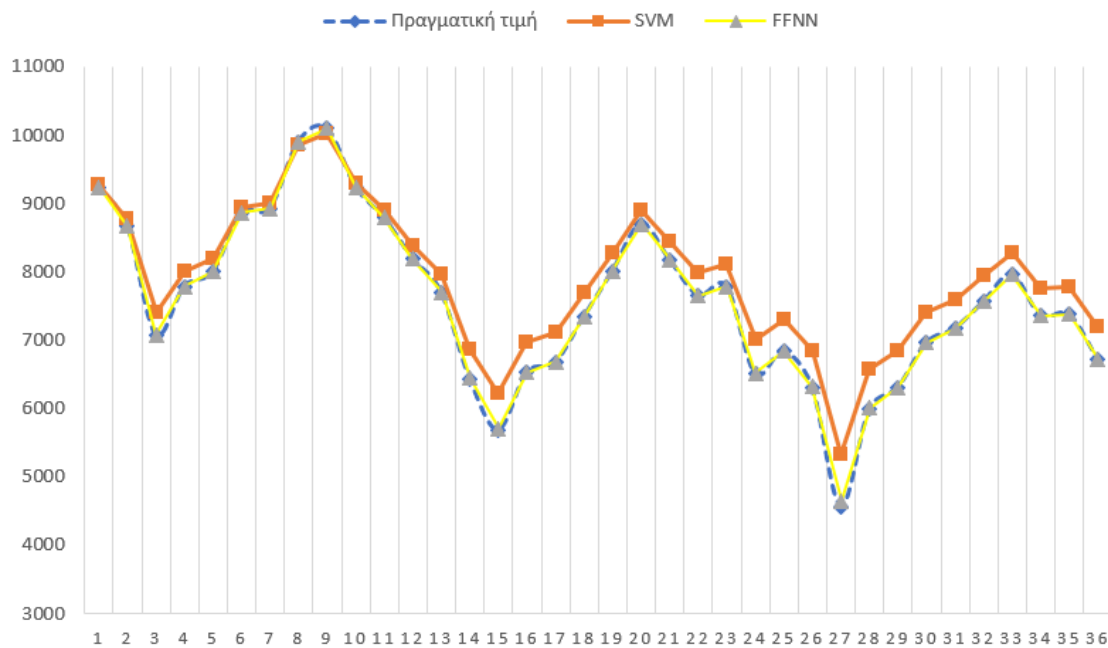
4.5.9 Σύγκριση Προβλέψεων Βέλτιστων Μοντέλων SVM και FFNN

Στην ενότητα αυτή, πραγματοποιείται σύγκριση των αποτελεσμάτων των προβλέψεων των βέλτιστων μοντέλων των Μηχανών Διανυσμάτων Υποστήριξης (SVM) και των νευρωνικών δικτύων εμπρόσθιας διάδοσης (FFNN). Όπως φαίνεται από τα αποτελέσματα, το βέλτιστο μοντέλο είναι αυτό των νευρωνικών δικτύων. Στον παρακάτω πίνακα, συγκρίνονται οι δύο προσεγγίσεις που εφαρμόστηκαν στην μελέτη αυτή.

Πίνακας 4.8. Συγκριτικός πίνακας που απεικονίζει, τα βελτιστοποιημένα μοντέλα των Μηχανών Διανυσμάτων Υποστήριξης (SVM) με τον αλγόριθμο Bayes και των νευρωνικών δικτύων εμπρόσθιας διάδοσης (FFNN) με τον αλγόριθμο PSO.

Μοντέλο	Μέθοδος Βελτιστοποίησης	Σφάλμα MSE	Σφάλμα RMSE
<i>SVM-BAYES</i>	Αλγόριθμος Bayes	10.9945	3.3158
<i>FFNN-PSO</i>	Αλγόριθμος PSO	0.01343	0.1159

Η παρακάτω εικόνα απεικονίζει τις προβλέψεις του βέλτιστου μοντέλου των Μηχανών Διανυσμάτων Υποστήριξης (πορτοκαλί χρώμα γραμμής), του βέλτιστου μοντέλου των τεχνητών νευρωνικών δικτύων (κίτρινο χρώμα γραμμής) και της πραγματικής τιμής (μπλε χρώμα γραμμής) για τα τελευταία 36 χρονικά βήματα (μήνες) του δείγματος από 01/01/2013 έως 31/12/2015. Όπως φαίνεται από τα αποτελέσματα, το βέλτιστο μοντέλο είναι αυτό των νευρωνικών δικτύων.



Εικόνα 4.11. Συγκριτική γραφική παράσταση των προβλέψεων του βέλτιστου μοντέλου νευρωνικού δικτύου, του βέλτιστου μοντέλου SVM και της πραγματικής τιμής για τα τελευταία 36 χρονικά βήματα (μήνες) του δείγματος.

4.6 Συμπεράσματα και Συζήτηση

Η υιοθέτηση τεχνικών τεχνητής νοημοσύνης στην πρόβλεψη της εγκληματικότητας μπορεί να αποβεί πολύτιμη στη δημόσια διοίκηση. Στην παρούσα έρευνα, χρησιμοποιήθηκαν μοντέλα τεχνητών νευρωνικών δικτύων και μηχανές διανυσμάτων υποστήριξης, για να προβλεφθεί η εγκληματικότητα αδικημάτων ιδιοκτησίας σε μηνιαία βάση, λαμβάνοντας υπόψη τον παράγοντα της ανεργίας ως σημαντικού κοινωνικοοικονομικού παράγοντα που επηρεάζει τα ποσοστά της εγκληματικότητας σύμφωνα με την υπάρχουσα βιβλιογραφία (Bertiz Armenico, et al., 2015; Blum and Li, 1991; Bogomolov et al., 2014).

Διάφοροι αλγόριθμοι μάθησης και τοπολογίες δικτύου εξετάστηκαν με την μέθοδο αναζήτησης πλέγματος, προκειμένου να κατασκευαστεί το καλύτερο μοντέλο τεχνητού

νευρωνικού δικτύου, το οποίο βελτιστοποιήθηκε χρησιμοποιώντας μία μετα-ευριστική μέθοδο, την Βελτιστοποίηση Σμήνους Σωματιδίων (Particle Swarm Optimization, PSO), η οποία αρχικά προτάθηκε από τους Eberhart and Kennedy (1995), ώστε να επιτευχθούν καλύτερες προβλέψεις.

Επιλέχθηκε ο αλγόριθμος Βελτιστοποίησης Σμήνους Σωματιδίων, διότι σε σύγκριση με άλλες μεθόδους βελτιστοποίησης νευρωνικών δικτύων, όπως ο γενετικός αλγόριθμος πλεονεκτεί ως προς την ευκολία εφαρμογής και την γρήγορη ρύθμιση λίγων παραμέτρων (Ou and Lin, 2006). Επίσης, ο αλγόριθμος Βελτιστοποίησης Σμήνους Σωματιδίων υπερέχει σε υπολογιστική ικανότητα έναντι του γενετικού αλγόριθμου (Hassan et al., 2005).

Επίσης, κατασκευάστηκαν μοντέλα προβλέψεις βασισμένα στις μηχανές διανυσμάτων υποστήριξης και χρησιμοποιήθηκε ο αλγόριθμος βελτιστοποίησης Bayes για μεγαλύτερη ακρίβεια των αποτελεσμάτων. Τα αποτελέσματα των δύο τεχνικών τεχνητής νοημοσύνης συγκρίθηκαν και βρέθηκε ότι καλύτερα αποτελέσματα παρήγαγαν τα τεχνητά νευρωνικά δίκτυα εμπρόσθιας διάδοσης.

Είναι σημαντικό, το ότι, με το εκπαιδευμένο μοντέλο που κατασκευάστηκε, έχει μοντελοποιηθεί η σχέση των παραμέτρων εισόδου με την παράμετρο εξόδου, και μπορεί να χρησιμοποιηθεί για τις συγκεκριμένες παραμέτρους που μελετήθηκε, ως εργαλείο διοίκησης, από τις αρμόδιες αρχές, για να δοκιμαστούν διαφορετικά σενάρια με δεδομένα παραγόντων εισόδου του ‘σήμερα’ (τρέχων μήνας) που επηρεάζουν την τιμή εξόδου (πρόβλεψη εξαρτημένης μεταβλητή), της εγκληματικότητας του ‘αύριο’ (επόμενος μήνας). Το στοιχείο αυτό, μπορεί να βοηθήσει αρκετά, αφού οι δημόσιες και δικαστικές αρχές έχουν το χρόνο μέχρι να επέλθει το επόμενο χρονικό βήμα, για να λάβουν τα κατάλληλα μέτρα στην κατεύθυνση της μείωσης της εγκληματικότητας. Επίσης, το μοντέλο πρόβλεψης δημιουργήθηκε με το λογισμικό MATLAB και μοντελοποιήθηκε με το Simulink (Dabney and Harman, 2004), μια επέκταση του Matlab για την μοντελοποίηση συστημάτων, και είναι επαναχρησιμοποιήσιμο.

Η πρόβλεψη της εγκληματικότητας μπορεί να σχετίζεται με την λήψη αποφάσεων για την καταπολέμηση της εγκληματικότητας, και την καλύτερη οργάνωση των αστυνομικών επιχειρήσεων (Camacho-Collados and Liberatore, 2015; Junior et al., 2017; Tambe et al., 2016). Σε σύγκριση με άλλες έρευνες (Alwee et al., 2013a, 2013b; Bertiz Armenico, et al., 2015) τα αποτελέσματα της παρούσας μελέτης έδειξαν μεγαλύτερη ακρίβεια πρόβλεψης της εγκληματικότητας, σύμφωνα με τη σύγκριση της Ρίζας του Μέσου Τετραγωνικού Σφάλματος (RMSE). Επίσης, η έρευνα αυτή μπορεί να συμβάλει στην βελτίωση της νέας μορφής αστυνόμευσης που εισήχθη τα τελευταία χρόνια, την «προβλεπτική αστυνόμευση» (predictive policing), η οποία επικεντρώνεται στο τι πρόκειται να συμβεί και για την λήψη αποφάσεων στο πώς θα αντιμετωπίσει αποτελεσματικά το έγκλημα με βάση τις προβλεφθείσες πληροφορίες (Beck and McCue, 2009; Perry, 2013).

Η πρόβλεψη της εγκληματικότητας στο αστικό περιβάλλον είναι ένας σημαντικός παράγοντας που μπορεί να συμβάλλει σε μία ασφαλέστερη ζωή, βοηθώντας τις Αρχές να λάβουν προληπτικά μέτρα για τη βελτίωση της διαχείρισης της δημόσιας ασφάλειας. Συγκεκριμένα, η πρόβλεψη της εγκληματικότητας στις αστικές περιοχές μπορεί να είναι πολύτιμη και εποικοδομητική στη διαδικασία λήψης αποφάσεων σχετικά με τις επιχειρήσεις των αστυνομικών και δικαστικών αρχών καθώς και στο σχεδιασμό στρατηγικών και πρακτικών κατά της εγκληματικότητας.

ΚΕΦΑΛΑΙΟ 5: Η Τεχνητή Νοημοσύνη και τα GIS στην Χωρική Πρόβλεψη της Εγκληματικότητας στη Δημόσια και Δικαστική Διοίκηση

5.1 Συνοπτική Περιγραφή της Έρευνας

Η υιοθέτηση νέων εξελιγμένων μεθόδων και τεχνικών για την πρόβλεψη της εγκληματικότητας μπορεί να είναι πολύ ελπιδοφόρα για την εφαρμογή στρατηγικών για την πρόληψη της. Σε αυτή την έρευνα αναπτύχθηκε μια μεθοδολογία για την χωρική πρόβλεψη της εγκληματικότητας στο αστικό περιβάλλον, χρησιμοποιώντας μεθόδους τεχνητής νοημοσύνης, ώστε τελικά να αναπτυχθούν τα καλύτερα μοντέλα πρόβλεψης σε συνδυασμό με μεθόδους χωρικής ανάλυσης, όπως: ανάλυση πλησιέστερων γειτόνων για τον προσδιορισμό των χωρικών μοντέλων και της κατανομής της εγκληματικότητας, ανάλυση θερμών σημείων, προκειμένου να προσδιοριστούν οι στατιστικά σημαντικές περιοχές με υψηλή συγκέντρωση εγκληματικότητας. Η προτεινόμενη μεθοδολογία επικεντρώνεται στη χωρική πρόβλεψη εγκληματικότητας, και βασίζεται στην κατασκευή και τη σύγκριση πολλαπλών μοντέλων νευρωνικών δικτύων που αναπτύχθηκαν ερευνώντας διαφόρους τύπους αλγορίθμων μάθησης και πολλαπλές αρχιτεκτονικές δικτύου.

5.2 Εισαγωγή – Σχετική Έρευνα

Η αξία της πρόβλεψης της εγκληματικότητας έχει γίνει πολύ σημαντική στη δημόσια και δικαστική διοίκηση, στη λήψη αποφάσεων με στόχο τη βελτίωση της

αποτελεσματικότητας των μεθόδων πρόληψης της εγκληματικότητας και την ανάπτυξη κατάλληλων στρατηγικών για την αντιμετώπιση της εγκληματικότητας. Έχουν αναπτυχθεί νέες καινοτόμες στρατηγικές διαχείρισης της εγκληματικότητας που επικεντρώνονται στην πρόβλεψη της εγκληματικότητας ώστε να ενισχύσουν την αποτελεσματικότητα των αστυνομικών επιχειρήσεων (Kennedy et al., 2011, Kennedy and Van Brunschot, 2009). Η ενσωμάτωση των γεωγραφικών συστημάτων πληροφοριών και των τεχνολογιών της πληροφορίας στην πρόβλεψη της εγκληματικότητας και στη δημόσια διοίκηση μελετήθηκε από διάφορους ερευνητές (Kouziokas, 2016a, 2016c, 2016d, Wang et al., 2013, Leitner, 2013). Επιπλέον, διάφορες μελέτες έχουν εφαρμόσει τεχνητά νευρωνικά δίκτυα στην πρόβλεψη της εγκληματικότητας στη δημόσια διοίκηση τα τελευταία χρόνια (Kouziokas, 2016b, Kouziokas et al., 2016; Olligschlaeger, 1997). Διάφορες τεχνικές και μεθοδολογίες έχουν αναπτυχθεί από ερευνητές για την πρόβλεψη της εγκληματικότητας, προκειμένου να προβλεφθούν οι περιοχές με υψηλό ποσοστό εγκληματικότητας. Αυτές οι μεθοδολογίες περιλαμβάνουν τεχνικές αυτοπαλινδρόμησης (Shoesmith, 2013), τεχνικές βασισμένες στα γεωγραφικά συστήματα πληροφοριών (Helbich et al., 2013), στην τεχνητή νοημοσύνη (Olligschlaeger, 1997) και τεχνικές εξόρυξης δεδομένων (McCue, 2014, Chen et al., 2004).

Η πλειοψηφία των ερευνητών επικεντρώθηκε στις τεχνικές που βασίζονται στα GIS χρησιμοποιώντας χωρικές αναλύσεις (Zhou et al., 2014, Andresen, 2006) και μεθόδους χαρτογράφησης θερμών σημείων (θεματική χαρτογράφηση, εκτίμηση πυκνότητας πυρήνα (KDE)) (Chainey et al., 2002; Levine, 2006; Nakaya and Yano, 2010; Ratcliffe, 2004) για τον εντοπισμό και την πρόβλεψη των χωρικών προτύπων εγκληματικότητας καθώς και των περιοχών με υψηλή συγκέντρωση περιστατικών εγκλημάτων (Chainey et al., 2008; Levine, 2008; Ratcliffe and McCullagh, 1999).

Το οργανωμένο έγκλημα εφαρμόζει διάφορες τεχνικές για την εκτέλεση εγκληματικών πράξεων και παράνομων δραστηριοτήτων, ανάλογα με διάφορους παράγοντες (Abadinsky, 2012, Kleemans et al., 2012). Τα τεχνητά νευρωνικά δίκτυα υιοθετήθηκαν στην πρόβλεψη της εγκληματικότητας από διάφορους ερευνητές σε συνδυασμό με την ανάλυση θερμών σημείων και άλλες τεχνικές ομαδοποίησης (Cohen et al., 2007, Corcoran et al., 2003, Olligschlaeger, 1997).

Ο Olligschlaeger (1997), περιγράφει πως τα τεχνητά νευρωνικά δίκτυα μπορούν να ενσωματωθούν σε ένα σύστημα έγκαιρης προειδοποίησης που εφαρμόζεται από το Γραφείο Αστυνομίας του Πίτσμπουργκ το οποίο βασίζεται σε γεωγραφικά συστήματα πληροφοριών για την χωρική πρόβλεψη περιστατικών εγκλημάτων σε σημεία αγορών ναρκωτικών. Χωρικά δεδομένα χρησιμοποιήθηκαν ως δεδομένα εισόδου για την εκπαίδευση του τεχνητού νευρωνικού δικτύου και η έξοδος απεικονίστηκε σε χάρτη που δείχνει τις περιοχές που προβλέφθηκαν με υψηλό αριθμό 911 κλήσεων για εγκλήματα ναρκωτικών.

Οι Cohen et al. (2007), δημιούργησαν ένα μοντέλο – δείκτη για την πρόβλεψη εγκλημάτων ιδιοκτησίας και βίας, προκειμένου να υποστηρίξει τις επιχειρήσεις της αστυνομίας για την καταστολή του εγκλήματος. Το αναπτυχθέν μοντέλο περιλαμβάνει ένα μοντέλο γραμμικής παλινδρόμησης και ένα μη γραμμικό νευρωνικό δίκτυο εμπρόσθιας διάδοσης, ενώ τα αποτελέσματα αυτής της έρευνας έδειξαν ότι το μοντέλο παλινδρόμησης είναι το καλύτερο για την πρόβλεψη μεγάλων πτώσεων της εγκληματικότητας και το τεχνητό νευρωνικό δίκτυο είναι καλύτερο στο να προβλέψει μεγάλες αυξήσεις της εγκληματικότητας.

Οι Corcoran et al. (2003), χρησιμοποίησαν έναν αλγόριθμο εξέτασης της επίπτωσης των περιστατικών εγκλημάτων στον εντοπισμό χωρικών συστάδων με υψηλά επίπεδα εγκληματικότητας. Οι συστάδες που εντοπίστηκαν χρησιμοποιήθηκαν για την εκπαίδευση τεχνητών νευρωνικών δικτύων βασισμένων στην αυτοπαλινδρόμηση για την πρόβλεψη περιστατικών σε κάθε συστάδα με βάση τα αναπτυχθέντα προβλεπτικά μοντέλα.

Σύμφωνα με την μέχρι τώρα έρευνα, τα τεχνητά νευρωνικά δίκτυα χρησιμοποιήθηκαν ως εργαλείο για την πρόβλεψη του αριθμού των περιστατικών εγκληματικότητας εντός των θερμών σημείων που βρέθηκαν σύμφωνα με παρελθοντικά χωρικά δεδομένα εγκληματικότητας με την εφαρμογή χωρικών μεθόδων για τον εντοπισμό των θερμών σημείων (hotspots) (Cohen et al., 2007, Corcoran et al., 2003, Olligschlaeger, 1997). Η διαφοροποιημένη προσέγγιση αυτής της έρευνας σε σύγκριση με άλλες μελέτες είναι ότι θεωρεί τα θερμά σημεία ως δυναμικά και κινητά χαρακτηριστικά και όχι στατικά. Η νέα προσέγγιση αυτής της έρευνας είναι ότι μελετά,

τη χωρική πρόβλεψη των θερμών σημείων της εγκληματικότητας που θεωρούνται ως δυναμικές οντότητες και όχι την πρόβλεψη περιστατικών εγκληματικότητας εντός των παραγόμενων θερμών σημείων σε αντίθεση με τις άλλες έρευνες (Cohen et al., 2007, Corcoran et al., 2003, Olligschlaeger, 1997).

Η έρευνα αυτή συνδυάζει την τεχνητή νοημοσύνη, τη χωρική ανάλυση και τις τεχνικές χαρτογράφησης των θερμών σημείων για την κατασκευή του βέλτιστου μοντέλου πρόβλεψης. Στα επόμενα τμήματα θα περιγραφεί η μεθοδολογία της έρευνας, τα αποτελέσματα και η συζήτηση.

5.3 Θεωρητικό πλαίσιο

5.3.1 Ανάλυση με τα γεωγραφικά συστήματα πληροφοριών

Τα συστήματα γεωγραφικών πληροφοριών εφαρμόστηκαν από διάφορους ερευνητές για να υποστηρίξουν μεθόδους ανάλυσης της εγκληματικότητας (Andresen, 2006; Hirschfield et al., 1995; Xue and Brown, 2003). Οι σημαντικότερες τεχνικές που βασίζονται στο GIS, οι οποίες χρησιμοποιήθηκαν από τους ερευνητές για την ανάλυση της εγκληματικότητας, σύμφωνα με τη βιβλιογραφία, περιλαμβάνουν: την πολυκριτηριακή ανάλυση, με τον ορισμό αρκετών κριτηρίων για την εφαρμογή χωρικών αναλύσεων για την επίλυση διαφόρων προβλημάτων λήψης αποφάσεων (Pattyn & Wouters, 2008), την ανάλυση εγγύτερων γειτόνων (Nearest Neighbor Analysis) για την αναγνώριση χωρικών προτύπων, τον δείκτη μέτρησης της χωρικής αυτοσυσχέτισης Moran's I (Bowers, 2003), την ακτινική ανάλυση για την ανάλυση ενός χωρικού προβλήματος μέσα σε μια συγκεκριμένη απόσταση γύρω από ένα συγκεκριμένο τόπο (Hirschfield, 2008), την θεματική χαρτογράφηση, για τη μελέτη των χωρικών κατανομών των περιστατικών εγκληματικότητας και την εκτίμηση πυκνότητας πυρήνα (KDE) για τη μελέτη της χωρικής πυκνότητας των χαρακτηριστικών της εγκληματικότητας (Chainey et al., 2008).

5.3.2 Ανάλυση θερμών σημείων

Η Ανάλυση Hotspot είναι μια μέθοδος που χρησιμοποιεί τεχνικές χωρικής ανάλυσης για τον εντοπισμό περιοχών με υψηλή συγκέντρωση χαρακτηριστικών

συμβάντων. Οι πιο συνηθισμένες μέθοδοι χωρικής ανάλυσης για την εφαρμογή της ανάλυσης θερμών σημείων είναι: η εκτίμηση της πυκνότητας του πυρήνα (Kernel Density Estimation), η στατιστική Getis-Ord G_i^* και η θεματική χαρτογράφηση των γεωγραφικών περιοχών (Chainey et al., 2008). Η ανάλυση θερμών σημείων χρησιμοποιήθηκε σε αυτή την έρευνα, προκειμένου να εντοπιστούν χωρικά τα θερμά σημεία της εγκληματικότητας όπου τα περιστατικά εγκληματικότητας έχουν μεγάλη πυκνότητα. Για να υπολογιστούν τα θερμά σημεία εγκλήματος, η στατιστική Getis-Ord G_i^* χρησιμοποιήθηκε. Ο στατιστικός δείκτης Getis-Ord G_i^* χρησιμοποιεί παραμέτρους από τα χαρακτηριστικά εισόδου για τον εντοπισμό χωρικών συστάδων με υψηλές τιμές χαρακτηριστικών (Ord and Getis, 1995; Getis and Ord, 1992). Η παραγόμενη στατιστική σημασία στα χαρακτηριστικά εξόδου προσαρμόστηκε χρησιμοποιώντας τη μέθοδο διόρθωσης False Discovery Rate (FDR), η οποία μειώνει τα κρίσιμα όρια της τιμής p έτσι ώστε να ληφθούν υπόψη οι πολλαπλές δοκιμές και οι χωρικές εξαρτήσεις (Benjamini and Hochberg, 1995). Οι τιμές της βαθμολογίας z και της τιμής p για κάθε χαρακτηριστικό προσδιορίζουν τη στατιστική σημαντικότητα (Scott and Janikas, 2010).

5.3.3 Χωρική κατανομή

Η μέθοδος απόστασης εγγύτερου σημείου είναι μια μέθοδος χωρικής ανάλυσης που βασίζεται στον υπολογισμό των αποστάσεων μεταξύ κάθε σημείου και του πλησιέστερου γειτονικού σημείου και στη συνέχεια τις συγκρίνει με την αναμενόμενη μέση απόσταση ενός τυχαίου δείγματος χαρακτηριστικών ενός χωρικού προτύπου πλήρους τυχαιότητας. Οι Clark and Evans (1954), πρότειναν τον δείκτη R ως μέτρο της ομαδοποίησης των προτύπων χωρικών χαρακτηριστικών. Εκφράζει την αναλογία της παρατηρούμενης μέσης εγγύτερης γειτονικής απόστασης προς την αναμενόμενη για τυχαίο πρότυπο της ίδιας έντασης. Εάν ο δείκτης R είναι μικρότερος από ένα, το πρότυπο υποδηλώνει ομαδοποίηση. Αν ο δείκτης R είναι μεγαλύτερος από ένα, το πρότυπο εμφανίζει διασπορά. Ο δείκτης R υπολογίζεται με την εξίσωση:

$$R = \frac{\overline{d_o}}{\overline{d_e}} \quad (5.1)$$

όπου $\overline{d_o}$ είναι η παρατηρούμενη μέση απόσταση μεταξύ κάθε χαρακτηριστικού (σημείου) και του πλησιέστερου γείτονα:

$$\overline{d_o} = \sum_{i=1}^N d_i / N \quad (5.2)$$

d_i είναι η απόσταση μεταξύ κάθε χαρακτηριστικού (σημείου) i και του πλησιέστερου γείτονα.

$\overline{d_e}$ είναι η αναμενόμενη μέση απόσταση για τα χαρακτηριστικά ενός τυχαίου προτύπου (κατανομή Poisson):

$$\overline{d_e} = 0.5 / \sqrt{\frac{N}{A}} \quad (5.3)$$

όπου N είναι ο συνολικός αριθμός ή τα χαρακτηριστικά και A η συνολική επιφάνεια της περιοχής μελέτης.

Η βαθμολογία Z_R υπολογίζεται από τον τύπο:

$$Z_R = \frac{\overline{d_o} - \overline{d_e}}{SE} \quad (5.4)$$

όπου η παράμετρος SE δίνεται από τον τύπο:

$$SE = \frac{0.26136}{\sqrt{\frac{N}{A}}} \quad (5.5)$$

5.3.4 Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης

Ο Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης (Feedforward Multilayer Perceptron (FFMLP)) χρησιμοποιήθηκε σε αυτή την έρευνα, καθώς θεωρείται μία από τις καλύτερες επιλογές για την πρόβλεψη χρονοσειρών με την χρήση νευρωνικών δικτύων (Tang & Fishwick, 1993). Η δομή ενός τυπικού νευρωνικού δικτύου αποτελείται από το επίπεδο εισόδου, το κρυφό επίπεδο ή επίπεδα με συγκεκριμένο αριθμό νευρώνων και το επίπεδο εξόδου (Svozil, Kvasnička, & Pospíchal, 1997). Σε ένα πολυεπίπεδο νευρωνικό δίκτυο πολλαπλών στρωμάτων εμπρόσθιας διάδοσης, οι νευρώνες συνδέονται μόνο προς τα εμπρός. Τα επίπεδα του νευρωνικού δικτύου περιέχουν συνδέσεις με τα επόμενα επίπεδα, αλλά δεν έχουν συνδέσεις προς τα πίσω. Πιο αναλυτικά στοιχεία για τον πολυεπίπεδο αισθητήρα εμπρόσθιας διάδοσης αναφέρθηκαν σε προηγούμενα κεφάλαια.

5.3.5 Ο αλγόριθμος Levenberg-Marquardt

Ο αλγόριθμος Levenberg-Marquardt θεωρείται ως ένας από τους ταχύτερους αλγόριθμους εκπαίδευσης νευρωνικών δικτύων (Marquardt, 1963) και χρησιμοποιείται συνήθως για την επίλυση μη γραμμικών προβλημάτων (Lourakis, 2005, Ngia & Sjoberg, 2000).

Επίσης, ο αλγόριθμος Levenberg – Marquardt παρήγαγε καλύτερα αποτελέσματα από άλλους αλγορίθμους μάθησης που δοκιμάστηκαν και συγκεκριμένα ο αλγόριθμος Ελαστικής Οπισθοδιάδοσης (Resilient Backpropagation), ο αλγόριθμος BFGS Quasi-Newton και ο αλγόριθμος Κλιμακούμενων Συζυγών Κλίσεων (Scale Conjugate Gradient) χρησιμοποιώντας διαφορετικές τοπολογίες σχετικά με τον αριθμό των νευρώνων στα κρυφά επίπεδα, τον αριθμό των κρυφών επιπέδων και τις πιο κοινές συναρτήσεις μεταφοράς (transfer functions).

Ο αλγόριθμος Levenberg-Marquardt αποτελεί μια επαναληπτική μέθοδο βελτιστοποίησης που εντοπίζει το τοπικό ελάχιστο μιας συνάρτησης με πολλαπλές μεταβλητές που διαμορφώνεται ως το άθροισμα των τετραγώνων μη γραμμικών συναρτήσεων. Η βελτιστοποίηση Levenberg - Marquardt είναι ένας συνδυασμός δύο

τεχνικών ελαχιστοποίησης: η μέθοδος απότομης καθόδου (Steepest Descent Method) και η μέθοδος Gauss-Newton. Στον αλγόριθμο βαθμωτής κλίσης, το άθροισμα των τετραγωνικών σφαλμάτων μειώνεται με την ενημέρωση των παραμέτρων στην κατεύθυνση απότομης κατάβασης. Στον αλγόριθμο Gauss-Newton, το άθροισμα των τετραγωνικών σφαλμάτων μειώνεται με την εύρεση του ελαχίστου ορίου της τετραγωνικής συνάρτησης ελαχίστων τετραγώνων (Lourakis & Argyros, 2005). Ο κανόνας ενημέρωσης των βαρών του αλγόριθμου Levenberg Marquardt υπολογίζεται ως ο συνδυασμός της μεθόδου της απότομης καθόδου και του αλγορίθμου Gauss-Newton με τον ακόλουθο τύπο (Wilamowski & Irwin, 2011, Hagan & Menhaj, 1994):

$$\Delta w = (J^T J + \mu I)^{-1} J^T e \quad (5.6)$$

όπου w είναι το διάνυσμα βάρους, μ ο συντελεστής συνδυασμού και ο I ο μοναδιαίος πίνακας. Το J είναι ο Ιακωβιανός πίνακας (Jacobian matrix) και e είναι το διάνυσμα σφάλματος. Τα J και e δίδονται από τους ακόλουθους πίνακες:

$$J = \begin{bmatrix} \frac{\partial e_{1,1}}{\partial w_1} & \frac{\partial e_{1,1}}{\partial w_2} & \dots & \frac{\partial e_{1,1}}{\partial w_N} \\ \frac{\partial e_{1,2}}{\partial w_1} & \frac{\partial e_{1,2}}{\partial w_2} & \dots & \frac{\partial e_{1,2}}{\partial w_N} \\ \dots & \dots & \dots & \dots \\ \frac{\partial e_{1,M}}{\partial w_1} & \frac{\partial e_{1,M}}{\partial w_2} & \dots & \frac{\partial e_{1,M}}{\partial w_N} \\ \dots & \dots & \dots & \dots \\ \frac{\partial e_{P,1}}{\partial w_1} & \frac{\partial e_{P,1}}{\partial w_2} & \dots & \frac{\partial e_{P,1}}{\partial w_N} \\ \frac{\partial e_{P,2}}{\partial w_1} & \frac{\partial e_{P,2}}{\partial w_2} & \dots & \frac{\partial e_{P,2}}{\partial w_N} \\ \dots & \dots & \dots & \dots \\ \frac{\partial e_{P,M}}{\partial w_1} & \frac{\partial e_{P,M}}{\partial w_2} & \dots & \frac{\partial e_{P,M}}{\partial w_N} \end{bmatrix} \quad (5.7)$$

$$e = \begin{bmatrix} e_{1,1} \\ e_{1,2} \\ \dots \\ e_{1,M} \\ \dots \\ e_{P,1} \\ e_{P,2} \\ \dots \\ e_{P,M} \end{bmatrix} \quad (5.8)$$

όπου P είναι το μέγεθος του προτύπου εκπαίδευσης, το M το μέγεθος της εξόδου και το N ο αριθμός των βαρών.

Το διάνυσμα σφάλματος $e_{p,m}$ υπολογίζεται από τον τύπο:

$$e_{p,m} = d_{p,m} - o_{p,m} \quad (5.9)$$

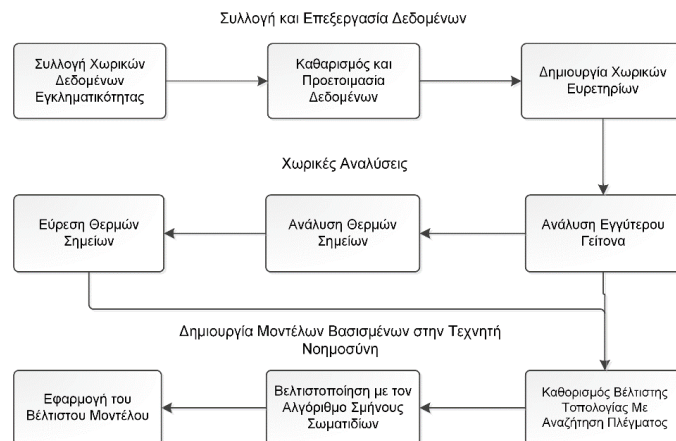
όπου το $d_{p,m}$ είναι η επιθυμητή έξοδος και $o_{p,m}$ είναι η πραγματική έξοδος σε ένα πρότυπο εκπαίδευσης p με έξοδο δικτύου m (Hagan & Menhaj, 1994).

5.4 Μεθοδολογία

Η μεθοδολογία έρευνας αποτελείται από τρία βασικά στάδια. Στο πρώτο στάδιο, τα στοιχεία της εγκληματικότητας συλλέχθηκαν και ελέγχθηκαν για τυχόν κενά και ασυνέπειες. Από τα στοιχεία της εγκληματικότητας εξήχθησαν αυτά που αφορούσαν εγκλήματα ιδιοκτησίας. Τα παραγόμενα δεδομένα χρησιμοποιήθηκαν για τη δημιουργία χωρικών ευρετηρίων χρησιμοποιώντας το λογισμικό ArcGIS.

Στο δεύτερο στάδιο, πραγματοποιήθηκε η ανάλυση απόστασης εγγύτερων γειτόνων και υπολογίστηκε ο δείκτης R που αποτελεί μέτρο της χωρικής ομαδοποίησης των περιστατικών εγκληματικότητας. Στη συνέχεια, εφαρμόστηκε η ανάλυση θερμών σημείων για τον εντοπισμό των στατιστικά σημαντικών περιοχών αυξημένης έντασης περιστατικών εγκληματικότητας (Getis and Ord, 1992). Στο τρίτο στάδιο, τα δεδομένα που παρήχθησαν στο προηγούμενο στάδιο, χρησιμοποιήθηκαν ως παράμετροι εισόδου για τη δημιουργία προβλεπτικών μοντέλων Τεχνητών Νευρωνικών Δικτύων με

πολλαπλή έξοδο, αφού θα προβλέπουν τις συντεταγμένες x και y . Τα Τεχνητά Νευρωνικά Δίκτυα εφαρμόζονται στην περίπτωση αυτή, αφού που μπορούν να μοντελοποιήσουν και μη γραμμικές σχέσεις πολλαπλής εισόδου με πολλαπλή έξοδο (Multiple Input Multiple Output – MIMO) (Svozil et al., 1997). Οι Μηχανές Διανυσμάτων Υποστήριξης (ΜΔΥ) δεν χρησιμοποιήθηκαν, αφού οι ΜΔΥ πολλαπλής εισόδου – πολλαπλής εξόδου (MIMO) αποτελούν ακόμα αντικείμενο έρευνας (Xu et al., 2013). Για την ανάπτυξη των μοντέλων τεχνητής νοημοσύνης χρησιμοποιήθηκε το λογισμικό MATLAB. Τα δεδομένα κανονικοποιήθηκαν χρησιμοποιώντας την κανονικοποίηση ελαχίστου – μεγίστου στο διάστημα $[0,1]$, που θεωρείται ότι εξάγει ακριβέστερα αποτελέσματα από άλλες μεθόδους κανονικοποίησης (Al Shalabi et al., 2006), εφαρμόζοντας τον μαθηματικό τύπο (4.1) και καθορίστηκε το σύνολο εκπαίδευσης, το σύνολο επικύρωσης και το σύνολο δοκιμής. Έπειτα, χρησιμοποιήθηκε ο αλγόριθμος αναζήτησης πλέγματος για την σύγκριση τοπολογιών (αριθμός νευρώνων κρυφών επιπέδων, τύπος αλγορίθμου εκπαίδευσης, κλπ), για την εύρεση της καλύτερης τοπολογίας του μοντέλου πρόβλεψης. Στη συνέχεια, τα βάρη του τεχνητού νευρωνικού δικτύου βελτιστοποιήθηκαν με τον αλγόριθμο Βελτιστοποίησης Σμήνους Σωματιδίων. Όλα τα στάδια της μεθοδολογίας της έρευνας αναλύονται στις επόμενες ενότητες. Η παρακάτω εικόνα, απεικονίζει μια επισκόπηση της μεθοδολογίας που ακολουθήθηκε.



Εικόνα 5.1 Επισκόπηση των σταδίων της μεθοδολογίας που ακολουθήθηκε για την χωρική πρόβλεψη της εγκληματικότητας.

5.5 Αποτελέσματα

5.5.1 Συλλογή Δεδομένων – Περιοχή Μελέτης

Τα δεδομένα εγκληματικότητας ανακτήθηκαν από την επίσημη ιστοσελίδα της πόλης του Σικάγου³, όπου αναρτήθηκαν από το Αστυνομικό Τμήμα του Σικάγου. Επιλέχθηκε η πόλη του Σικάγου, δεδομένου ότι η εγκληματικότητα εκεί είναι αρκετά αυξημένη και σε ημερήσια βάση ο αριθμός των χωρικών περιστατικών είναι ικανός για να πραγματοποιηθεί επεξεργασία και εντοπισμός των θερμών σημείων με το λογισμικό ArcGIS, σε αντίθεση με την Ελλάδα και με άλλες περιοχές του κόσμου, που δεν υπάρχει διαθέσιμος ικανός αριθμός χωρικών δεδομένων εγκληματικότητας σε ημερήσια βάση για χρονικό διάστημα επαρκές, ώστε να μπορούν να εκπαιδευτούν τα νευρωνικά μοντέλα.

Τα δεδομένα περιλαμβάνουν πληροφορίες σχετικά με τον τύπο εγκλήματος και τις συντεταγμένες του εγκλήματος. Οι επιλεγμένοι τύποι εγκλημάτων για την έρευνα αυτή αφορούσαν εγκλήματα ιδιοκτησίας: διάρρηξη, κλοπή αυτοκινήτων, ληστεία, κλοπή. Συγκεντρώθηκαν συνολικά 44.280 καταχωρήσεις περιστατικών για την περίοδο: 01/01/2014 έως 30/06/2014 σε ημερήσια βάση. Τα συλλεχθέντα στοιχεία της εγκληματικότητας καθαρίστηκαν και προεπεξεργάστηκαν για τυχόν ασυνέπειες και διπλότυπα. Τα δεδομένα εγκλημάτων, εισήχθησαν στο λογισμικό ArcGIS για τη δημιουργία χωρικών ευρετηρίων. Ο παρακάτω πίνακας, δείχνει τον αριθμό των χωρικών δεδομένων εγκληματικότητας που συλλέχθηκαν ανά μήνα και τον συνολικό αριθμό των δεδομένων που αναλύθηκαν σε αυτή την έρευνα.

Πίνακας 5.1: Δεδομένα εγκληματικότητας που χρησιμοποιήθηκαν στην έρευνα αυτή.

<i>Μήνας (2014)</i>	<i>Εγκλήματα Ιδιοκτησίας (Διαρρήξεις, κλοπές, ληστείες, κλοπές αυτοκινήτων)</i>	<i>Μέσος Ημερήσιος Αριθμός Εγκλημάτων</i>
Ιανουάριος	7142	230
Φεβρουάριος	6129	219
Μάρτιος	7066	228
Απρίλιος	7278	243
Μάιος	7995	258
Ιούνιος	8670	289
Σύνολο	44280	

³ <https://www.cityofchicago.org>

5.5.2 Χωρικές μέθοδοι ανάλυσης

Η ανάλυση απόστασης εγγύτερων γειτόνων χρησιμοποιήθηκε για τον υπολογισμό του δείκτη συγκέντρωσης R ως μέτρο για τον τύπο του χωρικού πρότυπου των περιστατικών εγκληματικότητας σε ημερήσια βάση. Η ανάλυση απόστασης εγγύτερων γειτόνων πραγματοποιήθηκε για τις ημέρες που αντιστοιχούσαν στα περιστατικά εγκληματικότητας για αυτή την έρευνα. Η πλειοψηφία των δεικτών R των ημερήσιων δεδομένων εγκλημάτων που υπολογίστηκαν, υποδεικνύει το ομαδοποιημένο χωρικό πρότυπο. Ο δείκτης R χρησιμοποιήθηκε ως μία από τις παραμέτρους εισόδου στα προβλεπτικά μοντέλα που αναπτύχθηκαν.

Στη συνέχεια, τα θερμά σημεία των περιστατικών εγκλήματος υπολογίστηκαν για κάθε ημέρα για την περίοδο από 1/1/2014 έως 30/6/2014 στην περιοχή μελέτης. Η μέθοδος ανάλυσης θερμών σημείων χρησιμοποιήθηκε για την εύρεση των σημείων αυξημένης έντασης της εγκληματικότητας. Ο στατιστικός δείκτης Getis-Ord G_i^* χρησιμοποιήθηκε για την επεξεργασία των δεδομένων εισόδου έτσι ώστε να υπολογιστούν τα θερμά σημεία εγκληματικότητας (Getis and Ord, 1992; Ord and Getis, 1995).

5.6 Τεχνητά νευρωνικά δίκτυα

5.6.1 Εφαρμογή

Ένας Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης χρησιμοποιήθηκε σε αυτή την έρευνα, καθώς θεωρείται ως ο πλέον κατάλληλος για στην πρόβλεψη προβλημάτων χρονοσειρών (Koskela et al., 1996; Tang and Fishwick, 1993). Ένας Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης αποτελείται συνήθως από τρία επίπεδα νευρώνων. Οι παράγοντες που πρέπει να καθοριστούν για την υλοποίηση ενός Τεχνητού Νευρωνικού Δικτύου περιλαμβάνουν τον καθορισμό των συνόλων εκπαίδευσης, επικύρωσης και δοκιμής καθώς και τον προσδιορισμό της μεθόδου εκπαίδευσης και της τοπολογίας (επίπεδο εισόδου, κρυφά επίπεδα και αριθμός νευρώνων, επίπεδο εξόδου) του δικτύου.

Το διάνυσμα εισόδου X_i περιλαμβάνει: την συντεταγμένη X του θερμού σημείου της προηγούμενης ημέρας που συμβολίζεται ως $coordinateX_{n-1}$, την συντεταγμένη Y του θερμού σημείου της προηγούμενης ημέρας που συμβολίζεται ως $coordinateY_{n-1}$, και τον δείκτη R της προηγούμενης ημέρας που συμβολίζεται ως R_{n-1} . Άρα θα έχουμε:

$$X_i = (coordinateX_{n-1}, coordinateY_{n-1}, R_{n-1}) \quad (5.10)$$

Οι τιμές εξόδου Y_i αφορούν τις συντεταγμένες των θερμών σημείων $coordinateX_n$ και $coordinateY_n$, της επόμενης ημέρας:

$$Y_i = (coordinateX_n, coordinateY_n) \quad (5.11)$$

5.6.2 Ορισμός συνόλων εκπαίδευσης, επικύρωσης και δοκιμών

Το σύνολο δεδομένων της περιοχής μελέτης χωρίστηκε σε τρία διαφορετικά μέρη. Το 70% του συνόλου δεδομένων χρησιμοποιήθηκε ως σύνολο εκπαίδευσης, το 15% του συνόλου δεδομένων χρησιμοποιήθηκε για το σύνολο επικύρωσης και το 15% για το σετ δοκιμής. Το σετ επικύρωσης είναι ένα ανεξάρτητο δείγμα που χρησιμοποιείται για την αξιολόγηση της γενίκευσης του νευρωνικού δικτύου και για τη διακοπή της διαδικασίας εκπαίδευσης, όταν η απόδοση δεν βελτιώνεται ώστε να αποφευχθεί η υπερ-εκπαίδευση (overfitting). Το σετ δοκιμής είναι επίσης ένα ανεξάρτητο δείγμα που δεν έχει καμία επίδραση στη φάση της εκπαίδευσης και είναι χρησιμοποιείται για να δοκιμαστεί το νευρωνικό δίκτυο που αναπτύχθηκε.

5.6.3 Τοπολογία βέλτιστου νευρωνικού μοντέλου πρόβλεψης

Ο αλγόριθμος Levenberg – Marquardt επιλέχθηκε ως ο βέλτιστος αλγόριθμος εκπαίδευσης για τα νευρωνικά μοντέλα δεδομένου ότι είχε τα καλύτερα αποτελέσματα και επειδή θεωρείται ένας από τους πιο γρήγορους αλγόριθμους ενημέρωσης βαρών (Lourakis, 2005, Zhao & Xu, 2004). Ο αλγόριθμος Levenberg – Marquardt χρησιμοποιήθηκε σε αυτή τη μελέτη, ο οποίος είχε καλύτερη ικανότητα γενίκευσης και

πιο γρήγορη διαδικασία εκπαίδευσης σε σύγκριση με άλλους αλγόριθμους, όπως ο αλγόριθμος Ελαστικής Οπισθοδιάδοσης (Resilient Backpropagation) (Riedmiller and Braun, 1993) και αντίστροφη μετατόπιση κλίσης (Baldi, 1995). Ερευνήθηκαν αρκετές δομές τοπολογιών σχετικά με τον αριθμό των κρυφών επιπέδων, τον αριθμό των νευρώνων κάθε κρυφού επιπέδου και τη φύση των συναρτήσεων ενεργοποίησης των κρυφών στρωμάτων. Η καλύτερη τοπολογία επιλέχθηκε σύμφωνα με τα αποτελέσματα απόδοσης (Πίνακας 5.2).

Η ρίζα του μέσου τετραγωνικού σφάλματος (RMSE) επιλέχθηκε έτσι ώστε να εκτιμηθεί το σφάλμα πρόβλεψης και να αξιολογηθεί η απόδοση κάθε νευρωνικού μοντέλου. Η μαθηματική έκφραση για τον υπολογισμό της RMSE δίνεται από την ακόλουθη εξίσωση:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_{p_i} - y_{r_i})^2} \quad (5.12)$$

Όπου y_{p_i} είναι η προβλεφθείσα τιμή, y_{r_i} είναι η πραγματική τιμή και N είναι ο αριθμός των τιμών εξόδου.

Για την βελτιστοποίηση της τοπολογίας του νευρωνικού δικτύου εφαρμόστηκε η μέθοδος αναζήτησης πλέγματος (Grid Search) (Basnayake et al., 2017). Το πλεονέκτημα της μεθόδου είναι ότι αυτή είναι πλήρης, δηλαδή ερευνά κάθε δυνατό συνδυασμό για να βρεθεί ο βέλτιστος. Επίσης, ένα άλλο πλεονέκτημα είναι ότι μπορεί να βελτιστοποιήσει υπερπαραμέτρους που αποτελούν μίγμα κατηγορικών και μη κατηγορικών υπερπαραμέτρων σε αντίθεση με άλλες βελτιστοποιήσεις αυτόματης μηχανικής μάθησης (AutoML) όπως η βελτιστοποίηση Bayes, οι οποίες αντιμετωπίζουν προβλήματα όταν υπάρχει συνδυασμός τέτοιου είδους υπερπαραμέτρων (Feurer et al., 2015; Guyon et al., 2016; Mendoza et al., 2016). Όπως για παράδειγμα, οι κατηγορικές μεταβλητές στο παρόν πρόβλημα θα είναι: τύπος συνάρτησης μεταφοράς, τύπος αλγορίθμου εκπαίδευσης που πρέπει να βελτιστοποιηθούν σε συνδυασμό με μη κατηγορικές μεταβλητές όπως: ο αριθμός νευρώνων στο κρυφό επίπεδο. Οι υπερπαραμέτροι που ερευνήθηκαν αφορούν: τον αριθμό νευρώνων σε κάθε κρυφό επίπεδο, τις πιο συνηθισμένες συναρτήσεις

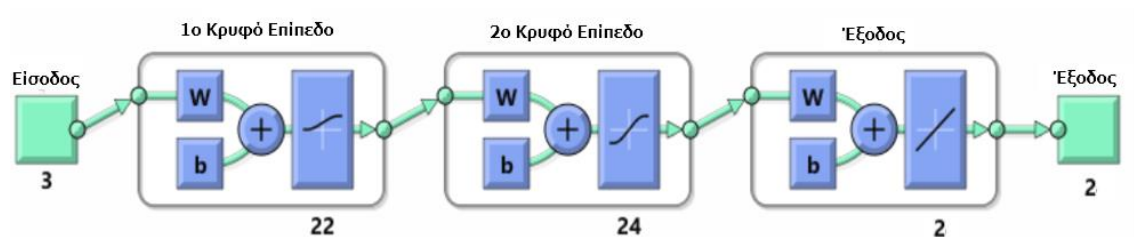
μεταφοράς για τα ένα έως δύο κρυφά επίπεδα: Η εφαπτομενική – σιγμοειδής συνάρτηση μεταφοράς (Tan-Sigmoid Transfer Function (TSTF)), η λογιστική – σιγμοειδής συνάρτηση μεταφοράς (Log-Sigmoid Transfer Function (LSTF)), η γραμμική συνάρτηση μεταφοράς (Linear Transfer Function (LTF)), η Elliot σιγμοειδής συνάρτηση μεταφοράς (Elliot Sigmoid Transfer Function (ESTF)) και η θετική γραμμική συνάρτηση μεταφοράς (Positive Linear Transfer Function (PLTF)). Επίσης, ερευνήθηκαν οι πιο συνηθισμένοι αλγόριθμοι μάθησης για την εκπαίδευση κάθε νευρωνικού μοντέλου και ήταν οι: αλγόριθμος Κλιμακούμενων Συζυγών Κλίσεων (Scaled Conjugate Gradient – SCG), BFGS Quasi-Newton (BFGS-QN), Ελαστικής Οπισθοδιάδοσης (Resilient Backpropagation – RB), Levenberg Marquardt (LM) Gradient Descent with Adaptive Learning Rate (GDALR), και Gradient Descent with Momentum and Adaptive Learning Rate Backpropagation (GDMALRB).

Ο πίνακας 5.2 παρουσιάζει τα βέλτιστα μοντέλα τεχνητών νευρωνικών δικτύων που βρέθηκαν με την μέθοδο αναζήτησης πλέγματος, μεταξύ όλων που δοκιμάστηκαν σύμφωνα με την ικανότητα γενίκευσης με μέτρηση της ρίζας του μέσου τετραγωνικού σφάλματος, για διαφορετικούς αλγορίθμους μάθησης, τοπολογίες και συναρτήσεις ενεργοποίησης κρυφών επιπέδων που χρησιμοποιήθηκαν για την κατασκευή των μοντέλων.

Πίνακας 5.2. Τα καλύτερα μοντέλα νευρωνικών δικτύων που ερευνήθηκαν σύμφωνα με τη ικανότητα γενίκευσής τους με διαφορετικούς αλγορίθμους εκπαίδευσης, τοπολογίες και συναρτήσεις ενεργοποίησης.

Μοντέλο	Αλγόριθμος Εκπαίδευσης	Αριθμός Κρυφών Επιπέδων	Αριθμός Νευρώνων στο 1 ^ο Κρυφό Επίπεδο	Αριθμός Νευρώνων στο 2 ^ο Κρυφό Επίπεδο	Συναρτήσεις Ενεργοποίησης	RMSE Σφάλμα Εκπαίδευσης	RMSE Σφάλμα Επικύρωσης
ANN1	RP	2	30	19	LAF – TSAF	65.5781	53.6101
ANN2	LM	2	22	24	LSAF – TSAF	1.2449	1.7187
ANN3	SCG	2	15	19	LAF – TSAF	448.7179	251.3825
ANN4	BFGS-QN	2	23	21	LAF – TSAF	427.2857	242.6973
ANN5	GDALRB	2	30	28	LAF – TSAF	1.1647e+03	825.7744
ANN6	GDMALRB	2	19	23	LAF – LSAF	1.2377e+03	692.8101

Το βέλτιστο μοντέλο τεχνητού νευρωνικού δικτύου με την καλύτερη απόδοση ήταν το ANN2 που έχει τα ακόλουθα χαρακτηριστικά: δύο κρυφά επίπεδα, τον αλγόριθμο Levenberg – Marquardt ως αλγόριθμο εκπαίδευσης, ενώ το πρώτο κρυφό επίπεδο αποτελείται από 22 νευρώνες και υλοποιεί ως συνάρτηση ενεργοποίησης την LSAF και το δεύτερο κρυφό επίπεδο που αποτελείται από 24 νευρώνες και υλοποιεί την TSAF ως συνάρτηση ενεργοποίησης. Τα αποτελέσματα έδειξαν πολύ καλή ακρίβεια, το σφάλμα RMSE του σετ επικύρωσης βρέθηκε 1.7187 και το σφάλμα RMSE του συνόλου εκπαίδευσης βρέθηκε 1.2449.



Εικόνα 5.2 Το σχήμα αναπαριστά την τοπολογία του βέλτιστου μοντέλου τεχνητού νευρωνικού δικτύου για την πρόβλεψη των συντεταγμένων των θερμών σημείων με την χρήση δύο κρυφών επιπέδων (hidden layers). Το πρώτο κρυφό επίπεδο αποτελείται από 22 νευρώνες και χρησιμοποιεί την λογιστική – σιγμοειδής συνάρτηση μεταφοράς, ενώ το δεύτερο επίπεδο αποτελείται από 24 νευρώνες και χρησιμοποιεί την εφαπτομενική – σιγμοειδής συνάρτηση μεταφοράς.

5.6.4 Βελτιστοποίηση Νευρωνικού Δικτύου με Σμήνος Σωματιδίων

Για την βελτιστοποίηση των βαρών του νευρωνικού δικτύου εφαρμόστηκε η μέθοδος Βελτιστοποίησης Σμήνους Σωματιδίων (Particle Swarm Optimization, PSO). Η βελτιστοποίηση Σμήνους Σωματιδίων προτάθηκε από τους Eberhart and Kennedy (1995). Επιλέχθηκε ο αλγόριθμος Βελτιστοποίησης Σμήνους Σωματιδίων, διότι σε σύγκριση με άλλες μεθόδους βελτιστοποίησης νευρωνικών δικτύων, όπως ο γενετικός αλγόριθμος πλεονεκτεί ως προς διάφορα σημεία όπως η ευκολία εφαρμογής και η γρήγορη ρύθμιση λιγότερων παραμέτρων (Ou and Lin, 2006).

Η Βελτιστοποίηση Σμήνους Σωματιδίων εξετάζει πληθυσμούς λύσεων για την εύρεση του βέλτιστου μίας συνάρτησης ή ενός προβλήματος. Ο αλγόριθμος Βελτιστοποίησης Σμήνους Σωματιδίων PSO βασίζεται σε μια στοχαστική τεχνική που εμπνεύστηκε από τη διερεύνηση της κοινωνικής συμπεριφοράς και της νοημοσύνης σμήνους στη φύση. Στον αλγόριθμο PSO, τα σωματίδια του σμήνους διασχίζουν το χώρο αναζήτησης για να αξιολογήσουν την αντικειμενική συνάρτηση (objective function). Κάθε σωματίδιο σμήνους χαρακτηρίζεται από την τρέχουσα θέση, την προηγούμενη καλύτερη θέση και την ταχύτητα. Ο αλγόριθμος βελτιστοποίησης σμήνους σωματιδίων (PSO) περιγράφηκε πιο αναλυτικά στην ενότητα 4.4.4.

Για την βελτιστοποίηση των παραμέτρων του Σμήνους Σωματιδίων, εφαρμόστηκε η μέθοδος αναζήτησης πλέγματος (Grid Search), που περιγράφηκε παραπάνω, για τις παραμέτρους που αφορούν: τον πληθυσμό των σωματιδίων του σμήνους, τους συντελεστές μάθησης c_1 , c_2 , τον συντελεστή βάρους αδράνειας (inertia weight coefficient). Οι πληθυσμοί των σωματιδίων σμήνους που εξετάστηκαν ήταν 20, 30, 40, 50, που είναι οι πιο συνηθισμένες τιμές σύμφωνα με τη βιβλιογραφία (Fourie and Groenwold, 2002). Οι συντελεστές μάθησης c_1 , c_2 , ορίστηκαν στην τιμή 2.0, όπως προτείνεται από τους Kennedy and Eberhart (1995). Οι τιμές του συντελεστή βάρους αδράνειας (inertia weight coefficient) εξετάστηκαν στο διάστημα [0.8, 1.4] με βήμα 0.1, όπως προτείνεται από τους Shi and Eberhart, (1998).

Οι παράμετροι που παρήγαγαν τα βέλτιστα αποτελέσματα ήταν: μέγεθος πληθυσμού: $P = 50$, βάρος αδράνειας $w = 1.10$, συντελεστής προσωπικής μάθησης $c_1 = 2.0$, ολικός συντελεστής εκμάθησης $c_2 = 2.0$. Ο παρακάτω πίνακας, παρουσιάζει τα πέντε καλύτερα μοντέλα του αλγορίθμου βελτιστοποίησης PSO, που βρέθηκαν χρησιμοποιώντας διαφορετικές τοπολογίες παραμέτρων. Το βέλτιστο μοντέλο της βελτιστοποίησης Σμήνους Σωματιδίων ήταν το μοντέλο 5. Η Ρίζα του Μέσου Τετραγωνικού Σφάλματος (RMSE) του βέλτιστου μοντέλου που δοκιμάστηκε στο σύνολο δοκιμής βρέθηκε να είναι 0.1023.

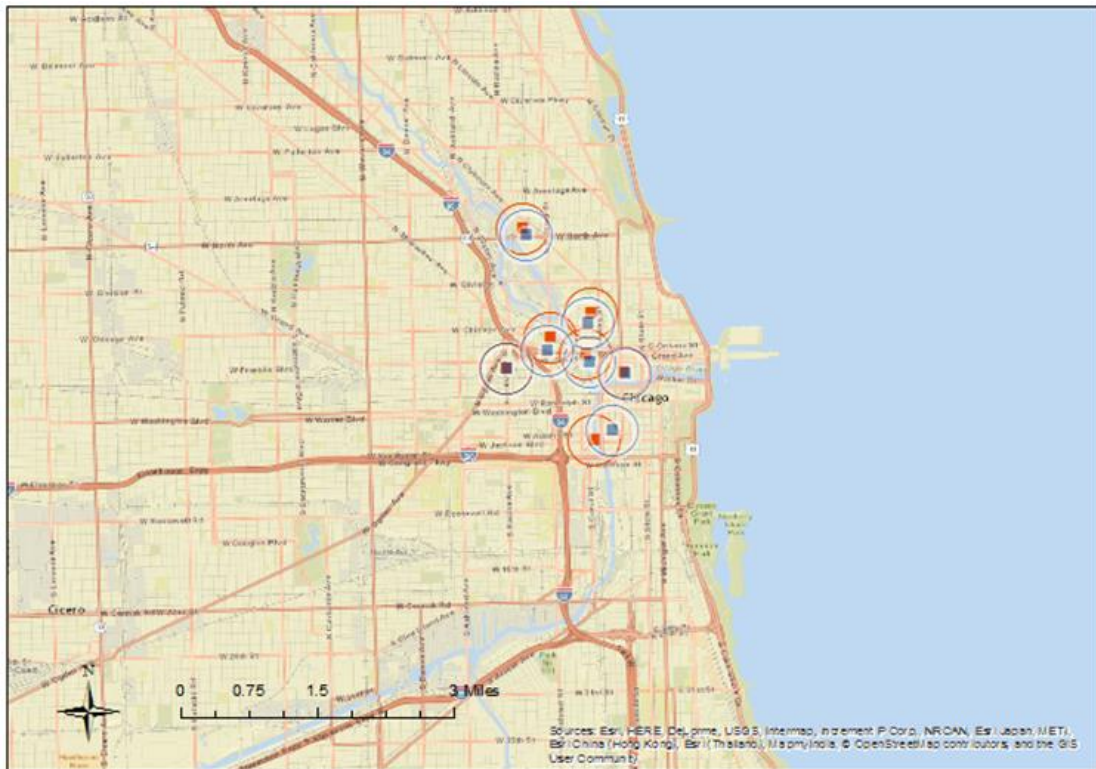
Πίνακας 5.3. Τα πέντε καλύτερα νευρωνικά μοντέλα του αλγορίθμου βελτιστοποίησης PSO που βρέθηκαν χρησιμοποιώντας διαφορετικές τοπολογίες παραμέτρων για την βελτιστοποίηση των βαρών του νευρωνικού δικτύου.

Μοντέλο	Πληθυσμός Σμήνους P	c_1	c_2	w	Βέλτιστο Κόστος (Best Cost)
<i>Μοντέλο 1</i>	20	2.0	2.0	0.90	1.1652
<i>Μοντέλο 2</i>	40	2.0	2.0	1.10	0.9582
<i>Μοντέλο 3</i>	40	2.0	2.0	0.90	0.8795
<i>Μοντέλο 4</i>	40	2.0	2.0	1.20	1.0152
<i>Μοντέλο 5</i>	50	2.0	2.0	1.10	0.1023

5.6.5 Προβλέψεις του Προτεινόμενου Μοντέλου

Το προτεινόμενο βέλτιστο μοντέλο νευρωνικού δικτύου χρησιμοποιήθηκε για την πρόβλεψη των θερμών σημείων εγκλημάτων σε ημερήσια βάση στην περιοχή της πόλης του Σικάγου.

Η εικόνα 5.3, αναπαριστά ένα χάρτη σύγκρισης μεταξύ των κέντρων, των πραγματικών και των προβλεφθέντων θερμών σημείων από το βέλτιστο μοντέλο πρόβλεψης. Τα τελικά αποτελέσματα έδειξαν πολύ καλή ακρίβεια προβλέψεων. Στην επόμενη εικόνα, αναπαρίστανται τα κέντρα των θερμών σημείων που προβλέφθηκαν για την τελευταία εβδομάδα του συνόλου δοκιμής.



Εικόνα 5.3 Απεικόνιση χάρτη σύγκρισης μεταξύ των κέντρων, των πραγματικών και των προβλεφθέντων θερμών σημείων από το βέλτιστο μοντέλο πρόβλεψης. Τα πραγματικά κέντρα hotspot απεικονίζονται με κόκκινο χρώμα και αυτά που προβλέφθηκαν με μπλε χρώμα.

5.6.6 Συμπεράσματα – Συζήτηση

Οι εφαρμογές των γεωγραφικών συστημάτων πληροφοριών έχουν αναβαθμιστεί τα τελευταία χρόνια με την υιοθέτηση πιο εξελιγμένων τεχνολογιών όπως η τεχνητή νοημοσύνη (Ferentinou and Chalkias, 2013; Openshaw and Openshaw, 1997; Voženílek 2009). Τα θερμά σημεία εγκληματικότητας αποτελούν σημαντικό μέτρο της συγκέντρωσης της εγκληματικότητας στο χώρο και θεωρούνται ως υψηλής σημασίας πληροφορία για τις αρμόδιες αρχές. Τα θερμά σημεία της εγκληματικότητας ενδέχεται να διαφέρουν ανάλογα με την κατηγορία του εγκλήματος, την περιοχή και τον χρόνο που συμβαίνει. Για το λόγο αυτό, στην παρούσα έρευνα για την χωρική πρόβλεψη της

εγκληματικότητας επιλέχθηκε μια κατηγορία εγκλημάτων τα εγκλήματα ιδιοκτησίας (διάρρηξη, κλοπή αυτοκινήτων, ληστεία, κλοπή) για να ερευνηθεί σε συγκεκριμένη χρονική περίοδο και περιοχή. Προκειμένου να βελτιωθεί η ακρίβεια των αποτελεσμάτων, η χρονική περίοδος επιλέχθηκε έτσι ώστε να μην είναι τόσο μεγάλη για το λόγο ότι οι κοινωνικοοικονομικοί παράγοντες που μπορεί να επηρεάσουν την εγκληματικότητα θα είναι περίπου στα ίδια επίπεδα και ταυτόχρονα τα συλλεχθέντα δεδομένα για αυτή την περίοδο θα είναι αρκετά για να τροφοδοτήσουν και να εκπαιδεύσουν επαρκώς τα προβλεπτικά μοντέλα βασισμένα στην τεχνητή νοημοσύνη.

Η έρευνα αυτή διαφέρει από τις άλλες μελέτες που έχουν διεξαχθεί μέχρι σήμερα, σύμφωνα με τη βιβλιογραφία (Cohen et al., 2007, Corcoran et al., 2003, Olligschlaeger, 1997), δεδομένου ότι οι ερευνητές επικεντρώθηκαν στην πρόβλεψη του αριθμού των περιστατικών εγκλημάτων στα θερμά σημεία στις μεθοδολογίες που εφάρμοσαν, σε αντίθεση με την παρούσα έρευνα που επικεντρώνεται στην πρόβλεψη των χωρικών διαστάσεων των θερμών σημείων της εγκληματικότητας. Στην παρούσα έρευνα, τα θερμά σημεία (hotspots) της εγκληματικότητας θεωρούνται ως δυναμικές οντότητες και όχι ως στατικές σε αντίθεση με τις άλλες έρευνες που επισκοπήθηκαν (Corcoran et al., 2003, Olligschlaeger, 1997).

Επίσης, στην έρευνα αυτή, σύμφωνα με την προτεινόμενη μεθοδολογία, επιχειρείται η πρόβλεψη του χωρικού χαρακτήρα της εγκληματικότητας, λαμβάνοντας υπόψη τη χωρική κατανομή των περιστατικών εγκλημάτων, εφαρμόζοντας μεθόδους χωρικής ανάλυσης και τεχνητής νοημοσύνης.

Σύμφωνα με τη βιβλιογραφία διάφοροι ερευνητές χρησιμοποίησαν εργαλεία ανάλυσης βασισμένα στο GIS για την πρόβλεψη εγκληματικότητας που εντοπίζεται στο χώρο, όμως πολύ λίγες έρευνες συνδύασαν μεθόδους βασισμένες στα γεωγραφικά συστήματα πληροφοριών με μεθόδους τεχνητής νοημοσύνης. Ο Πίνακας 3, αναπαριστά τις έρευνες αυτές, τις μεθοδολογίες που χρησιμοποιήθηκαν και τις μετρικές ακρίβειας.

Πίνακας 5.4. Πίνακας σύγκρισης της υπάρχουσας βιβλιογραφίας σχετικά με την χρήση μεθόδων χωρικής ανάλυσης και νευρωνικών δικτύων για την πρόβλεψη της εγκληματικότητας χωρικά.

<i>Βιβλιογραφική Αναφορά</i>	<i>Μεθοδολογία</i>
Olligschlaeger (1997)	Χωρική Ανάλυση, Τεχνητά Νευρωνικά Δίκτυα, Παλινδρόμηση
Cohen et al. (2007)	Χωρική Ανάλυση, Τεχνητά Νευρωνικά Δίκτυα, Παλινδρόμηση
Corcoran et al. (2003)	Χωρική Ανάλυση, Τεχνητά Νευρωνικά Δίκτυα, Παλινδρόμηση, Γάμμα Τεστ, Τυχαίος Βηματισμός (Random Walk)

Τα αποτελέσματα αυτής της έρευνας έδειξαν μεγαλύτερη ακρίβεια πρόβλεψης σε σύγκριση με άλλες έρευνες που χρησιμοποίησαν νευρωνικά δίκτυα και μεθόδους χωρικής ανάλυσης (Cohen et al., 2007, Corcoran et al., 2003, Olligschlaeger, 1997).

Επίσης, η παρούσα έρευνα διαφέρει από τις άλλες σχετικές μελέτες, δεδομένου ότι βασίζεται στην προβλεπτική διοίκηση. Σύμφωνα με την προτεινόμενη μεθοδολογία, εφαρμόστηκαν βελτιστοποιημένες τεχνικές τεχνητής νοημοσύνης για την πρόβλεψη παραγόντων σημαντικών στην δημόσια διοίκηση, όπως στο κεφάλαιο αυτό της των θερμών σημείων εγκληματικότητας.

Η πρόβλεψη αφορά το επόμενο χρονικό βήμα που στην παρούσα μελέτη στο κεφάλαιο αυτό είναι η επόμενη ημέρα. Δηλαδή, με δεδομένα παραγόντων εισόδου που επηρεάζουν την τιμή εξόδου (εξαρτημένη μεταβλητή) του ‘σήμερα’ μπορεί να προβλεφθεί η τιμή του ‘αύριο’. Το στοιχείο αυτό, μπορεί να βοηθήσει αρκετά τις δημόσιες αρχές, αφού θα έχουν το χρόνο μέχρι να επέλθει το επόμενο χρονικό βήμα, για να λάβουν μέτρα προς στην κατεύθυνση της μείωσης της εγκληματικότητας στις περιοχές που προβλέφθηκαν ότι θα έχουν αυξημένη ένταση περιστατικών εγκλημάτων.

Η προτεινόμενη μεθοδολογία μπορεί να βοηθήσει τις αρμόδιες αρχές στη διαδικασία λήψης αποφάσεων σχετικά με την πρόληψη της εγκληματικότητας καθώς επίσης στις επιχειρήσεις επέμβασης της αστυνομίας. Επίσης, η προτεινόμενη έρευνα μπορεί να βοηθήσει στη λήψη αποφάσεων πολλές δημόσιες υπηρεσίες εκτός από τις δικαστικές και αστυνομικές αρχές. Για παράδειγμα, η χωρική πρόβλεψη της εγκληματικότητας μπορεί να βοηθήσει στον εντοπισμό των μη ασφαλών περιοχών στο συγκοινωνιακό δίκτυο όπου μετακινούνται μεγάλοι πληθυσμοί (Kouziokas, 2016, 2017c), στην χωροθέτηση εκπαιδευτικών μονάδων, κλπ. παιδικών σταθμών, ώστε να μην είναι κοντά στις περιοχές αυξημένης εγκληματικότητας, αλλά και σε άλλες περιπτώσεις στη λήψη αποφάσεων σε αστικές περιοχές, όπως για παράδειγμα στον εντοπισμό μη ασφαλών περιοχών σε τουριστικές περιοχές που έχουν σχέση με το μάρκετινγκ των πόλεων και την τουριστική τους ανάπτυξη.

Η προτεινόμενη μεθοδολογία και οι χρησιμοποιούμενες τεχνικές μπορούν να εφαρμοστούν από τις αρχές στη δημόσια και δικαστική διαδικασία λήψης αποφάσεων, κατά τον σχεδιασμό στρατηγικών πρόληψης της εγκληματικότητας. Επίσης, η προτεινόμενη μεθοδολογία μπορεί να διευκολύνει την υιοθέτηση των κατάλληλων πολιτικών για την πρόληψη της εξάπλωσης της εγκληματικότητας και τη συμβολή στη χάραξη στρατηγικών κατά της εγκληματικότητας, προκειμένου να διαφυλαχθούν οι ζωές των πολιτών πιο αποτελεσματικά.

Οι περιορισμοί της προτεινόμενης μεθοδολογίας αφορά τα χωρικά δεδομένα της εγκληματικότητας που πρέπει να είναι διαθέσιμα, αλλά και τον αριθμό των περιστατικών εγκλημάτων δεδομένου ότι η ανάλυση θερμών σημείων για να εφαρμοστεί, απαιτεί έναν ελάχιστο αριθμό συμβάντων.

Οι μελλοντικές επεκτάσεις της έρευνας αυτής αφορούν την ανάπτυξη και τον έλεγχο παρόμοιων μοντέλων πρόβλεψης για τη διερεύνηση του βαθμού χωρικής προβλεψιμότητας διάφορων τύπων εγκλημάτων σε άλλες πόλεις και περιοχές του κόσμου.

ΚΕΦΑΛΑΙΟ 6: Τεχνητή Νοημοσύνη στην Πρόβλεψη Περιβαλλοντικών Δεδομένων στην Δημόσια Διοίκηση

6.1 Συνοπτική Περιγραφή της μελέτης

Η παρούσα μελέτη επικεντρώνεται στην κατασκευή μοντέλων νευρωνικών δικτύων για την πρόβλεψη περιβαλλοντικών δεδομένων που αφορούν ατμοσφαιρικούς ρύπους, που είναι κρίσιμοι για τη λήψη περιβαλλοντικών αποφάσεων στη δημόσια διοίκηση και το περιβαλλοντικό σχεδιασμό. Η εφαρμογή της τεχνητής νοημοσύνης των τεχνολογιών της πληροφορίας σε διάφορους τομείς, όπως η περιβαλλοντική διοίκηση, έχει αυξηθεί σημαντικά τις τελευταίες δεκαετίες με την ανάπτυξη νέων τεχνικών μηχανικής μάθησης και εργαλείων για την κατασκευή μοντέλων νευρωνικών δικτύων (Kouziokas et al., 2017b, 2017c, 2017d, 2017e, 2017f, 2018). Σε αυτή την έρευνα, πραγματοποιείται η εφαρμογή των μοντέλων νευρωνικών δικτύων εμπρόσθιας διάδοσης για την πρόβλεψη περιβαλλοντικών δεδομένων.

Τα περιβαλλοντικά δεδομένα μπορούν να διαδραματίσουν σημαντικό ρόλο στη δημόσια διοίκηση, αλλά και στην περιβαλλοντική διαχείριση και τον σχεδιασμό, προωθώντας βιώσιμες στρατηγικές λήψης αποφάσεων. Τα επίπεδα των παραγόντων ρύπανσης επελέγησαν για να προβλεφθούν, δεδομένου ότι οι παράγοντες ρύπανσης συνδέονται με προβλήματα δημόσιας υγείας αυξημένης σημαντικότητας. Κατασκευάστηκαν διάφορα μοντέλα τεχνητών νευρωνικών δικτύων χρησιμοποιώντας διαφορετικές αρχιτεκτονικές σχετικά με τον αριθμό των νευρώνων στα κρυφά επίπεδα, τον αριθμό των κρυφών επιπέδων και τους νευρώνες εισόδου και εξόδου προκειμένου να δημιουργηθεί το βέλτιστο μοντέλο που θα προβλέπει αποτελεσματικά τα επίπεδα των περιβαλλοντικών δεδομένων. Ο πολυεπίπεδος αισθητήρας εμπρόσθιας διάδοσης (Feedforward Multilayer Perceptron) χρησιμοποιήθηκε σε αυτή την έρευνα καθώς θεωρείται ως ο καταλληλότερος για την πρόβλεψη χρονοσειρών σύμφωνα με πολλούς ερευνητές (Koskela et al., 1996; Tang and Fishwick, 1993). Το βέλτιστο μοντέλο πρόβλεψης που αναπτύχθηκε μπορεί να είναι πολύτιμο για τη δημόσια διοίκηση, καθώς μπορεί να χρησιμοποιηθεί ως εργαλείο για πιο αποτελεσματική περιβαλλοντική διαχείριση αλλά και για την υιοθέτηση προληπτικών μέτρων και πολιτικών για την προστασία της δημόσιας υγείας.

6.2 Εισαγωγή – Σχετική Έρευνα

Η σημασία των περιβαλλοντικών πληροφοριών είναι υψηλής σημασίας στη δημόσια διοίκηση, αλλά και στην περιβαλλοντική διαχείριση και το σχεδιασμό για διάφορους λόγους.

Τα περιβαλλοντικά δεδομένα και ειδικότερα οι πληροφορίες σχετικά με τους παράγοντες ρύπανσης μπορούν να μελετηθούν και να αναλυθούν προκειμένου να υιοθετηθούν οι κατάλληλες περιβαλλοντικές στρατηγικές για την προστασία της ανθρώπινης υγείας καθώς και του περιβάλλοντος και των φυσικών πόρων (Council Directive, 1992; Evans et al., 2013; Kalantzi et al., 2011).

Η τεχνητή νοημοσύνη έχει εφαρμοστεί σε διάφορους επιστημονικούς τομείς. Η ανάπτυξη νέων τεχνολογιών και εργαλείων μάθησης για τα νευρωνικά δίκτυα έχει αυξήσει ραγδαία την εφαρμογή της τεχνητής νοημοσύνης σε πολλούς τομείς.

Τα περιβαλλοντικά δεδομένα μπορούν να διαδραματίσουν σημαντικό ρόλο στη δημόσια διοίκηση, αλλά και στην περιβαλλοντική διαχείριση και τον σχεδιασμό, προωθώντας βιώσιμες στρατηγικές λήψης αποφάσεων στη διαχείριση του περιβάλλοντος (Azapagic and Perdan, 2000; Kouziokas, 2016d; Kouziokas, 2017e, 2017f; Kouziokas et al., 2018; Varotsos, 2005).

Οι Cortés et al. (2000) έδειξαν ότι η τεχνητή νοημοσύνη μπορεί να εφαρμοστεί στα συστήματα υποστήριξης περιβαλλοντικών αποφάσεων για την επίλυση διαφόρων ειδών περιβαλλοντικών ζητημάτων. Η ίδια έρευνα έδειξε ότι η ενσωμάτωση διαφόρων μεθόδων τεχνητής νοημοσύνης σε στατιστικά μοντέλα σε ένα ενιαίο σύστημα παρέχει μεγαλύτερη ακρίβεια και αξιοπιστία.

Σε μια άλλη έρευνα οι Kanevski et al. (2004), επικεντρώθηκαν στα ζητήματα που έχουν σχέση με την λήψη αποφάσεων και μελέτησαν την εξόρυξη και την μοντελοποίηση δεδομένων περιβάλλοντος με βάση αλγορίθμους μηχανικής μάθησης και μεθόδους γεωστατιστικής. Στη μελέτη αυτή, αναπτύχθηκαν μοντέλα τεχνητών νευρωνικών δικτύων για την ανάλυση και χαρτογράφηση χωρικών περιβαλλοντικών δεδομένων.

Επίσης, μια άλλη έρευνα οι Graf-Jaccottet and Jaunin (1998), μελέτησαν την εφαρμογή προβλεπτικών μεθόδων βασισμένων σε μοντέλα αυτοπαλινδρόμησης για την πρόβλεψη των τιμών του όζοντος για χρονικό διάστημα πέντε ετών σε ημερήσια βάση. Τα αποτελέσματα έδειξαν καλή ακρίβεια στην πρόβλεψη των τιμών.

Σε αυτή την έρευνα, εξετάζεται η δυνατότητα εφαρμογής μοντέλων πρόβλεψης τεχνητών νευρωνικών δικτύων για την πρόβλεψη περιβαλλοντικών παραγόντων ρύπανσης. Στις επόμενες ενότητες περιγράφεται το θεωρητικό υπόβαθρο, η ερευνητική μεθοδολογία, καθώς και οι παράγοντες που εξετάστηκαν για την κατασκευή του

βέλτιστου μοντέλου πρόβλεψης, όπως η τοπολογία του νευρωνικού δικτύου και ο αλγόριθμος εκπαίδευσης.

6.3 Θεωρητικό πλαίσιο

6.3.1 Τεχνητά Νευρωνικά Δίκτυα

Τα τεχνητά νευρωνικά δίκτυα (ANNs) ορίζονται ως τεχνητά υπολογιστικά συστήματα που προσομοιώνουν το νευρωνικά κύτταρα του ανθρώπινου εγκεφάλου. Ένα νευρωνικό δίκτυο επεξεργάζεται τα δεδομένα από τις παραμέτρους εισόδου.

Τα δεδομένα εισόδου μετακινούνται μέσω των συνδέσεων και των νευρώνων και αφού υποστούν την κατάλληλη επεξεργασία, παράγονται τα αποτελέσματα εξόδου σύμφωνα με τις παραμέτρους εισόδου (Svozil et al., 1997). Στην παρούσα μελέτη χρησιμοποιούνται τα τεχνητά νευρωνικά δίκτυα για την πρόβλεψη των επιπέδων των παραγόντων ρύπανσης του περιβάλλοντος.

6.3.2 Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης

Σε αυτήν την έρευνα χρησιμοποιήθηκε ένας Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης καθώς πολλές μελέτες έδειξαν ότι είναι ο πιο κατάλληλος για προβλήματα που έχουν να κάνουν με προβλέψεις των χρονοσειρών.

Σε ένα απλό νευρωνικό δίκτυο πολλαπλών στρωμάτων εμπρόσθιας διάδοσης, οι νευρώνες ή οι κόμβοι συνδέονται μόνο προς την κατεύθυνση προς τα εμπρός και αποτελούνται από ένα επίπεδο εισόδου, ένα κρυμμένο επίπεδο και ένα επίπεδο εξόδου. Κάθε επίπεδο αποτελείται από έναν αριθμό νευρώνων (Blum and Kwan, 1991; Hornik, 1991).

6.3.3 Αλγόριθμος Κλιμακούμενων Συζυγών Κλίσεων

Ο αλγόριθμος Κλιμακούμενων Συζυγών Κλίσεων (Scaled Conjugate Gradient, SCG) εισήχθη το 1993 από τον Møller (1993). Για την εκπαίδευση του τεχνητού νευρωνικού δικτύου χρησιμοποιήθηκε ο αλγόριθμος Κλιμακούμενων Συζυγών Κλίσεων (Scaled Conjugate Gradient – SCG), ο οποίος θεωρείται ως ένας από τους γρηγορότερους αλγορίθμους μάθησης μεταξύ πολλών άλλων αλγορίθμων εκπαίδευσης (Møller, 1993).

Οι τεχνικές συζυγών κλίσεων (conjugate gradient) παράγουν ταχύτερη σύγκλιση από τις μεθόδους καθόδου βασισμένης στην κλίση (gradient descent), εφαρμόζοντας μια αναζήτηση σε όλες τις κατευθύνσεις της κλίσης, προκειμένου να προσδιοριστεί το μέγεθος της κλιμάκωσης του βήματος. Ο αλγόριθμος Κλιμακούμενων Συζυγών Κλίσεων (Scaled Conjugate Gradient) ελαχιστοποιεί τις συναρτήσεις του στόχου και χρησιμοποιεί μια μέθοδο κλιμάκωσης του μεγέθους βημάτων για να επιτύχει μια ταχύτερη διαδικασία εκπαίδευσης (Andrei, 2007; Steihaug, 1983).

6.3.4 Περιβαλλοντικοί ρύποι

Από τους σημαντικότερους παράγοντες ρύπανσης του αέρα στο αστικό περιβάλλον είναι το όζον (O_3). Αυτοί οι παράγοντες έχουν αρνητικές επιπτώσεις στην ανθρώπινη υγεία, όπως έδειξαν οι έρευνες.

Σύμφωνα με διάφορες έρευνες, το όζον ως ρυπογόνος παράγοντας, έχει συνδεθεί με την καθημερινή θνησιμότητα σε διάφορα μέρη του κόσμου, όπως το Λονδίνο (Williams et al., 2014) και το Δέλτα του ποταμού Περλ της νότιας Κίνας (Tao et al., 2012). Επίσης, σύμφωνα με άλλη έρευνα (Mirowsky et al., 2016), οι προφλεγμονώδεις και οξειδωτικοί μεσολαβητές σε ανθρώπινα βρογχικά επιθηλιακά κύτταρα προκαλούνται από το όζον και το διοξείδιο του αζώτου στην ατμόσφαιρα.

Επίσης, μια άλλη έρευνα έδειξε τη θετική συσχέτιση του όζοντος και του ποσοστού θνησιμότητας σε 21 πόλεις της Ανατολικής Ασίας (Chen et al., 2014). Επιπλέον, μια έρευνα έδειξε πως το όζον συνδέεται με την απελευθέρωση των προφλεγμονωδών

μεσολαβητών από τα βρογχικά επιθηλιακά κύτταρα των μη ατοπικών και ατοπικών ασθενών με άσθμα (Bayram et al., 2001).

6.4 Μεθοδολογία Έρευνας

Η μεθοδολογία της έρευνας αποτελείται από τα παρακάτω στάδια: συλλογή και προετοιμασία των δεδομένων, τροφοδότηση των μοντέλων πρόβλεψης τεχνητής νοημοσύνης: νευρωνικών δικτύων και Μηχανών Διανυσμάτων Υποστήριξης, για την ανάπτυξη μοντέλων πρόβλεψης, βελτιστοποίηση των μοντέλων με τεχνικές AutoML, χρήση των βέλτιστων μοντέλων τεχνητής νοημοσύνης για την πρόβλεψη των συγκεντρώσεων των ρύπων. Στο πρώτο στάδιο, συγκεντρώθηκαν, προετοιμάστηκαν και κανονικοποιήθηκαν τα περιβαλλοντικά δεδομένα.

Στη δεύτερη φάση, χρησιμοποιήθηκαν τα τεχνητά νευρωνικά δίκτυα εμπρόσθιας διάδοσης για να κατασκευαστεί το βέλτιστο πρόβλεψης τεχνητής νοημοσύνης, το οποίο βελτιστοποιήθηκε με τον αλγόριθμο PSO. Επίσης, χρησιμοποιήθηκαν μοντέλα Μηχανών Διανυσμάτων Υποστήριξης, για την ανάπτυξη μοντέλων πρόβλεψης που βελτιστοποιήθηκαν με τον αλγόριθμο Bayes. Για την ανάπτυξη των μοντέλων τεχνητής νοημοσύνης χρησιμοποιήθηκε το λογισμικό MATLAB. Στην επόμενη φάση, χρησιμοποιήθηκαν τα βέλτιστα μοντέλα για την πρόβλεψη των συγκεντρώσεων του όζοντος και συγκρίθηκαν μεταξύ τους ως προς την ακρίβεια πρόβλεψης.

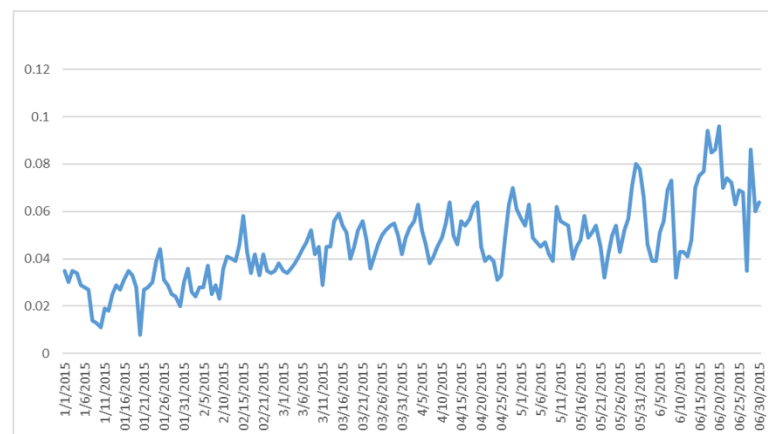
6.5 Αποτελέσματα

6.5.1 Συλλογή Δεδομένων – Περιοχή Μελέτης

Τα μοντέλα του τεχνητών νευρωνικών δικτύων κατασκευάστηκαν χρησιμοποιώντας ως παραμέτρους εισόδου αρκετούς παράγοντες που επηρεάζουν τα επίπεδα του όζοντος, όπως η μέγιστη θερμοκρασία, το μήκος της ημέρας (ώρες ηλιοφάνειας) και τα ιστορικά δεδομένα του όζοντος (O_3) και του διοξειδίου του αζώτου (NO_2) (Crouse et al., 2015). Επιλέχθηκε η πόλη του Λος Άντζελες για την

συλλογή των δεδομένων, διότι μετά από σχετική έρευνα στο διαδίκτυο, μόνο εκεί βρέθηκαν ιστορικά δεδομένα και των τεσσάρων παραγόντων εισόδου για ένα επαρκές χρονικό διάστημα που να μπορούν να εκπαιδευτούν τα μοντέλα τεχνητής νοημοσύνης, σε αντίθεση με την Ελλάδα, που δεν βρέθηκαν ιστορικά δεδομένα ηλιοφάνειας για ικανό χρονικό διάστημα, που να αρκεί για την εκπαίδευση των μοντέλων, και που αποτελεί έναν πολύ σημαντικό παράγοντα στις φωτοχημικές αντιδράσεις που πραγματοποιούνται για την παραγωγή του αέριου ρύπου του όζοντος (Crouse et al., 2015). Επίσης, πολλές μελέτες που αφορούν το φωτοχημικό νέφος, κύριο συστατικό του οποίου είναι το όζον έχουν επικεντρωθεί στην πόλη του Λος Άντζελες, αφού εκεί πρωτο-παρατηρήθηκε το φαινόμενο αυτό γι' αυτό και ονομάζεται επίσης και "νέφος τύπου Λος Άντζελες" (Calvert 1976).

Τα περιβαλλοντικά δεδομένα ανακτήθηκαν από την Υπηρεσία Προστασίας του Περιβάλλοντος των ΗΠΑ (U.S. Environmental Protection Agency 2018) για την πόλη του Λος Άντζελες. Τα δεδομένα των παραγόντων εισόδου επεξεργάστηκαν και προετοιμάστηκαν για να χρησιμοποιηθούν ως δεδομένα εισόδου σε μοντέλα τεχνητών νευρωνικών δικτύων. Τα περιβαλλοντικά δεδομένα συλλέχθηκαν και προετοιμάστηκαν (ελέγχθηκαν για ασυνέπειες, διπλότυπα και κενά) για τους πρώτους έξι μήνες του έτους 2015, σε ημερήσια βάση, για την πόλη του Λος Άντζελες. Η παρακάτω εικόνα, δείχνει τις ημερήσιες τιμές συγκέντρωσης του όζοντος σε ppm στην περιοχή μελέτης.



Εικόνα 6.1 Αναπαράσταση της διακύμανσης των τιμών της ημερήσιας συγκέντρωσης του όζοντος (ppm) στο δείγμα στην περιοχή μελέτης.

6.5.2 Διανύσματα Εισόδου - Εξόδου

Το διάνυσμα εισόδου, θα αποτελείται από την ημερήσια τιμή μέγιστης συγκέντρωση όζοντος 8 ωρών (ppm) της προηγούμενης ημέρας O_{3n-1} , την ημερήσια μέγιστη συγκέντρωση διοξειδίου του αζώτου (ppb) της προηγούμενης ημέρας NO_{2n-1} , το μήκος της ηλιοφάνειας της προηγούμενης ημέρας σε ώρες D_{n-1} και την μέγιστη ημερήσια θερμοκρασία της προηγούμενης ημέρας T_{n-1} .

$$X_i = (O_{3n-1}, NO_{2n-1}, D_{n-1}, T_{n-1}) \quad (6.1)$$

Οι τιμές εξόδου Y_i αφορούν τις τιμές του όζοντος της επόμενης ημέρας:

$$Y_i = (O_{3n}) \quad (6.2)$$

Τα περιβαλλοντικά δεδομένα χωρίστηκαν σε τρία διαφορετικά μέρη. Το 60% των πρωτογενών δεδομένων χρησιμοποιήθηκε ως σύνολο εκπαίδευσης, το 20% για το σύνολο επικύρωσης και 20% για το σύνολο δοκιμών. Το σύνολο δεδομένων εκπαίδευσης χρησιμοποιήθηκε για να εκπαιδεύσει το νευρωνικό δίκτυο με προηγούμενα ιστορικά δεδομένα. Το σύνολο επικύρωσης χρησιμοποιήθηκε για την αξιολόγηση των μοντέλων τεχνητών νευρωνικών δικτύων. Ο αλγόριθμος Κλιμακούμενων Συζυγών Κλίσεων (Scaled Conjugate Gradient Algorithm) χρησιμοποιήθηκε ως ο καλύτερος αλγόριθμος εκπαίδευσης σύμφωνα με τα αποτελέσματα σύγκρισης με άλλους αλγορίθμους.

6.5.3 Βελτιστοποίηση Τοπολογίας Νευρωνικού Μοντέλου

Για την βελτιστοποίηση της τοπολογίας του νευρωνικού δικτύου εφαρμόστηκε η μέθοδος αναζήτησης πλέγματος (Grid Search) (Basnayake et al., 2017). Το πλεονέκτημα της μεθόδου είναι ότι αυτή είναι πλήρης, δηλαδή ερευνά κάθε δυνατό συνδυασμό για να βρεθεί ο βέλτιστος. Επίσης, ένα άλλο πλεονέκτημα είναι ότι μπορεί να βελτιστοποιήσει υπερπαραμέτρους που αποτελούν μίγμα κατηγορικών και μη κατηγορικών υπερπαραμέτρων σε αντίθεση με άλλες βελτιστοποιήσεις αυτόματης μηχανικής μάθησης (AutoML) όπως η βελτιστοποίηση Bayes, οι οποίες αντιμετωπίζουν προβλήματα όταν υπάρχει συνδυασμός τέτοιου είδους υπερπαραμέτρων (Feurer et al., 2015; Guyon et al., 2016; Mendoza et al., 2016). Όπως για παράδειγμα, οι κατηγορικές μεταβλητές στο παρόν πρόβλημα θα είναι: τύπος συνάρτησης μεταφοράς, τύπος αλγορίθμου εκπαίδευσης

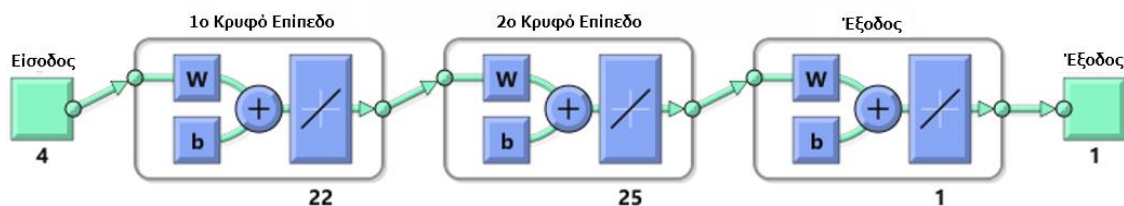
που πρέπει να βελτιστοποιηθούν σε συνδυασμό με μη κατηγορικές μεταβλητές όπως: ο αριθμός νευρώνων στο κρυφό επίπεδο. Οι υπερπαραμέτροι που ερευνηθήκαν αφορούν: τον αριθμό νευρώνων σε κάθε κρυφό επίπεδο, τις πιο συνηθισμένες συναρτήσεις μεταφοράς για τα ένα έως δύο κρυφά επίπεδα: Η εφαπτομενική – σιγμοειδής συνάρτηση μεταφοράς (Tan-Sigmoid Transfer Function – TSTF), η λογιστική – σιγμοειδής συνάρτηση μεταφοράς (Log-Sigmoid Transfer Function – LSTF), η γραμμική συνάρτηση μεταφοράς (Linear Transfer Function – LTF). Επίσης, ερευνήθηκαν οι πιο συνηθισμένοι αλγόριθμοι μάθησης για την εκπαίδευση κάθε νευρωνικού μοντέλου και ήταν οι: αλγόριθμος Κλιμακούμενων Συζυγών Κλίσεων (Scaled Conjugate Gradient – SCG), BFGS Quasi-Newton (BFGS-QN), Ελαστικής Οπισθοδιάδοσης (Resilient Backpropagation – RB), Levenberg Marquardt (LM). Η Ρίζα του Μέσου Τετραγωνικού Σφάλματος (Root Mean Square Error - RMSE) χρησιμοποιήθηκε για την εκτίμηση του σφάλματος πρόβλεψης για κάθε νευρωνικό μοντέλο που αναπτύχθηκε. Στον παρακάτω πίνακα, απεικονίζονται τα βέλτιστα νευρωνικά μοντέλα που βρέθηκαν για κάθε αλγόριθμο που δοκιμάστηκε χρησιμοποιώντας διαφορετικές τοπολογίες και συναρτήσεις ενεργοποίησης. Όπως φαίνεται από τον πίνακα το βέλτιστο μοντέλο με το μικρότερο σφάλμα επικύρωσης είναι το ANN1.

Πίνακας 6.1. Τα βέλτιστα νευρωνικά μοντέλα που βρέθηκαν για κάθε αλγόριθμο που δοκιμάστηκε χρησιμοποιώντας διαφορετικές τοπολογίες και συναρτήσεις ενεργοποίησης.

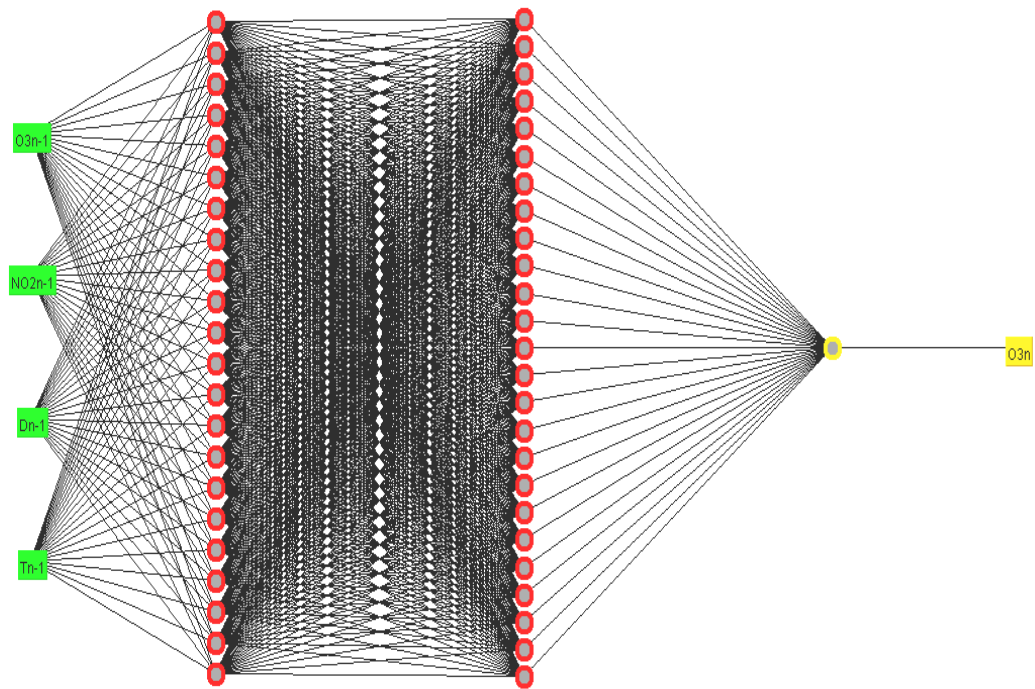
Μοντέλο	Αλγόριθμος Εκπαίδευσης	Αριθμός Κρυφών Επιπέδων	Αριθμός Νευρώνων στο 1 ^ο Κρυφό Επίπεδο	Αριθμός Νευρώνων στο 2 ^ο Κρυφό Επίπεδο	Συναρτήσεις Μεταφοράς (Transfer Functions)	Σφάλμα Εκπαίδευσης (Training Error)	Σφάλμα Επικύρωσης (Validation Error)
ANN1	SCG	2	22	25	LTF- LTF	0.208	0.180
ANN2	BFGS- QN	2	23	19	LTF- LTF	189.546	191.021
ANN3	RB	2	29	15	TSAF- LTF	89.157	95.248
ANN4	LM	2	23	21	TSAF- LTF	15.249	17.254

Στο βέλτιστο μοντέλο νευρωνικού δικτύου, ο αλγόριθμος εκπαίδευσης ήταν ο αλγόριθμος Κλιμακούμενων Συζυγών Κλίσεων. Η καλύτερη τοπολογία ήταν αυτή με δύο κρυφά επίπεδα. Η βέλτιστη τοπολογία αποτελείται από 22 νευρώνες στο πρώτο κρυφό επίπεδο και την Γραμμική Συνάρτηση Μεταφοράς (Linear Transfer Function – LTF) ως συνάρτηση μεταφοράς και με 25 νευρώνες στο δεύτερο κρυφό επίπεδο και την Γραμμική Συνάρτηση Μεταφοράς (Linear Transfer Function – LTF) ως συνάρτηση μεταφοράς.

Το βέλτιστο μοντέλο αξιολογήθηκε με τη Ρίζα του Μέσου Τετραγωνικού Σφάλματος (Root Mean Squared Error – RMSE), σύμφωνα με την σύγκριση που πραγματοποιήθηκε μεταξύ όλων των άλλων αναπτυχθέντων μοντέλων νευρωνικών δικτύων. Η Ρίζα του Μέσου Τετραγωνικού Σφάλματος (RMSE) του βέλτιστου μοντέλου βρέθηκε να είναι 0.180 στην εποχή 27. Στην επόμενη εικόνα, απεικονίζεται η τοπολογία του μοντέλου βέλτιστου τεχνητού νευρωνικού δικτύου.

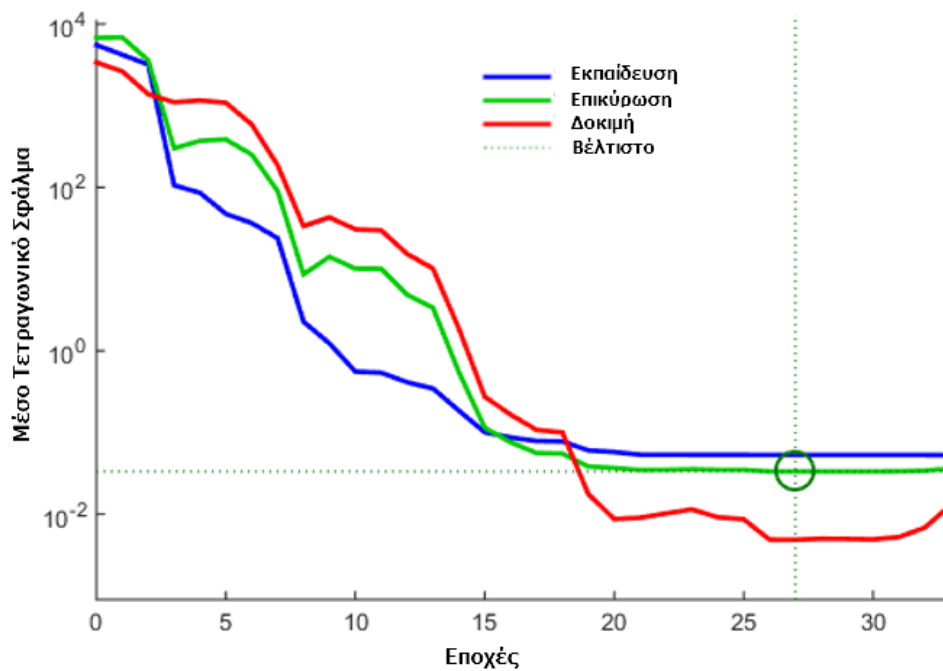


Εικόνα 6.2 Το σχήμα αναπαριστά την τοπολογία του βέλτιστου μοντέλου τεχνητού νευρωνικού δικτύου για την πρόβλεψη των τιμών του όζοντος O_3 με την χρήση δύο κρυφών επιπέδων (hidden layers). Το πρώτο κρυφό επίπεδο αποτελείται από 22 νευρώνες και χρησιμοποιεί την γραμμική συνάρτηση μεταφοράς, ενώ το δεύτερο επίπεδο αποτελείται από 25 νευρώνες και χρησιμοποιεί την γραμμική συνάρτηση μεταφοράς.



Εικόνα 6.3 Το σχήμα αναπαριστά την τοπολογία του βέλτιστου μοντέλου τεχνητού νευρωνικού δικτύου για την πρόβλεψη των τιμών του όζοντος O_3 , τα επίπεδα εισόδου, εξόδου, τα κρυφά επίπεδα και τους νευρώνες των κρυφών επιπέδων.

Στην επόμενη εικόνα, απεικονίζεται η γραφική παράσταση απόδοσης του βέλτιστου νευρωνικού μοντέλου κατά την διαδικασία της εκπαίδευσης με τον αλγόριθμο Κλιμακούμενων Συζυγών Κλίσεων, ως αλγόριθμος εκπαίδευσης, σύμφωνα με το Μέσο Τετραγωνικό Σφάλμα (Mean Squared Error (MSE)).



Εικόνα 6.4 Η γραφική παράσταση απόδοσης του βέλτιστου μοντέλου νευρωνικού δικτύου.

6.5.4 Βελτιστοποίηση Νευρωνικού Δικτύου με Σμήνος Σωματιδίων

Για την βελτιστοποίηση των βαρών του νευρωνικού δικτύου εφαρμόστηκε η μέθοδος Βελτιστοποίησης Σμήνους Σωματιδίων (Particle Swarm Optimization, PSO). Ο αλγόριθμος βελτιστοποίησης σμήνους σωματιδίων (PSO) περιγράφηκε πιο αναλυτικά στην ενότητα 4.4.4. Για την βελτιστοποίηση των παραμέτρων του Σμήνους Σωματιδίων, εφαρμόστηκε η μέθοδος αναζήτησης πλέγματος (Grid Search), που περιγράφηκε παραπάνω, για τις παραμέτρους που αφορούν: τον πληθυσμό των σωματιδίων του σμήνους, τους συντελεστές μάθησης c_1 , c_2 , τον συντελεστή βάρους αδράνειας (inertia weight coefficient). Οι πληθυσμοί των σωματιδίων σμήνους που εξετάστηκαν ήταν 20, 30, 40, 50, που είναι οι πιο συνηθισμένες τιμές σύμφωνα με τη βιβλιογραφία (Fourie and Groenwold, 2002). Οι συντελεστές μάθησης c_1 , c_2 , ορίστηκαν στην τιμή 2.0, όπως προτείνεται από τους Kennedy and Eberhart (1995). Οι τιμές του συντελεστή βάρους αδράνειας (inertia weight coefficient) εξετάστηκαν στο διάστημα [0.8, 1.4] με βήμα 0.1,

όπως προτείνεται από τους Shi and Eberhart, (1998). Οι παράμετροι που παρήγαγαν τα βέλτιστα αποτελέσματα ήταν: μέγεθος πληθυσμού: $P = 50$, βάρος αδράνειας $w = 1.00$, συντελεστής προσωπικής μάθησης $c_1 = 2.0$, ολικός συντελεστής εκμάθησης $c_2 = 2.0$. Ο παρακάτω πίνακας, παρουσιάζει τα πέντε καλύτερα μοντέλα του αλγορίθμου βελτιστοποίησης PSO, που βρέθηκαν χρησιμοποιώντας διαφορετικές τοπολογίες παραμέτρων. Το βέλτιστο μοντέλο της βελτιστοποίησης Σμήνους Σωματιδίων ήταν το μοντέλο 5. Η Ρίζα του Μέσου Τετραγωνικού Σφάλματος (RMSE) του βελτιστοποιημένου μοντέλου νευρωνικών δικτύων βρέθηκε να είναι 0.0982.

Πίνακας 6.2. Τα πέντε καλύτερα νευρωνικά μοντέλα του αλγορίθμου βελτιστοποίησης PSO που βρέθηκαν χρησιμοποιώντας διαφορετικές τοπολογίες παραμέτρων για την βελτιστοποίηση των βαρών του νευρωνικού δικτύου.

Μοντέλο	Πληθυσμός Σμήνους P	c_1	c_2	w	Βέλτιστο Κόστος (Best Cost)
<i>Μοντέλο 1</i>	20	2.0	2.0	1.30	0.1826
<i>Μοντέλο 2</i>	30	2.0	2.0	1.00	0.1985
<i>Μοντέλο 3</i>	40	2.0	2.0	0.90	0.1812
<i>Μοντέλο 4</i>	40	2.0	2.0	1.10	0.1355
<i>Μοντέλο 5</i>	50	2.0	2.0	1.00	0.0982

6.5.5 Βελτιστοποιημένα Μοντέλα Μηχανών Διανυσμάτων Υποστήριξης

Τα Μοντέλα Μηχανών Διανυσμάτων Υποστήριξης βελτιστοποιήθηκαν εφαρμόζοντας τον αλγόριθμο βελτιστοποίησης Bayes, όπως περιγράφηκε στην ενότητα 4.4.7, και υπολογίστηκαν οι υπερπαραμέτροι για την εύρεση των βέλτιστων μοντέλων. Χρησιμοποιήθηκαν οι κάτωθι βασικοί πυρήνες, οι οποίοι θεωρούνται ως οι πιο κύριοι στην υλοποίηση των Μηχανών Διανυσμάτων Υποστήριξης στην παλινδρόμηση: γραμμικός, πολυωνυμικός, ακτινικός και γκαουσιανός χρησιμοποιώντας το λογισμικό

MATLAB. Για την αξιολόγηση των μοντέλων χρησιμοποιήθηκε η ελάχιστη τιμή της αντικειμενικής συνάρτησης. Με τον αλγόριθμο βελτιστοποίησης Bayes βελτιστοποιήθηκαν οι κάτωθι υπερ-παράμετροι των μοντέλων Μηχανών Διανυσμάτων Υποστήριξης: Η τιμή κόστους (C), η παράμετρος epsilon (ϵ), η παράμετρος γάμμα (γ) των πυρήνων εκτός του γραμμικού και ο βαθμός του πολυωνύμου για τον πολυωνυμικό πυρήνα.

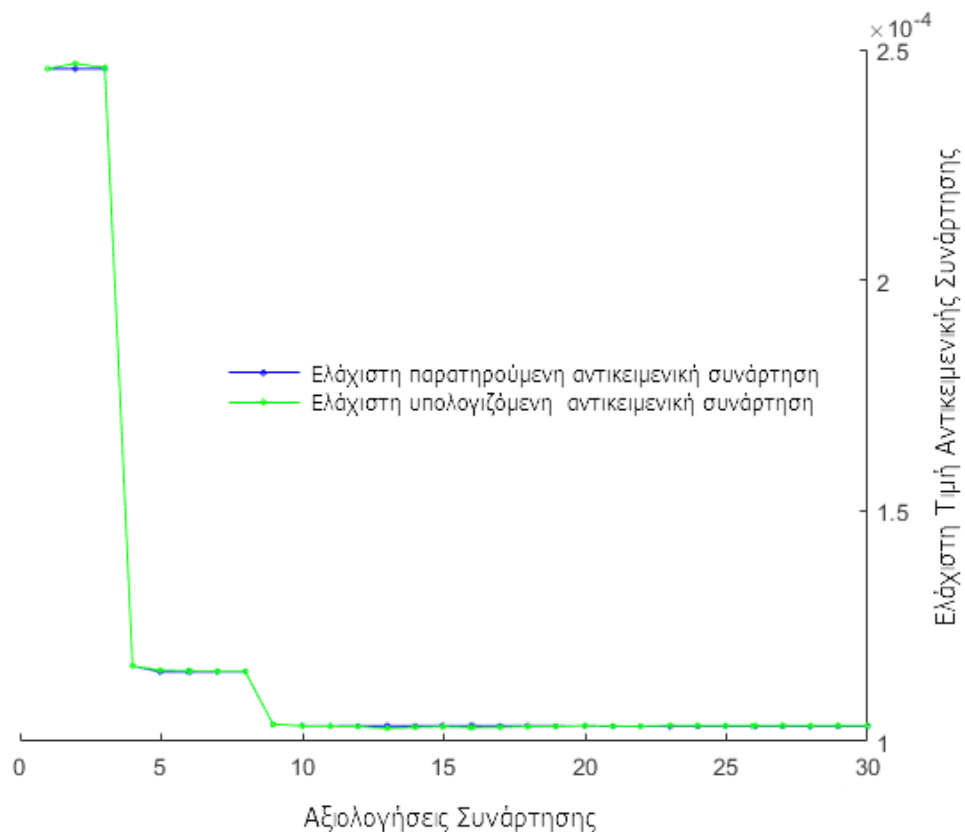
Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον Γκαουσιανό πυρήνα παρουσιάζονται στον παρακάτω πίνακα.

Πίνακας 6.3. Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον Γκαουσιανό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμενική Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζόμ.	C	γ	ϵ
1	0.00024583	0.00024583	0.00024583	636.210	0.445	0.000177
2	0.00025704	0.00024583	0.00024696	1.676	110.480	0.049278
3	0.00025709	0.00024583	0.00024608	0.423	0.011	0.000672
4	0.00011632	0.00011632	0.00011634	39.161	966.400	0.000017
5	0.00011509	0.00011509	0.00011542	42.493	951.970	0.000049
6	0.00030296	0.00011509	0.00011528	0.003	668.760	0.007003
7	0.00025704	0.00011509	0.00011520	257.530	992.230	0.634410
8	0.00025704	0.00011509	0.00011516	54.016	20.726	0.166310
9	0.00010371	0.00010371	0.00010374	344.150	939.090	0.000025
10	0.00010335	0.00010335	0.00010333	979.970	903.780	0.000060
11	0.00010345	0.00010335	0.00010331	954.030	988.780	0.000016
12	0.00010348	0.00010335	0.00010327	732.900	386.860	0.000031
13	0.00010396	0.00010335	0.00010283	975.150	668.320	0.000030

Πίνακας 6.3. (συνέχεια) Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον Γκαουσιανό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμενική Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζόμ.	C	γ	ε
14	0.00010353	0.00010335	0.00010309	553.570	848.520	0.000030
15	0.00010376	0.00010335	0.00010320	704.210	602.310	0.000034
16	0.00010367	0.00010335	0.00010304	945.690	437.550	0.000015
17	0.00010361	0.00010335	0.00010309	836.960	970.730	0.000040
18	0.00010404	0.00010335	0.00010324	852.590	656.410	0.000020
19	0.00025704	0.00010335	0.00010324	759.640	0.001	0.933820
20	0.00010461	0.00010335	0.00010345	782.690	592.470	0.000033
21	0.00010326	0.00010326	0.00010332	977.710	935.460	0.000030
22	0.00025704	0.00010326	0.00010332	0.001	0.001	0.969000
23	0.00010381	0.00010326	0.00010345	670.280	993.520	0.000034
24	0.00025377	0.00010326	0.00010343	0.001	1.841	0.000016
25	0.00025704	0.00010326	0.00010342	834.420	0.001	0.000017
26	0.00025704	0.00010326	0.00010344	0.001	1.283	1.449000
27	0.00010366	0.00010326	0.00010342	381.710	349.290	0.000015
28	0.00010338	0.00010326	0.00010340	442.360	546.120	0.000025
29	0.00010452	0.00010326	0.00010344	670.250	505.350	0.000017
30	0.00025751	0.00010326	0.00010344	0.001	0.001	0.000016



Εικόνα 6.5. Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον Γκαουσιανό πυρήνα (Gaussian Kernel) Μηχανών Διανυσμάτων Υποστήριξης.

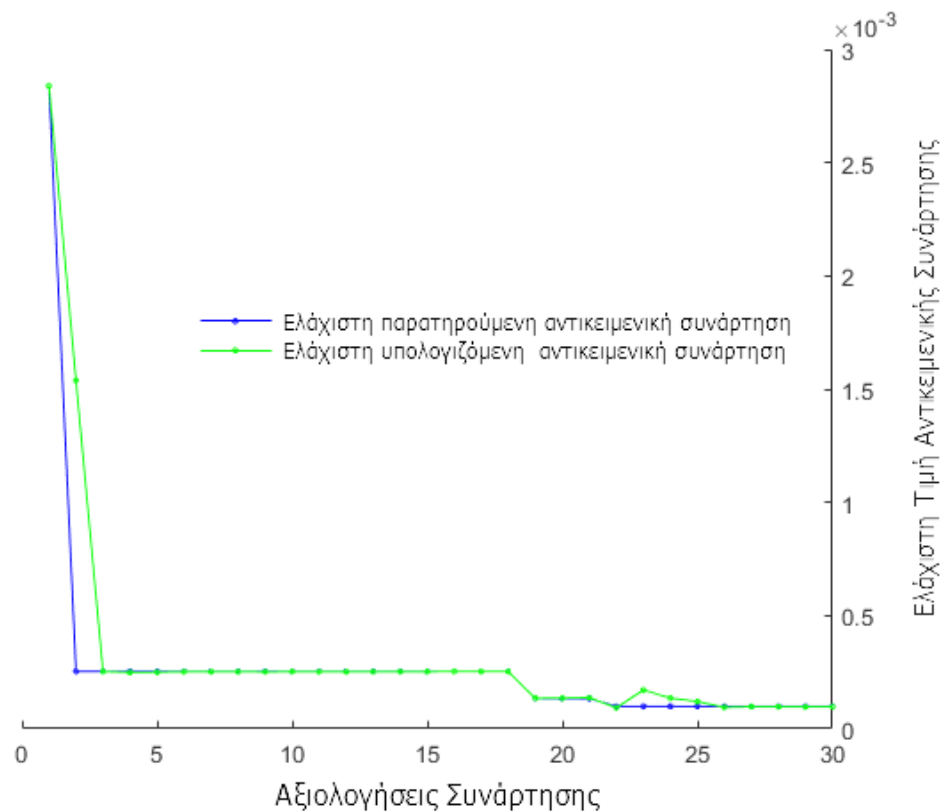
Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον ακτινικό πυρήνα παρουσιάζονται στον παρακάτω πίνακα.

Πίνακας 6.4. Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον ακτινικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμενική Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζόμ.	C	γ	ϵ
1	0.00283940	0.00283940	0.00283940	583.470	8.621	0.000039
2	0.00025335	0.00025335	0.00153760	46.389	0.001	0.072654

Πίνακας 6.4. (συνέχεια) Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον ακτινικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμενική Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζόμ.	C	γ	ε
3	0.00025335	0.00025335	0.00025266	0.112	54.297	0.058402
4	0.00025335	0.00025335	0.00024929	454.720	37.135	1.268000
5	0.00025335	0.00025335	0.00025000	6.379	0.028	0.345690
6	0.00025335	0.00025335	0.00025201	0.625	1.597	0.132370
7	0.00025335	0.00025335	0.00025209	0.036	0.307	0.694630
8	0.00025335	0.00025335	0.00025210	45.750	0.010	1.479800
9	0.00025335	0.00025335	0.00025164	0.002	897.280	0.112620
10	0.00025335	0.00025335	0.00025200	1.149	0.011	0.102330
11	0.00025335	0.00025335	0.00025204	9.805	8.543	0.778090
12	0.00025335	0.00025335	0.00025208	133.830	173.300	1.482600
13	0.00025335	0.00025335	0.00025231	0.001	0.075	0.094545
14	0.00025335	0.00025335	0.00025230	335.390	4.872	0.247950
15	0.00025335	0.00025335	0.00025232	805.180	0.087	0.797270
16	0.00025346	0.00025335	0.00025356	1.011	0.103	0.001195
17	0.00032908	0.00025335	0.00025305	49.348	2.432	0.004577
18	0.00025342	0.00025335	0.00025305	3.050	0.013	0.001746
19	0.00013375	0.00013375	0.00013472	0.012	19.958	0.001470
20	0.00025562	0.00013375	0.00013552	0.001	0.020	0.001677
21	0.00023832	0.00013375	0.00013764	0.085	449.560	0.003676
22	0.00009898	0.00009898	0.00009274	817.510	873.000	0.001547
23	0.00025777	0.00009898	0.00017148	0.011	926.310	0.001568
24	0.00009941	0.00009898	0.00013559	497.780	949.960	0.001564
25	0.00009875	0.00009875	0.00012061	918.120	862.360	0.001710
26	0.00009913	0.00009875	0.00009500	538.510	858.070	0.001663
27	0.00032710	0.00009875	0.00009753	815.900	20.050	0.001486
28	0.00025335	0.00009875	0.00009747	943.290	962.250	0.120680
29	0.00025562	0.00009875	0.00009743	0.003	0.001	0.015913
30	0.00025335	0.00009875	0.00009744	0.001	867.370	1.381200



Εικόνα 6.6. Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον ακτινικό πυρήνα (RBF Kernel) Μηχανών Διανυσμάτων Υποστήριξης.

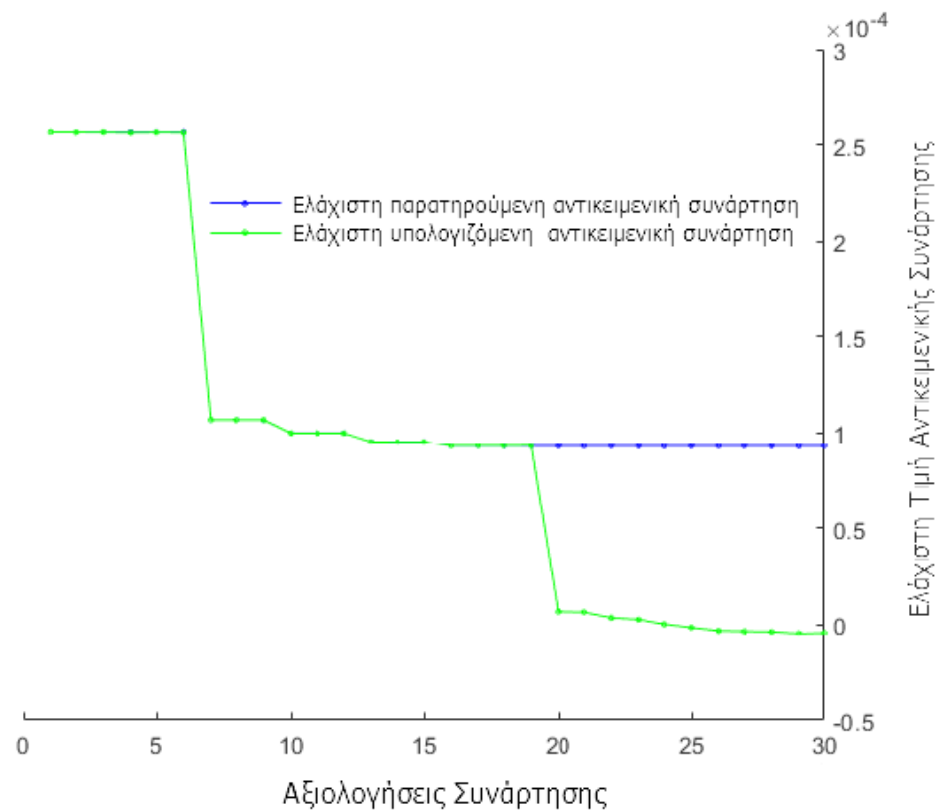
Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον πολωνυμικό πυρήνα παρουσιάζονται στον παρακάτω πίνακα.

Πίνακας 6.5. Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον πολωνυμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμ. Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζ.	C	γ	ϵ	Βαθμός Πολυων. r
1	0.00025632	0.00025632	0.00025632	359.050	134.510	0.290520	3
2	0.00025632	0.00025632	0.00025632	884.270	0.063	0.000080	2
3	0.00028283	0.00025632	0.00025629	0.002	92.253	0.042812	2

Πίνακας 6.5. (συνέχεια) Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον πολωνυμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμ. Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζ.	C	γ	ε	Βαθμός Πολυν. r
4	0.00025632	0.00025632	0.00025607	7.691	0.010	1.206400	1
5	0.00025632	0.00025632	0.00025626	48.780	0.001	0.000534	3
6	0.00025632	0.00025632	0.00025612	47.389	0.002	0.000050	3
7	0.00010664	0.00010664	0.00010668	36.279	389.730	0.014475	4
8	0.00011102	0.00010664	0.00010669	24.775	901.590	0.000153	2
9	0.00025632	0.00010664	0.00010668	38.185	998.780	1.430400	5
10	0.00009986	0.00009986	0.00009983	27.359	390.990	0.001703	1
11	0.00010500	0.00009986	0.00009983	45.609	944.710	0.001481	2
12	0.00014420	0.00009986	0.00009984	10.715	995.610	0.005810	4
13	0.00009526	0.00009526	0.00009529	43.966	80.511	0.002903	2
14	0.00009607	0.00009526	0.00009523	37.467	59.195	0.000334	3
15	0.00009531	0.00009526	0.00009522	27.972	46.073	0.000888	5
16	0.00008308	0.00008308	0.00009313	16.365	32.520	0.000020	3
17	0.00009515	0.00008308	0.00009317	17.334	38.323	0.000068	5
18	0.00009630	0.00008308	0.00009319	29.396	54.399	0.000016	2
19	0.00009684	0.00008308	0.00009325	14.361	85.672	0.000017	2
20	0.09798100	0.00008308	0.00000679	4.584	2.896	0.000017	3
21	0.00025632	0.00008308	0.00000653	3.760	0.028	0.000121	3
22	0.00025632	0.00008308	0.00000362	0.264	0.004	0.000359	4
23	0.00025632	0.00008308	0.00000278	1.847	0.045	0.000020	2
24	0.00025632	0.00008308	0.00000042	0.206	0.016	0.018636	3
25	0.00012754	0.00008308	0.00000137	1.074	237.940	0.005445	3
26	0.00025632	0.00008308	0.00000301	0.046	0.003	0.000020	2
27	0.00025632	0.00008308	0.00000331	1.983	613.000	0.630790	6
28	0.00025632	0.00008308	0.00000353	0.235	0.001	0.086386	2
29	0.00025632	0.00008308	0.00000447	0.024	0.006	0.000182	2
30	0.00014007	0.00008308	0.00000415	0.852	302.740	0.000025	3



Εικόνα 6.7. Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον πολωνομικό πυρήνα (Polynomial Kernel) Μηχανών Διανυσμάτων Υποστήριξης.

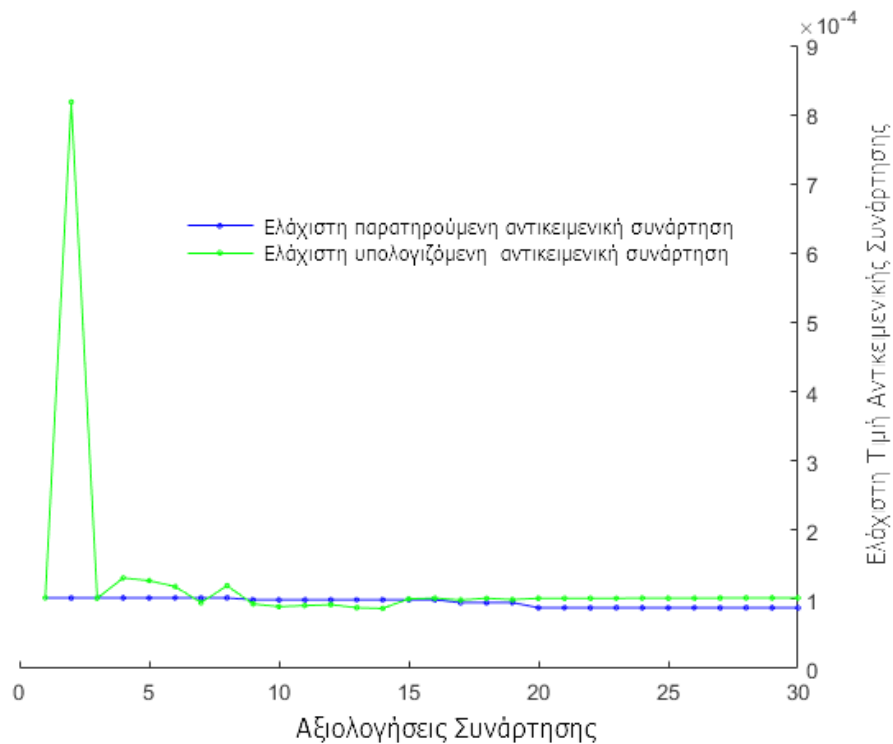
Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον γραμμικό πυρήνα παρουσιάζονται στον παρακάτω πίνακα.

Πίνακας 6.6. Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον γραμμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη	Αντικειμενική	Βέλτιστη	Βέλτιστη	C	ϵ
Αλγορίθμοι	Συνάρτηση	Παρατηρ.	Υπολογιζόμ.		
1	0.00010208	0.00010208	0.00010208	130.540	0.000077

Πίνακας 6.6. (συνέχεια) Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον γραμμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμενική Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζόμ.	C	ε
2	0.01397600	0.00010208	0.00081820	9.172	0.004158
3	0.00033164	0.00010208	0.00010159	0.001	0.001818
4	0.00025552	0.00010208	0.00013088	51.368	0.441130
5	0.00025552	0.00010208	0.00012699	0.012	0.052979
6	0.00021431	0.00010208	0.00011836	3.454	0.000018
7	0.00010206	0.00010206	0.00009524	957.770	0.000015
8	0.00033709	0.00010206	0.00011963	0.001	0.000015
9	0.00009931	0.00009931	0.00009350	952.510	0.000115
10	0.00010014	0.00009931	0.00008940	818.100	0.000027
11	0.00010065	0.00009931	0.00009103	928.670	0.000029
12	0.00009948	0.00009931	0.00009240	952.910	0.000027
13	0.00025552	0.00009931	0.00008790	937.220	0.287020
14	0.00036149	0.00009931	0.00008692	918.740	0.000161
15	0.01250000	0.00009931	0.00010070	948.560	0.000016
16	0.00025552	0.00009931	0.00010182	17.530	1.440600
17	0.00009524	0.00009524	0.00009894	62.560	0.006273
18	0.03314400	0.00009524	0.00010157	536.200	0.000511
19	0.00025552	0.00009524	0.00009956	0.183	1.467800
20	0.00008784	0.00008784	0.00010147	27.270	0.000016
21	0.00025552	0.00008784	0.00010160	2.807	1.453800
22	0.00025552	0.00008784	0.00010163	0.228	1.478400
23	0.00010960	0.00008784	0.00010163	615.670	0.011596
24	0.00025552	0.00008784	0.00010174	0.167	1.478600
25	0.00012234	0.00008784	0.00010172	911.870	0.000286
26	0.00025552	0.00008784	0.00010180	9.263	1.300000
27	0.00025552	0.00008784	0.00010198	0.008	1.374400
28	0.00025552	0.00008784	0.00010205	2.851	1.454600
29	0.00025552	0.00008784	0.00010208	0.129	0.133660
30	0.00025552	0.00008784	0.00010215	839.040	0.072165



Εικόνα 6.8. Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον γραμμικό πυρήνα (Linear Kernel) Μηχανών Διανυσμάτων Υποστήριξης.

Όπως παρατηρούμε το καλύτερο μοντέλο των Μηχανών Διανυσμάτων Υποστήριξης (SVM) που βρέθηκε με τον αλγόριθμο Bayes είναι αυτό του πολυωνυμικού πυρήνα με τα χαρακτηριστικά υπερπαραμέτρων που αναπαρίστανται στον παρακάτω πίνακα.

Πίνακας 6.7. Τα βέλτιστα μοντέλα των Μηχανών Διανυσμάτων Υποστήριξης (SVM) που βρέθηκαν με τον αλγόριθμο Bayes για κάθε πυρήνα σύμφωνα με την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Πυρήνας Μοντέλου	Τιμή Αντικειμενικής Συνάρτησης
Γραμμικός	0.00008784
Πολυωνυμικός	0.00008308
Γκαουσιανός	0.00010326
Ακτινικός	0.00009875

6.5.6 Προτεινόμενο Υβριδικό Μοντέλο

Στην παρούσα διατριβή, προτείνεται ένα υβριδικό μοντέλο Τεχνητής Νοημοσύνης για την πρόβλεψη συνδυάζοντας τα βελτιστοποιημένα μοντέλα νευρωνικών δικτύων και Μηχανών Διανυσμάτων Υποστήριξης. Το μοντέλο που προτείνεται, βασίζεται στην εξαγωγή των βαρών του βέλτιστου νευρωνικού δικτύου εμπρόσθιας διάδοσης, που έχουν βελτιστοποιηθεί με την εφαρμογή του αλγορίθμου Βελτιστοποίησης Σμήνους Σωματιδίων. Στην συνέχεια, ο πίνακας των βαρών πολλαπλασιάζεται με τον ανάστροφο πίνακά του ώστε να δημιουργηθεί ένας πίνακας Gram (Cristianini et al., 2002; Shawe-Taylor et al., 2002; Shawe-Taylor et al., 2005).

Ο τελικός προτεινόμενος πυρήνας Μηχανών Διανυσμάτων Υποστήριξης, αποτελείται από το γινόμενο του πίνακα Gram που δημιουργήθηκε με την προηγούμενη διαδικασία, με τον βελτιστοποιημένο πολωνυμικό πυρήνα που βρέθηκε εφαρμόζοντας τον αλγόριθμο βελτιστοποίησης Bayes. Θα έχουμε:

$$K_{proposed}(x_i, x_j) = W_{NeuralGramMatrix} * K_{PolynomialBest}(x_i, x_j) \quad (6.3)$$

Όπου:

$K_{proposed}(x_i, x_j)$, συμβολίζει τον προτεινόμενο πυρήνα

$W_{NeuralGramMatrix}$, αναπαριστά τον πίνακα Gram που δημιουργήθηκε από το γινόμενο του πίνακα βαρών του νευρωνικού δικτύου εμπρόσθιας διάδοσης που βελτιστοποιήθηκαν με τον αλγόριθμο Βελτιστοποίησης Σμήνους Σωματιδίων, με τον ανάστροφο του.

$K_{PolynomialBest}(x_i, x_j)$, αναπαριστά τον βέλτιστο πολωνυμικός πυρήνα που βρέθηκε με την βελτιστοποίηση Bayes.

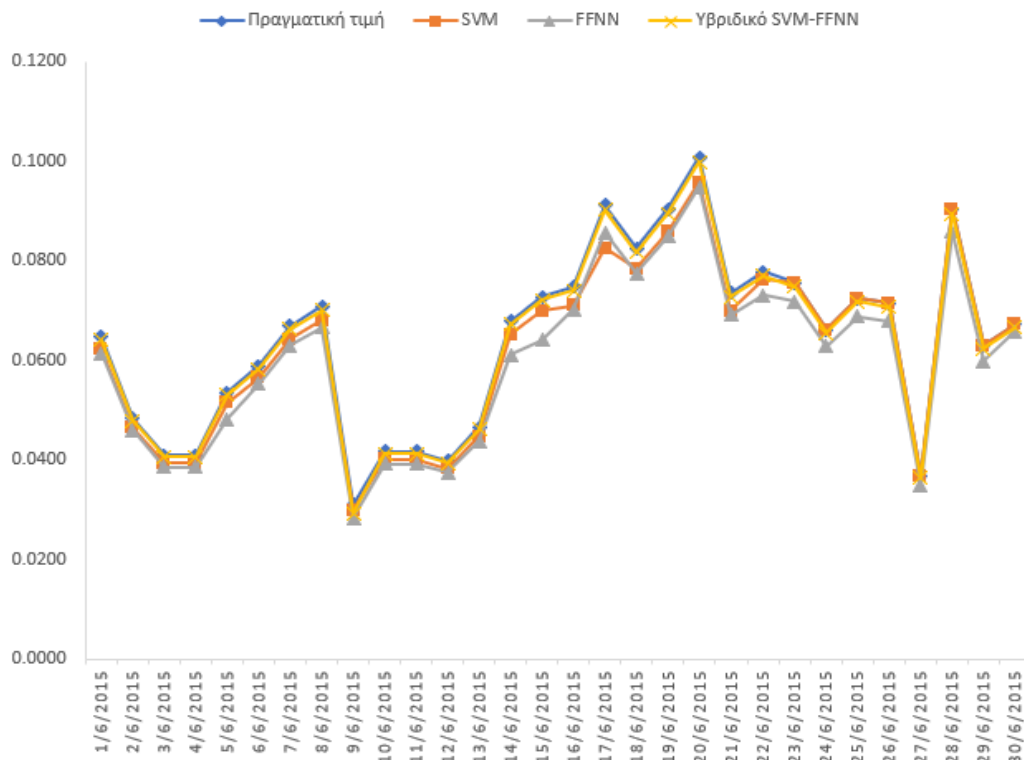
6.5.7 Σύγκριση Βέλτιστων Μοντέλων SVM, FFNN και του Υβριδικού Μοντέλου

Στην ενότητα αυτή, πραγματοποιείται σύγκριση των αποτελεσμάτων των προβλέψεων του όζοντος, των βέλτιστων μοντέλων των Μηχανών Διανυσμάτων Υποστήριξης (SVM) και των νευρωνικών δικτύων εμπρόσθιας διάδοσης (FFNN). Όπως φαίνεται από τα αποτελέσματα, το βέλτιστο μοντέλο είναι αυτό των Μηχανών Διανυσμάτων Υποστήριξης. Στον παρακάτω πίνακα, συγκρίνονται οι δύο προσεγγίσεις που εφαρμόστηκαν στην μελέτη αυτή.

Πίνακας 6.8. Συγκριτικός πίνακας που απεικονίζει, τα βελτιστοποιημένα μοντέλα των Μηχανών Διανυσμάτων Υποστήριξης (SVM) με τον αλγόριθμο Bayes και των νευρωνικών δικτύων εμπρόσθιας διάδοσης (FFNN) με τον αλγόριθμο PSO, ως προς την ακρίβεια πρόβλεψης.

Μοντέλο	Μέθοδος Βελτιστοποίησης	Σφάλμα MSE	Σφάλμα RMSE
<i>SVM-BAYES</i>	Αλγόριθμος Bayes	9.51e-05	9.75e-03
<i>FFNN-PSO</i>	Αλγόριθμος PSO	9.64e-03	9.82e-02
<i>ΥΒΡΙΔΙΚΟ SVM-FFNN</i>	Αλγόριθμοι PSO, BAYES	2.10e-07	4.58e-04

Η παρακάτω εικόνα απεικονίζει τις προβλέψεις του βέλτιστου μοντέλου των Μηχανών Διανυσμάτων Υποστήριξης (κόκκινο χρώμα γραμμής), του υβριδικού μοντέλου (πορτοκαλί χρώμα γραμμής), του βέλτιστου μοντέλου των τεχνητών νευρωνικών δικτύων (γκρι χρώμα γραμμής) και της πραγματικής τιμής (μπλε χρώμα γραμμής) για τα τελευταία 30 χρονικά βήματα (ημέρες) του δείγματος από 01/06/2013 έως 30/06/2015. Όπως φαίνεται από τα αποτελέσματα, το βέλτιστο μοντέλο είναι αυτό των Μηχανών Διανυσμάτων Υποστήριξης.



Εικόνα 6.9. Συγκριτική γραφική παράσταση των προβλέψεων του βέλτιστου μοντέλου νευρωνικού δικτύου, του βέλτιστου μοντέλου SVM, του υβριδικού μοντέλου και της πραγματικής τιμής για τα τελευταία 30 χρονικά βήματα (ημέρες) του δείγματος.

6.6 Συμπεράσματα και Συζήτηση

Αρκετές έρευνες έδειξαν, ότι το όζον ως ατμοσφαιρικός παράγοντας ρύπανσης συνδέεται με διάφορα είδη ασθενειών και αναπνευστικών προβλημάτων (Cuinica et al., 2014; Crouse et al., 2015). Η έρευνα αυτή μπορεί να συμβάλει επίσης, στην πρόληψη αυτών των επιβλαβών επιπτώσεων στην ανθρώπινη υγεία, υιοθετώντας τις κατάλληλες στρατηγικές διαχείρισης και ελαχιστοποιώντας τις συγκεντρώσεις των ρύπων στην ατμόσφαιρα.

Η υιοθέτηση της τεχνητής νοημοσύνης στη διαχείριση της ρύπανσης του περιβάλλοντος στη δημόσια διοίκηση, αποτελεί σήμερα μια πολύ σημαντική ανάγκη με

τον αυξημένο αριθμό περιβαλλοντικών πληροφοριών. Σε αυτή την μελέτη, ερευνήθηκε η πρόβλεψη του όζοντος εφαρμόζοντας Τεχνητά Νευρωνικά Δίκτυα βελτιστοποιημένα με τον αλγόριθμο PSO και Μηχανές Διανυσμάτων Υποστήριξης βελτιστοποιημένες με τον αλγόριθμο Bayes. Επίσης, προτάθηκε ένα υβριδικό μοντέλο συνδυάζοντας τις δύο παραπάνω μεθόδους, για την πρόβλεψη των επιπέδων του όζοντος.

Τα αποτελέσματα του υβριδικού μοντέλου, έδειξαν καλύτερη ακρίβεια προβλέψεων των συγκεντρώσεων του όζοντος σε σχέση με άλλες σχετικές έρευνες (Graf-Jaccottet and Jaunin, 1998). Τα βέλτιστα μοντέλα μηχανικής μάθησης εφαρμόστηκαν λαμβάνοντας υπόψη ως παραμέτρους εισόδου, διάφορους παράγοντες που θα μπορούσαν να επηρεάσουν τα επίπεδα του όζοντος και επίσης δοκιμάζοντας αρκετές τοπολογίες των μοντέλων προκειμένου να βρεθεί η βέλτιστη.

Σύμφωνα με την προτεινόμενη μεθοδολογία που εφαρμόζεται στην παρούσα έρευνα, βελτιστοποιημένες τεχνικές τεχνητής νοημοσύνης χρησιμοποιούνται για την πρόβλεψη ενός παράγοντα σημαντικού στην δημόσια διοίκηση, όπως στο κεφάλαιο αυτό του όζοντος. Η πρόβλεψη αφορά την τιμή του όζοντος του επόμενου χρονικού βήματος, που στην περίπτωση αυτή είναι η επόμενη ημέρα. Δηλαδή, με δεδομένα παραγόντων εισόδου του ‘σήμερα’ που επηρεάζουν την τιμή εξόδου (πρόβλεψη εξαρτημένης μεταβλητής) της συγκέντρωσης του όζοντος του ‘αύριο’. Το στοιχείο αυτό, μπορεί να βοηθήσει αρκετά, αφού οι δημόσιες και περιβαλλοντικές αρχές έχουν το χρόνο μέχρι να επέλθει το επόμενο χρονικό βήμα, για να λάβουν μέτρα στην κατεύθυνση της μείωσης του ρυπογόνου παράγοντα. Για παράδειγμα, η λήψη αποφάσεων για την επόμενη ημέρα, όπως η απόφαση για η μείωση της κυκλοφορίας στο κέντρο της πόλης, ώστε να μειωθούν οι εκπομπές των αέριων ρύπων, όπως του διοξειδίου του αζώτου, που επηρεάζουν τις τιμές του όζοντος (Carslaw et al., 2007). Επίσης, μπορούν να ληφθούν αποφάσεις για την μείωση των ωρών λειτουργίας των εργοστασίων της περιοχής που αποτελούν πηγή παραγωγής οξειδίων του αζώτου (Abdul-Wahab et al., 2000). Το μοντέλο πρόβλεψης μπορεί να χρησιμοποιηθεί για τον έλεγχο σεναρίων, όπως για παράδειγμα η μείωση της τιμής ενός παράγοντα εισόδου, πόση μείωση θα επιφέρει στη τιμή του όζοντος με βάση το μοντέλο.

Συνοψίζοντας, οι προβλέψεις που πραγματοποιήθηκαν με το προτεινόμενο υβριδικό μοντέλο τεχνητής νοημοσύνης, σχετικά με τα επίπεδα του όζοντος, θα είναι χρήσιμες στις Δημόσιες Αρχές στη λήψη αποφάσεων σε θέματα που αφορούν τον αστικό περιβαλλοντικό σχεδιασμό καθώς επίσης και στην λήψη προληπτικών μέτρων για την εφαρμογή αποτελεσματικών πολιτικών περιβαλλοντικής διαχείρισης με σκοπό την ελαχιστοποίηση της ρύπανσης της ατμόσφαιρας και την προστασία της δημόσιας υγείας.

ΚΕΦΑΛΑΙΟ 7: Τεχνητή Νοημοσύνη σε Προβλέψεις Κοινωνικο-οικονομικών παραγόντων

7.1 Συνοπτική Περιγραφή της μελέτης

Η τεχνητή νοημοσύνη έχει εφαρμοστεί σε πολλά επιστημονικά πεδία. τα τελευταία χρόνια με την ανάπτυξη νέων τεχνολογιών νευρωνικών δικτύων και τεχνικών μηχανικής μάθησης. Σε αυτή την έρευνα, τα τεχνητά νευρωνικά δίκτυα εφαρμόζονται για την ανάπτυξη προβλεπτικών νευρωνικών μοντέλων για την πρόβλεψη της ανεργίας. Στην παρούσα μελέτη, θα εφαρμοστεί αρχιτεκτονική του Νευρωνικού Δικτύου Εμπρόσθιας Διάδοσης, δεδομένου ότι θεωρείται από τις πιο κατάλληλες σε προβλέψεις χρονοσειρών (Koskela et al., 1996; Tang and Fishwick, 1993). Το βέλτιστο νευρωνικό μοντέλο πρόβλεψης αξιολογήθηκε με δοκιμές διαφορετικών τοπολογιών του νευρωνικού δικτύου που αφορούν τον αριθμό των νευρώνων, τον αριθμό των κρυφών στρωμάτων και επίσης τη φύση των συναρτήσεων μεταφοράς στα κρυφά στρώματα. Διάφοροι κοινωνικοοικονομικοί παράγοντες ερευνήθηκαν προκειμένου να ληφθούν υπόψη ως παράμετροι εισόδου για να κατασκευαστεί το βέλτιστο μοντέλο πρόβλεψης. Τα αποτελέσματα έδειξαν πολύ καλή ακρίβεια προβλέψεων σχετικά με την τα δεδομένα ανεργίας. Η προτεινόμενη μεθοδολογία μπορεί να βοηθήσει τις αρμόδιες δημόσιες υπηρεσίες να υιοθετήσουν προληπτικά μέτρα για την πρόληψη της περαιτέρω αύξησης της ανεργίας που θα είχε αρνητικές επιπτώσεις στην κοινωνία.

7.2 Εισαγωγή – Σχετική Έρευνα

Η ταχεία ανάπτυξη της τεχνολογίας των πληροφοριών και της επικοινωνιών έχει επηρεάσει πολύπλευρα τις πρακτικές της δημόσιας διοίκησης προς την κατεύθυνση της υιοθέτησης βελτιωμένων στρατηγικών βιώσιμης αστικής ανάπτυξης και διαχείρισης. Η ραγδαία αύξηση των πληροφοριών στη δημόσια διοίκηση, τα τελευταία χρόνια, οδήγησε τους δημόσιους φορείς στην υιοθέτηση νέων τεχνολογιών και πληροφοριακών συστημάτων προκειμένου να αντιμετωπιστούν τα προβλήματα δημόσιας διαχείρισης με πιο συστηματικό τρόπο.

Αρκετές μελέτες έχουν δείξει την αυξημένη ανάγκη μεταρρύθμισης της δημόσιας διοίκησης με την εφαρμογή πρακτικών βασισμένων στην Τεχνολογία των Πληροφοριών και της Επικοινωνιών (ΤΠΕ) προκειμένου να προωθηθεί η ριζική αναμόρφωση του δημόσιου τομέα (Dunleavy et al. 2006; Kouziokas 2016b, 2016c, 2017a). Η ανάπτυξη νέων τεχνικών τεχνητής νοημοσύνης αύξησε σημαντικά την υιοθέτηση συστημάτων τεχνητής νοημοσύνης σε πολλούς επιστημονικούς τομείς.

Τα νευρωνικά δίκτυα έχουν εφαρμοστεί από πολλούς ερευνητές ως προηγμένο εργαλείο πρόβλεψης σε διάφορα θέματα δημόσιας διοίκησης όπως η περιβαλλοντική διοίκηση (Kouziokas et al., 2016), τα μέσα μαζικής μεταφοράς (Kouziokas, 2016a, Kouziokas 2017i, 2018a), το οικονομικό μάνατζμεντ (Kouziokas, 2017b, 2017g, 2018) και τη λήψη αποφάσεων (Cortès et al., 2000, Kouziokas and Perakis 2017, Metaxiotis et al., 2003).

Η εφαρμογή διαφόρων τεχνικών πρόβλεψης, όπως τα τεχνητά νευρωνικά δίκτυα, στην πρόβλεψη των ποσοστών ανεργίας έχει μελετηθεί από διάφορους ερευνητές (Chen 2008, Franses et al., 2004, Karathanasopoulos et al., 2015, Kouziokas 2017f, 2018b; Olmedo 2014).

Ο Chen (2008) μελέτησε την εφαρμογή του μη γραμμικού γκρίζου μοντέλου Bernoulli, προκειμένου να προβλέψει τα ποσοστά ανεργίας σε δέκα χώρες. Τα αποτελέσματα έδειξαν ότι μπορούν να βοηθήσουν τις κυβερνήσεις να αναπτύξουν βιώσιμες στρατηγικές σχετικά με την οικονομία και την εργασία.

Οι Franses et al. (2004), χρησιμοποίησαν το μοντέλο αυτόματης αντιστάθμισης χρονοσειρών λαμβάνοντας υπόψη διάφορους παράγοντες προκειμένου να προβλέψουν τα ποσοστά ανεργίας σε τρεις χώρες. Τα αποτελέσματα έδειξαν ότι η προτεινόμενη τεχνική αποδίδει καλά αποτελέσματα πρόβλεψης.

Ο Olmedo (2014), εφάρμοσε την ανάλυση πλησιέστερων γειτόνων και τα τεχνητά νευρωνικά δίκτυα για να προβλέψει τα ποσοστά ανεργίας. Τα αποτελέσματα έδειξαν πολύ καλή ακρίβεια πρόβλεψης.

Οι Karathanasopoulos et al. (2015), μελέτησαν την πρόβλεψη της ανεργίας στις Ηνωμένες Πολιτείες συνδυάζοντας ακτινικά νευρωνικά δίκτυα και μηχανές διανυσματικής υποστήριξης. Τα αποτελέσματα έδειξαν ότι αυτή η μεθοδολογία παρέχει αποτελέσματα με αυξημένη ακρίβεια.

Στην παρούσα έρευνα, χρησιμοποιήθηκαν μοντέλα τεχνητών νευρωνικών δικτύων για την πρόβλεψη των ποσοστών ανεργίας στο Ηνωμένο Βασίλειο. Λήφθηκαν υπόψη διάφοροι παράγοντες που επηρεάζουν τα ποσοστά ανεργίας, ώστε να χρησιμοποιηθούν ως παράγοντες εισόδου κατά την τροφοδότηση των μοντέλων νευρωνικών δικτύων. Επίσης, ερευνήθηκαν πολλαπλές αρχιτεκτονικές νευρωνικών δικτύων σχετικά με τον τύπο των συναρτήσεων μεταφοράς και τον αριθμό των νευρώνων των κρυφών στρωμάτων, προκειμένου να βρεθεί το βέλτιστο μοντέλο πρόβλεψης. Στα επόμενα κεφάλαια παρουσιάζεται και συζητείται η μεθοδολογία της έρευνας, το θεωρητικό πλαίσιο και τα αποτελέσματα.

7.3 Θεωρητικό πλαίσιο

7.3.1 Τεχνητά Νευρωνικά Δίκτυα

Τα Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ) είναι τεχνητά υπολογιστικά συστήματα που μπορούν να προσομοιάσουν τη δομή του ανθρώπινου εγκεφαλικού συστήματος. Σε ένα τεχνητό νευρωνικό δίκτυο τα δεδομένα εισόδου επεξεργάζονται για να παράγουν τις τιμές εξόδου.

Το πλεονέκτημα των νευρωνικών δικτύων είναι ότι μπορούν να μοντελοποιήσουν και σχέσεις μη γραμμικής φύσης μεταξύ της εισόδου και της εξόδου. Οι πληροφορίες διοχετεύονται μέσω των συνδέσεων του δικτύου προκειμένου να παράγουν μια έξοδο σύμφωνα με την είσοδο (Basheer and Hajmeer 2000). Τα τεχνητά νευρωνικά δίκτυα χρησιμοποιούνται στη μελέτη αυτή για να προβλέψουν τα ποσοστά ανεργίας λαμβάνοντας υπόψη διάφορους κοινωνικοοικονομικούς παράγοντες που επηρεάζουν τα ποσοστά ανεργίας σύμφωνα με τη βιβλιογραφία.

7.3.2 Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης

Σε αυτή τη μελέτη χρησιμοποιήθηκε ένας Πολυεπίπεδος Αισθητήρας Εμπρόσθιας Διάδοσης (Feed-Forward Multilayer Perceptron), δεδομένου ότι πολλοί ερευνητές έχουν δείξει ότι είναι ο καταλληλότερος για προβλήματα πρόβλεψης χρονοσειρών (Hornik 1991). Το χαρακτηριστικό γνώρισμα ενός νευρωνικού δικτύου πολλαπλών στρωμάτων εμπρόσθιας διάδοσης είναι ότι όλοι οι νευρώνες συνδέονται μόνο με κατεύθυνση προς τα εμπρός (Svozil et al 1997). Μία τυπική δομή ενός νευρωνικού δικτύου πολλαπλών στρωμάτων εμπρόσθιας διάδοσης αποτελείται από ένα στρώμα εισόδου, ένα στρώμα εξόδου και ένα ή περισσότερα κρυφά στρώματα. Κάθε στρώμα αποτελείται από έναν αριθμό νευρώνων (Koskela et al., 1996).

7.3.3 Αλγόριθμος εκπαίδευσης Levenberg Marquardt

Ο αλγόριθμος Levenberg-Marquardt χρησιμοποιήθηκε στην παρούσα έρευνα ως αλγόριθμος εκπαίδευσης των νευρωνικών δικτύων που αναπτύχθηκαν, δεδομένου ότι συνήθως χρησιμοποιείται για την επίλυση μη γραμμικών προβλημάτων και προβλημάτων χρονοσειρών (Lourakis, 2005). Επιπλέον, ένας άλλος λόγος για την επιλογή του αλγορίθμου Levenberg – Marquardt ως αλγορίθμου μάθησης για την εκπαίδευση του νευρωνικού δικτύου πολλαπλών στρωμάτων, είναι ότι θεωρείται ένας από τους ταχύτερους αλγόριθμους εκπαίδευσης σε σύγκριση με άλλους.

Ο αλγόριθμος Levenberg – Marquardt είναι μια επαναληπτική μέθοδος βελτιστοποίησης και συνδυάζει τα πλεονεκτήματα ελαχιστοποίησης δύο αλγορίθμων, του αλγορίθμου Gauss-Newton και επίσης του αλγορίθμου απότομης καθόδου που χρησιμοποιείται για την επίλυση μη γραμμικών προβλημάτων με ελάχιστα τετράγωνα (Marquardt 1963).

7.3.4 Ανεργία

Σύμφωνα με τη Διεθνή Οργάνωση Εργασίας, το ποσοστό ανεργίας είναι το ποσοστό του πληθυσμού που είναι οικονομικά ενεργός αλλά είναι άνεργοι (εκείνοι που εργάζονται, αλλά και ο πληθυσμός που αναζητά και είναι διαθέσιμος για εργασία). Η επόμενη εικόνα απεικονίζει τα ποσοστά ανεργίας στην Ευρώπη από το 2004 έως το 2015

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
EU-28	9.3	9.0	8.2	7.2	7.0	9.0	9.6	9.7	10.5	10.9	10.2	9.4
Euro area	9.3	9.1	8.4	7.5	7.6	9.6	10.2	10.2	11.4	12.0	11.6	10.9
Belgium	8.4	8.5	8.3	7.5	7.0	7.9	8.3	7.2	7.6	8.4	8.5	8.5
Bulgaria	12.1	10.1	9.0	6.9	5.6	6.8	10.3	11.3	12.3	13.0	11.4	9.2
Czech Republic	8.3	7.9	7.1	5.3	4.4	6.7	7.3	6.7	7.0	7.0	6.1	5.1
Denmark	5.5	4.8	3.9	3.8	3.4	6.0	7.5	7.6	7.5	7.0	6.6	6.2
Germany	10.4	11.2	10.1	8.5	7.4	7.6	7.0	5.8	5.4	5.2	5.0	4.6
Estonia	10.1	8.0	5.9	4.6	5.5	13.5	16.7	12.3	10.0	8.6	7.4	6.2
Ireland	4.5	4.4	4.5	4.7	6.4	12.0	13.9	14.7	14.7	13.1	11.3	9.4
Greece	10.6	10.0	9.0	8.4	7.8	9.6	12.7	17.9	24.5	27.5	26.5	24.9
Spain	11.0	9.2	8.5	8.2	11.3	17.9	19.9	21.4	24.8	26.1	24.5	22.1
France	8.9	8.9	8.8	8.0	7.4	9.1	9.3	9.2	9.8	10.3	10.3	10.4
Croatia	13.9	13.0	11.6	9.9	8.6	9.2	11.7	13.7	16.0	17.3	17.3	16.3
Italy	8.0	7.7	6.8	6.1	6.7	7.7	8.4	8.4	10.7	12.1	12.7	11.9
Cyprus	4.6	5.3	4.6	3.9	3.7	5.4	6.3	7.9	11.9	15.9	16.1	15.0
Latvia	11.7	10.0	7.0	6.1	7.7	17.5	19.5	16.2	15.0	11.9	10.8	9.9
Lithuania	10.9	8.3	5.8	4.3	5.8	13.8	17.8	15.4	13.4	11.8	10.7	9.1
Luxembourg	5.0	4.6	4.6	4.2	4.9	5.1	4.6	4.8	5.1	5.9	6.0	6.4
Hungary	6.1	7.2	7.5	7.4	7.8	10.0	11.2	11.0	11.0	10.2	7.7	6.8
Malta	7.2	6.9	6.8	6.5	6.0	6.9	6.9	6.4	6.3	6.4	5.8	5.4
Netherlands	5.7	5.9	5.0	4.2	3.7	4.4	5.0	5.0	5.8	7.3	7.4	6.9
Austria	5.5	5.6	5.3	4.9	4.1	5.3	4.8	4.6	4.9	5.4	5.6	5.7
Poland	19.1	17.9	13.9	9.6	7.1	8.1	9.7	9.7	10.1	10.3	9.0	7.5
Portugal	7.8	8.8	8.9	9.1	8.8	10.7	12.0	12.9	15.8	16.4	14.1	12.6
Romania	8.0	7.1	7.2	6.4	5.6	6.5	7.0	7.2	6.8	7.1	6.8	6.8
Slovenia	6.3	6.5	6.0	4.9	4.4	5.9	7.3	8.2	8.9	10.1	9.7	9.0
Slovakia	18.4	16.4	13.5	11.2	9.6	12.1	14.5	13.7	14.0	14.2	13.2	11.5
Finland	8.8	8.4	7.7	6.9	6.4	8.2	8.4	7.8	7.7	8.2	8.7	9.4
Sweden	7.4	7.7	7.1	6.1	6.2	8.3	8.6	7.8	8.0	8.0	7.9	7.4
United Kingdom	4.7	4.8	5.4	5.3	5.6	7.6	7.8	8.1	7.9	7.6	6.1	5.3

Εικόνα 7.1 Ανεργία στην χώρες της Ευρωπαϊκής Ένωσης από το 2004 έως το 2015. Πηγή: Eurostat⁴

⁴ <http://ec.europa.eu/eurostat/>

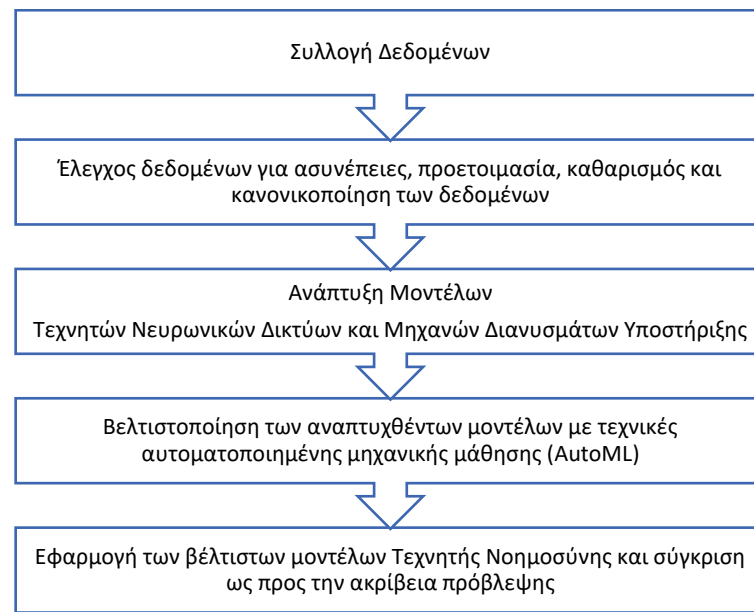
7.4 Μεθοδολογία Έρευνας

Η μεθοδολογία της έρευνας χωρίζεται σε τέσσερα στάδια: συλλογή δεδομένων, καθαρισμός και προετοιμασία δεδομένων, ανάπτυξη μοντέλων τεχνητής νοημοσύνης σύγκριση και τέλος εφαρμογή του βέλτιστου μοντέλου για την πρόβλεψη των ποσοστών ανεργίας στο Ηνωμένο Βασίλειο. Στο πρώτο στάδιο συλλέχθηκαν δεδομένα σχετικά με τα ποσοστά ανεργίας καθώς και δεδομένα σχετικά με τους κοινωνικοοικονομικούς παράγοντες που επηρεάζουν τα ποσοστά ανεργίας, όπως η αύξηση του ΑΕΠ, οι εξαγωγές αγαθών και υπηρεσιών.

Στο δεύτερο στάδιο, τα δεδομένα ελέγχθηκαν για πιθανά κενά και ασυνέπειες και κανονικοποιήθηκαν για να τροφοδοτήσουν τα μοντέλα πρόβλεψης τεχνητών νευρωνικών δικτύων. Στην τρίτη φάση, αναπτύχθηκαν μοντέλα πρόβλεψης βασισμένα στα νευρωνικά δίκτυα εμπρόσθιας διάδοσης και στις Μηχανές Διανυσμάτων Υποστήριξης.

Στο τέταρτο στάδιο, πραγματοποιείται η βελτιστοποίηση των μοντέλων με τεχνικές αυτοματοποιημένης μηχανικής μάθησης (AutoML). Για την βελτιστοποίηση του νευρωνικού δικτύου χρησιμοποιήθηκαν ο αλγόριθμος αναζήτησης πλέγματος και ο αλγόριθμος βελτιστοποίησης Σμήνους Σωματιδίων. Για την βελτιστοποίηση των υπερπαραμέτρων των Μηχανών Διανυσμάτων Υποστήριξης χρησιμοποιήθηκε ο αλγόριθμος βελτιστοποίησης Bayes.

Στην τελευταία φάση, το βέλτιστο μοντέλα πρόβλεψης με κάθε μέθοδο τεχνητής νοημοσύνης χρησιμοποιήθηκαν για να προβλέψουν τα ποσοστά ανεργίας στο Ηνωμένο Βασίλειο, τα οποία συγκρίθηκαν ως προς την ακρίβεια. Μια επισκόπηση της μεθοδολογίας της έρευνας απεικονίζεται στην εικόνα 7.3.



Εικόνα 7.2 Επισκόπηση σταδίων εφαρμοσθείσας μεθοδολογίας για την ανάπτυξη και εφαρμογή των μοντέλων τεχνητής νοημοσύνης.

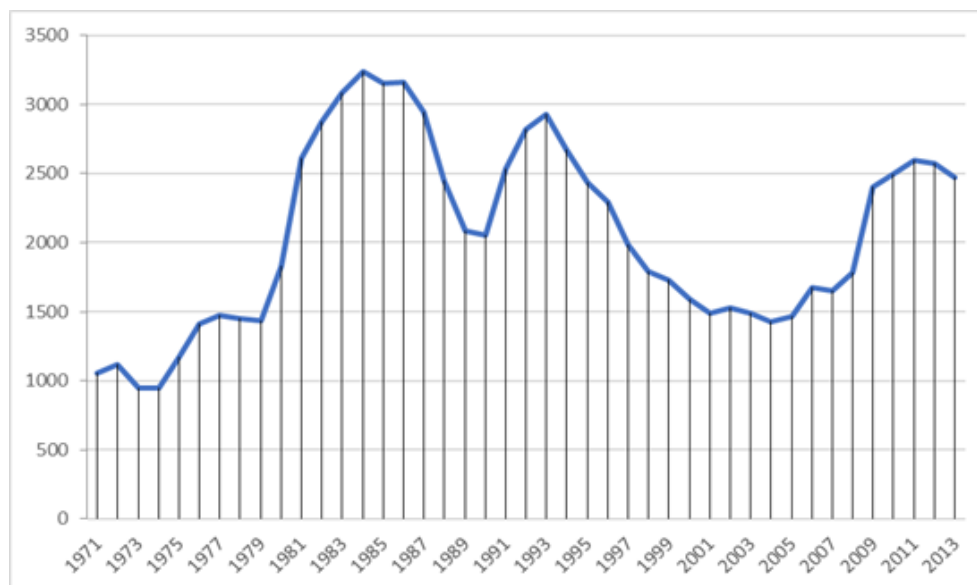
7.5 Αποτελέσματα

7.5.1 Συλλογή Δεδομένων – Περιοχή Μελέτης

Τα δεδομένα που αφορούσαν κοινωνικοοικονομικούς παράγοντες όπως η ετήσια αύξηση του ΑΕΠ και οι εξαγωγές αγαθών και υπηρεσιών (ετήσια αύξηση επί τοις εκατό) συλλέχθηκαν από τον επίσημο ιστότοπο του Γραφείου Εθνικών Στατιστικών (ONS) του Ηνωμένου Βασιλείου. Τα στοιχεία για τα ποσοστά ανεργίας συλλέχθηκαν επίσης από τον επίσημο ιστότοπο του Γραφείου Εθνικών Στατιστικών (ONS) του Ηνωμένου Βασιλείου (U.K. Office for National Statistics 2018).

Επιλέχθηκε το Ηνωμένο Βασίλειο, διότι μόνο εκεί βρέθηκαν δεδομένα και για τους τρεις παράγοντες εισόδου: την τιμή της αύξησης του ΑΕΠ, την τιμή της αύξησης των εξαγωγών αγαθών και υπηρεσιών και του ποσοστού της ανεργίας για τόσο μεγάλο βάθος χρόνου, σε αντίθεση με την Ελλάδα και άλλες χώρες, που δεν υπάρχει διαθεσιμότητα ικανού αριθμού δεδομένων και των τριών παραγόντων εισόδου για τόσο μεγάλο χρονικό διάστημα, ώστε να εκπαιδευτούν επαρκώς τα μοντέλα τεχνητής νοημοσύνης.

Τα δεδομένα που συλλέχθηκαν καλύπτουν τη χρονική περίοδο από το έτος 1971 έως το έτος 2013. Τα δεδομένα προεπεξεργάστηκαν και προετοιμάστηκαν για να χρησιμοποιηθούν για τη τροφοδότηση των μοντέλων τεχνητών νευρωνικών δικτύων. Επίσης, τα δεδομένα ελέγχθηκαν για κενά, ασυμφωνίες και διπλότυπα. Η επόμενη εικόνα, δείχνει την μεταβολή των τιμών της ανεργίας (σε χιλιάδες) στο Ηνωμένο Βασίλειο από το 1971 έως το 2013.



Εικόνα 7.3 Το σχήμα αναπαριστά την μεταβολή των τιμών της ανεργίας (σε χιλιάδες) στο Ηνωμένο Βασίλειο από το 1971 έως το 2013.

7.5.2 Μοντέλα Πρόβλεψης Τεχνητών Νευρωνικών Δικτύων

Καταρχάς, διερευνήθηκε η σχετική βιβλιογραφία προκειμένου να εντοπιστούν οι παράμετροι που επηρεάζουν τα ποσοστά ανεργίας που θα χρησιμοποιηθούν ως μεταβλητές εισόδου στα μοντέλα νευρωνικών δικτύων. Σύμφωνα με τη βιβλιογραφία, πολλοί κοινωνικοοικονομικοί παράγοντες επηρεάζουν τα επίπεδα των ποσοστών ανεργίας. Από τους σημαντικότερους παράγοντες είναι: η αύξηση του ΑΕΠ και η αύξηση των εξαγωγών αγαθών και υπηρεσιών (Bayar 2014; Ozughalu and Ogwumike 2013). Τα

μοντέλα νευρωνικών δικτύων αναπτύχθηκαν χρησιμοποιώντας τους προαναφερθέντες κοινωνικοοικονομικούς παράγοντες ως μεταβλητές εισόδου που επηρεάζουν τα επίπεδα των ποσοστών ανεργίας καθώς επίσης και οι ιστορικές τιμές των ποσοστών ανεργίας.

Συγκεκριμένα, το διάνυμα εισόδου θα αποτελείται από την τιμή της αύξησης του ΑΕΠ του προηγούμενου έτους που συμβολίζεται ως $GDPI_{n-1}$, την τιμή της αύξησης των εξαγωγών αγαθών και υπηρεσιών του προηγούμενου έτους που συμβολίζεται ως E_{n-1} και του ποσοστού της ανεργίας του προηγούμενου έτους που συμβολίζεται ως U_{n-1} :

$$X_i = (GDPI_{n-1}, E_{n-1}, U_{n-1}) \quad (7.1)$$

Οι τιμές εξόδου Y_i αφορούν τις τιμές που προβλέπει το κάθε μοντέλο σχετικά με την ανεργία του επόμενου έτους:

$$Y_i = (U_n) \quad (7.2)$$

Τα δεδομένα χωρίστηκαν σε τρία διαφορετικά μέρη. Το 60% των δεδομένων χρησιμοποιήθηκε ως σύνολο εκπαίδευσης, το 20% για το σύνολο επικύρωσης και το 20% για το σετ δοκιμής. Το σύνολο δεδομένων εκπαίδευσης χρησιμοποιήθηκε για την εκπαίδευση των μοντέλων νευρωνικών δικτύων χρησιμοποιώντας τα συλλεχθέντα δεδομένα.

Το σετ επικύρωσης χρησιμοποιήθηκε για την αξιολόγηση των μοντέλων νευρωνικών δικτύων. Ο αλγόριθμος Levenberg – Marquardt χρησιμοποιήθηκε ως αλγόριθμος μάθησης, καθώς θεωρείται ένας από τους γρηγορότερους αλγορίθμους εκπαίδευσης σε σύγκριση με άλλους αλγορίθμους μάθησης και επιπλέον παρουσίασε καλύτερα αποτελέσματα σε σύγκριση με τους άλλους αλγορίθμους που ερευνήθηκαν.

7.5.3 Βελτιστοποίηση Τοπολογίας του Νευρωνικού Δικτύου

Προκειμένου να βρεθεί το βέλτιστο μοντέλο νευρωνικού δικτύου, εκτελέστηκαν πολλαπλές δοκιμές. Αρχικά, η απόδοση κάθε αναπτυχθέντος μοντέλου νευρωνικού

δικτύου ερευνήθηκε χρησιμοποιώντας διαφορετικές αρχιτεκτονικές νευρωνικών δικτύων. Οι παράμετροι δοκιμής των δομών του νευρωνικού δικτύου ήταν ο αριθμός των κρυφών στρωμάτων και οι συναρτήσεις μεταφοράς στα κρυμμένα στρώματα. Οι αρχιτεκτονικές με ένα και δύο κρυφά στρώματα ερευνήθηκαν και η βέλτιστη τοπολογία που παρήγαγε τα καλύτερα αποτελέσματα πρόβλεψης ήταν αυτή με δύο κρυφά επίπεδα.

Οι συναρτήσεις μεταφοράς που μελετήθηκαν είναι: Η Εφαπτομενική – σιγμοειδής συνάρτηση μεταφοράς (Tan-Sigmoid Transfer Function – TSTF), η λογιστική – σιγμοειδής συνάρτηση μεταφοράς (Log-Sigmoid Transfer Function – LSTF), η γραμμική συνάρτηση μεταφοράς (Linear Transfer Function – LTF), η Elliot σιγμοειδής συνάρτηση μεταφοράς (Elliot Sigmoid Transfer Function – ESTF) και η θετική γραμμική συνάρτηση μεταφοράς (Positive Linear Transfer Function – PLTF).

Επίσης, δοκιμάστηκαν οι κυριότεροι αλγόριθμοι μηχανικής μάθησης για την εκπαίδευση κάθε νευρωνικού μοντέλου και συγκεκριμένα οι: ο αλγόριθμος Κλιμακούμενων Συζυγών Κλίσεων (Scaled Conjugate Gradient – SCG), BFGS Quasi-Newton (BFGS-QN), ο αλγόριθμος Ελαστικής Οπισθοδιάδοσης (Resilient Backpropagation – RB) και ο αλγόριθμος Levenberg Marquardt (LM). Επίσης, εξετάστηκαν τοπολογίες νευρωνικών δικτύων με αριθμό κρυφών επιπέδων (hidden layers) από ένα έως δύο, και αριθμό νευρώνων κάθε κρυφού επιπέδου από έναν έως πενήντα.

Η Ρίζα του Μέσου Τετραγωνικού Σφάλματος (Root Mean Square Error - RMSE) χρησιμοποιήθηκε για την εκτίμηση του σφάλματος πρόβλεψης για κάθε νευρωνικό μοντέλο που αναπτύχθηκε.

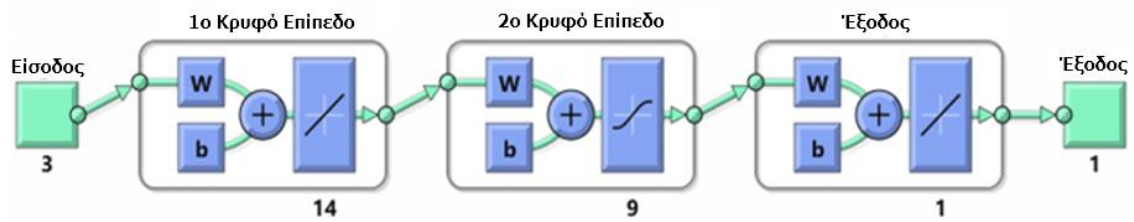
Στον παρακάτω πίνακα, απεικονίζονται τα βέλτιστα νευρωνικά μοντέλα που βρέθηκαν για κάθε αλγόριθμο που δοκιμάστηκε χρησιμοποιώντας διαφορετικές τοπολογίες και συναρτήσεις ενεργοποίησης. Όπως φαίνεται από τον πίνακα το βέλτιστο μοντέλο με το μικρότερο σφάλμα επικύρωσης είναι το ANN4.

Πίνακας 7.1. Τα βέλτιστα νευρωνικά μοντέλα που βρέθηκαν για κάθε αλγόριθμο που δοκιμάστηκε χρησιμοποιώντας διαφορετικές τοπολογίες και συναρτήσεις ενεργοποίησης.

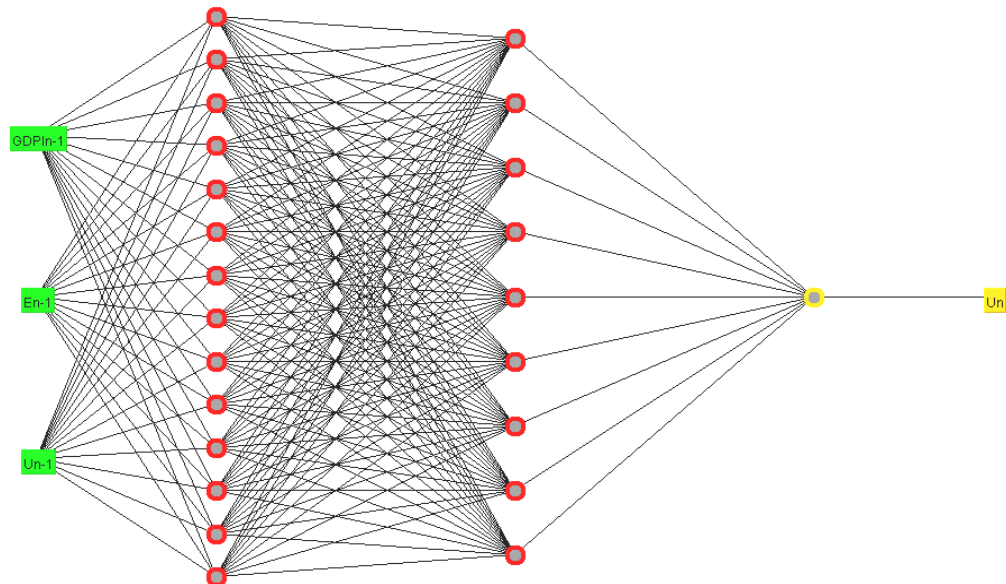
Μοντέλο	Αλγόριθμος	Αριθμός Κρυφών Επιπέδων	Αριθμός Νευρώνων στο 1 ^ο Κρυφό Επίπεδο	Αριθμός Νευρώνων στο 2 ^ο Κρυφό Επίπεδο	Συναρτήσεις Μεταφοράς (<i>Transfer</i> <i>Functions</i>)	RMSE Σφάλμα Εκπαίδευσης (<i>Training</i> <i>Error</i>)	RMSE Σφάλμα Επικύρωσης (<i>Validation</i> <i>Error</i>)
ANN1	SCG	2	12	19	TSAF- LTF	21.0154	25.2541
ANN2	BFGS- QN	2	14	26	LTF- LTF	325.5156	321.0214
ANN3	RB	2	22	17	LTF- LTF	125.1287	132.5510
ANN4	LM	2	14	9	TSAF- LTF	0.4812	0.4957

Στο βέλτιστο μοντέλο νευρωνικού δικτύου, ο αλγόριθμος εκπαίδευσης ήταν ο αλγόριθμος Levenberg Marquardt. Η βέλτιστη αρχιτεκτονική του νευρωνικού δικτύου, βρέθηκε να είναι αυτή με δύο κρυφά επίπεδα, με 14 νευρώνες και την γραμμική συνάρτηση μεταφοράς (LTF) ως συνάρτηση μεταφοράς στο πρώτο κρυφό επίπεδο και 9 νευρώνες και την εφαπτομενική – σιγμοειδής συνάρτηση μεταφοράς (TSTF) στο δεύτερο κρυφό επίπεδο ως συνάρτηση μεταφοράς.

Το μέσο τετραγωνικό σφάλμα (MSE) του βέλτιστου μοντέλου βρέθηκε να είναι 0.2457 στην εποχή 8 και η ρίζα του μέσου τετραγωνικού σφάλματος (RMSE) ήταν 0.4957. Τα αποτελέσματα έδειξαν πολύ καλή ακρίβεια προβλέψεων. Στην εικόνα 7.5, απεικονίζεται η τοπολογία του νευρωνικού δικτύου του βέλτιστου μοντέλου, ενώ στην επόμενη εικόνα, αναπαρίστανται σχηματικά οι συνδέσεις των νευρώνων και των κρυφών επιπέδων του μοντέλου.

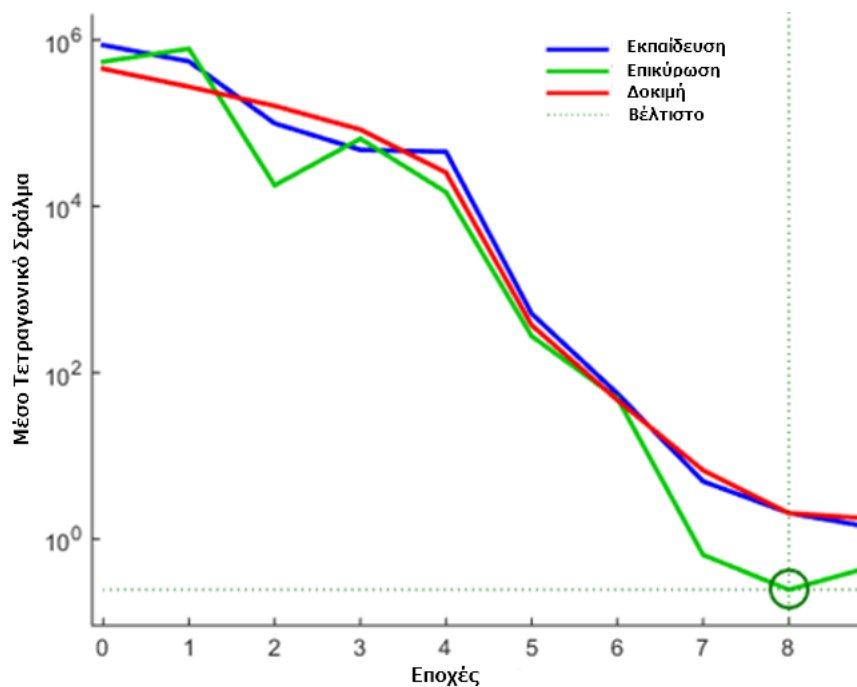


Εικόνα 7.4 Το σχήμα αναπαριστά την τοπολογία του βέλτιστου μοντέλου τεχνητού νευρωνικού δικτύου για την πρόβλεψη των τιμών της ανεργίας με την χρήση δύο κρυφών επιπέδων (hidden layers). Το πρώτο κρυφό επίπεδο αποτελείται από 14 νευρώνες και χρησιμοποιεί την γραμμική συνάρτηση μεταφοράς, ενώ το δεύτερο επίπεδο αποτελείται από 9 νευρώνες και χρησιμοποιεί την εφαπτομενική – σιγμοειδής συνάρτηση μεταφοράς.



Εικόνα 7.5 Η εικόνα αναπαριστά σχηματικά την τοπολογία και τις συνδέσεις των νευρώνων και του αριθμού των νευρώνων των κρυφών επιπέδων του βέλτιστου μοντέλου τεχνητού νευρωνικού δικτύου για την πρόβλεψη των τιμών της ανεργίας.

Η επόμενη εικόνα, αναπαριστά το διάγραμμα απόδοσης των συνόλων εκπαίδευσης, επικύρωσης και δοκιμής, χρησιμοποιώντας το βέλτιστο μοντέλο νευρωνικού δικτύου σύμφωνα με το ελάχιστο μέσο σφάλμα (MSE).



Εικόνα 7.6 Η γραφική παράσταση απόδοσης του βέλτιστου μοντέλου νευρωνικού δικτύου, σύμφωνα με το Μέσο Τετραγωνικό Σφάλμα (Mean Squared Error).

7.5.4 Βελτιστοποίηση Νευρωνικού Δικτύου με Σμήνος Σωματιδίων

Για την βελτιστοποίηση των βαρών του νευρωνικού δικτύου εφαρμόστηκε η μέθοδος Βελτιστοποίησης Σμήνους Σωματιδίων (Particle Swarm Optimization, PSO). Ο αλγόριθμος βελτιστοποίησης σμήνους σωματιδίων (PSO) περιγράφηκε πιο αναλυτικά στην ενότητα 4.4.4. Για την βελτιστοποίηση των παραμέτρων του Σμήνους Σωματιδίων, εφαρμόστηκε η μέθοδος αναζήτησης πλέγματος (Grid Search), που περιγράφηκε παραπάνω, για τις παραμέτρους που αφορούν: τον πληθυσμό των σωματιδίων του σμήνους, τους συντελεστές μάθησης c_1 , c_2 , τον συντελεστή βάρους αδράνειας (inertia weight coefficient). Οι πληθυσμοί των σωματιδίων σμήνους που εξετάστηκαν ήταν 20, 30, 40, 50, που είναι οι πιο συνηθισμένες τιμές σύμφωνα με τη βιβλιογραφία (Fourie and Groenwold, 2002). Οι συντελεστές μάθησης c_1 , c_2 , ορίστηκαν στην τιμή 2.0, όπως προτείνεται από τους Kennedy and Eberhart (1995). Οι τιμές του συντελεστή βάρους αδράνειας (inertia weight coefficient) εξετάστηκαν στο διάστημα [0.8, 1.4] με βήμα 0.1, όπως προτείνεται από τους Shi and Eberhart, (1998). Οι παράμετροι που παρήγαγαν τα

βέλτιστα αποτελέσματα ήταν: μέγεθος πληθυσμού: $P = 50$, βάρος αδράνειας $w = 1.10$, συντελεστής προσωπικής μάθησης $c_1 = 2.0$, παγκόσμιος συντελεστής εκμάθησης $c_2 = 2.0$. Ο παρακάτω πίνακας, παρουσιάζει τα πέντε καλύτερα μοντέλα του αλγορίθμου βελτιστοποίησης PSO, που βρέθηκαν χρησιμοποιώντας διαφορετικές τοπολογίες παραμέτρων. Το βέλτιστο μοντέλο της βελτιστοποίησης Σμήνους Σωματιδίων ήταν το μοντέλο 5. Η Ρίζα του Μέσου Τετραγωνικού Σφάλματος (RMSE) του βελτιστοποιημένου μοντέλου που δοκιμάστηκε στο σύνολο δοκιμής βρέθηκε να είναι 0.1659.

Πίνακας 7.2. Τα πέντε καλύτερα νευρωνικά μοντέλα του αλγορίθμου βελτιστοποίησης PSO που βρέθηκαν χρησιμοποιώντας διαφορετικές τοπολογίες παραμέτρων για την βελτιστοποίηση των βαρών του νευρωνικού δικτύου.

Μοντέλο	Πληθυσμός Σμήνους P	c_1	c_2	w	Βέλτιστο Κόστος (Best Cost)
<i>Μοντέλο 1</i>	20	2.0	2.0	1.10	0.3402
<i>Μοντέλο 2</i>	30	2.0	2.0	1.30	0.4565
<i>Μοντέλο 3</i>	30	2.0	2.0	0.80	0.3854
<i>Μοντέλο 4</i>	40	2.0	2.0	0.90	0.4931
<i>Μοντέλο 5</i>	50	2.0	2.0	1.10	0.1659

7.5.5 Βελτιστοποιημένα Μοντέλα Μηχανών Διανυσμάτων Υποστήριξης

Τα Μοντέλα Μηχανών Διανυσμάτων Υποστήριξης βελτιστοποιήθηκαν εφαρμόζοντας τον αλγόριθμο βελτιστοποίησης Bayes, όπως περιγράφηκε στην ενότητα 4.4.7, και υπολογίστηκαν οι υπερπαραμέτροι για την εύρεση των βέλτιστων μοντέλων. Χρησιμοποιήθηκαν οι κάτωθι βασικοί πυρήνες, οι οποίοι θεωρούνται ως οι πιο συνήθεις στην υλοποίηση των Μηχανών Διανυσμάτων Υποστήριξης στην παλινδρόμηση: γραμμικός, πολυωνυμικός, ακτινικός και γκαουσιανός χρησιμοποιώντας το λογισμικό MATLAB. Για την αξιολόγηση των μοντέλων χρησιμοποιήθηκε η ελάχιστη τιμή της αντικειμενικής συνάρτησης. Με τον αλγόριθμο βελτιστοποίησης Bayes βελτιστοποιήθηκαν οι κάτωθι υπερ-παραμέτροι των μοντέλων Μηχανών Διανυσμάτων Υποστήριξης: Η τιμή κόστους (C), η παράμετρος epsilon (ϵ), η παράμετρος γάμμα (γ)

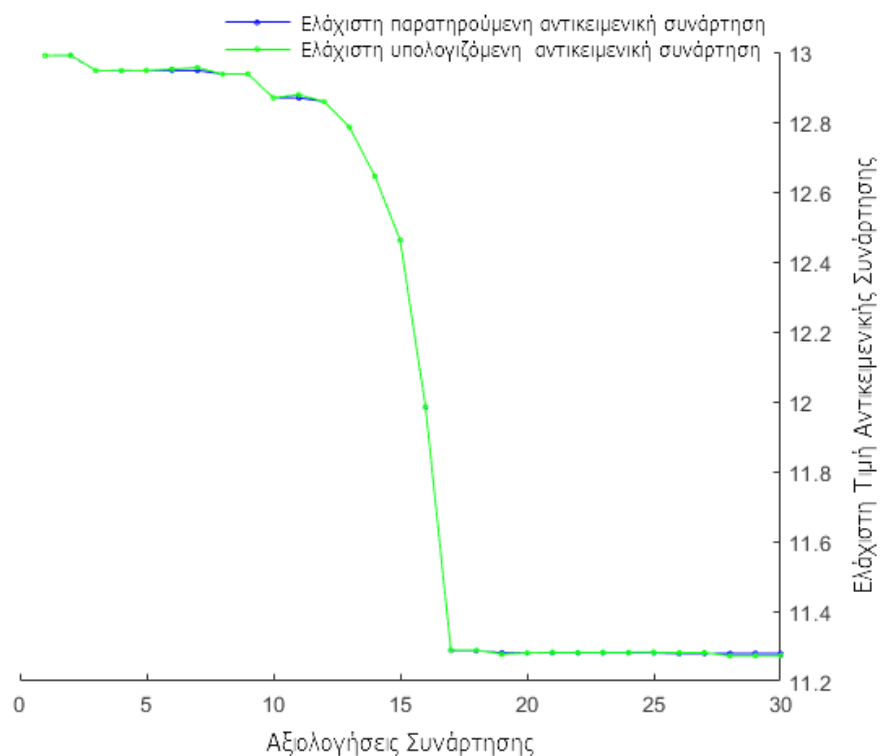
των πυρήνων εκτός του γραμμικού και ο βαθμός του πολυνύμου για τον πολωνυμικό πυρήνα. Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον Γκαουσιανό πυρήνα παρουσιάζονται στον παρακάτω πίνακα.

Πίνακας 7.3. Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον Γκαουσιανό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμενική ή Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζ.	C	γ	ϵ
1	12.990	12.990	12.990	732.840	0.002	2.4318
2	12.994	12.990	12.991	0.002	0.035	951.3900
3	12.947	12.947	12.947	137.020	13.231	549.3200
4	12.987	12.947	12.948	672.100	0.002	10442.0000
5	13.014	12.947	12.947	0.075	75.072	140.4900
6	13.015	12.947	12.951	57.914	1.015	7.1810
7	12.974	12.947	12.955	149.480	3.689	554.7500
8	12.937	12.937	12.937	134.450	20.639	586.7200
9	12.987	12.937	12.937	176.300	61.303	3860.2000
10	12.869	12.869	12.869	128.210	54.912	626.0300
11	13.014	12.869	12.878	126.100	0.002	70.2960
12	12.858	12.858	12.858	113.000	71.554	650.4800
13	12.785	12.785	12.785	189.140	96.125	673.8100
14	12.645	12.645	12.645	221.210	187.220	667.3600
15	12.461	12.461	12.462	300.100	277.430	655.7100
16	11.983	11.983	11.984	349.810	947.840	500.8300
17	11.287	11.287	11.289	519.440	985.940	0.9275
18	13.018	11.287	11.288	0.001	962.970	0.8878
19	11.281	11.281	11.277	932.350	987.150	0.9900
20	11.370	11.281	11.280	852.450	770.090	0.9119
21	12.060	11.281	11.282	41.231	989.920	1.4465
22	13.018	11.281	11.282	0.225	0.001	1.6622
23	11.310	11.281	11.282	706.130	919.860	1.1608
24	12.987	11.281	11.282	0.001	12.304	68378.0000

Πίνακας 7.3. (συνέχεια) Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον Γκαουσιανό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμενική Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζόμ.	C	γ	ϵ
25	12.987	11.281	11.282	0.278	0.133	71830.0000
26	11.279	11.279	11.282	968.300	999.230	2.9175
27	12.987	11.279	11.282	0.001	0.001	70312.0000
28	11.281	11.279	11.272	807.230	976.920	7.2176
29	11.293	11.279	11.273	654.340	954.150	5.1325
30	12.987	11.279	11.273	1.160	911.740	81042.0000



Εικόνα 7.7. Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον Γκαουσιανό πυρήνα (Gaussian Kernel) Μηχανών Διανυσμάτων Υποστήριξης.

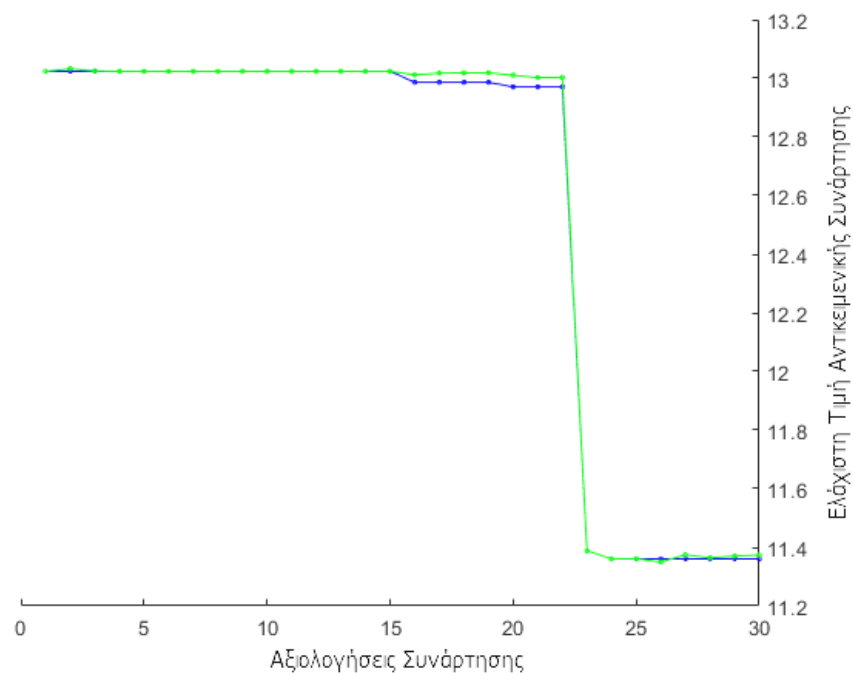
Επίσης, βελτιστοποιήθηκαν οι παράμετροι του ακτινικού πυρήνα, με τον αλγόριθμο Bayes, χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον ακτινικό πυρήνα παρουσιάζονται στον παρακάτω πίνακα.

Πίνακας 7.4. Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον ακτινικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμενική Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζ.	C	γ	ε
1	13.024	13.024	13.024	5.149	3.783	30788.0000
2	13.200	13.024	13.033	0.179	0.185	3.5250
3	13.114	13.024	13.025	344.950	0.008	28.2760
4	13.024	13.024	13.024	0.002	0.013	2980.8000
5	13.024	13.024	13.024	55.594	330.200	9433.4000
6	13.024	13.024	13.024	9.078	103.140	81689.0000
7	13.024	13.024	13.024	77.036	0.033	5465.5000
8	13.024	13.024	13.024	0.004	0.019	5254.1000
9	13.024	13.024	13.024	0.502	401.090	81818.0000
10	13.024	13.024	13.024	33.034	0.041	43842.0000
11	13.024	13.024	13.024	0.271	0.008	4910.1000
12	13.024	13.024	13.024	11.646	0.001	16635.0000
13	13.024	13.024	13.024	0.004	352.810	47386.0000
14	13.025	13.024	13.024	0.037	474.560	400.3500
15	13.024	13.024	13.024	3.012	1.291	898.6000
16	12.986	12.986	13.012	5.416	4.425	651.8700
17	13.036	12.986	13.017	21.413	0.001	855.0500
18	13.024	12.986	13.019	13.239	976.260	1155.4000
19	13.200	12.986	13.019	0.002	929.980	0.8194

Πίνακας 7.4. (συνέχεια) Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον ακτινικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη	Αντικειμενική	Βέλτιστη	Βέλτιστη	C	γ	ϵ
Αλγόριθμου	Συνάρτηση	Παρατηρ.	Υπολογιζ.			
20	12.971	12.971	13.011	1.120	563.630	573.4200
21	12.984	12.971	13.002	0.003	245.070	567.5700
22	13.200	12.971	13.003	0.975	0.001	0.8222
23	11.388	11.388	11.388	395.690	970.350	15.1090
24	11.361	11.361	11.361	991.830	980.910	3.8566
25	11.364	11.361	11.361	995.970	975.070	12.4840
26	11.398	11.361	11.350	375.760	991.920	6.3898
27	11.439	11.361	11.375	827.090	808.590	10.5340
28	11.374	11.361	11.364	985.350	941.980	8.8048
29	11.382	11.361	11.370	948.450	936.990	6.4442
30	11.470	11.361	11.374	308.170	877.960	0.8543



Εικόνα 7.8. Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον ακτινικό πυρήνα (RBF Kernel) Μηχανών Διανυσμάτων Υποστήριξης.

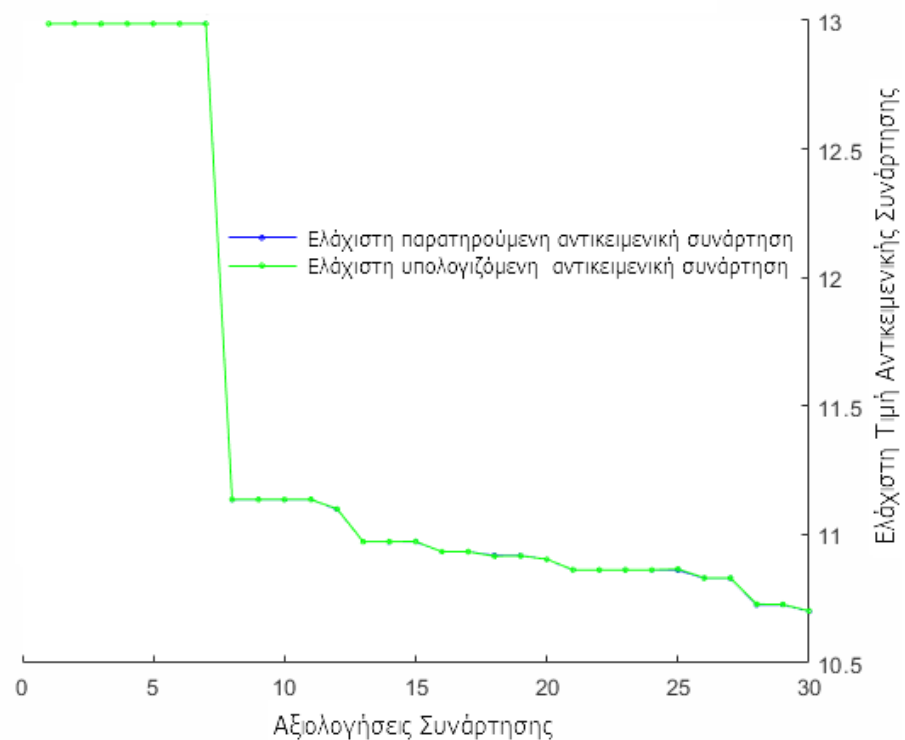
Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον πολωνυμικό πυρήνα παρουσιάζονται στον παρακάτω πίνακα.

Πίνακας 7.5. Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον πολωνυμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμ. Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογ.	C	γ	ε	Βαθμός Πολυν. r
1	12.987	12.987	12.987	732.840	0.002	2.4318	2
2	12.987	12.987	12.987	0.002	0.035	951.3900	5
3	12.987	12.987	12.987	137.020	13.231	549.3200	2
4	12.987	12.987	12.987	672.100	0.002	10442.0000	1
5	12.987	12.987	12.987	0.357	0.424	128.8400	2
6	12.987	12.987	12.987	0.398	0.088	152.8000	3
7	17.136	12.987	12.987	22.334	68.528	13.1370	4
8	11.136	11.136	11.136	8.893	546.260	0.8797	2
9	12.987	11.136	11.136	0.001	0.001	81818.0000	3
10	16.052	11.136	11.136	621.260	61.627	0.8184	4
11	12.987	11.136	11.136	4.605	0.005	0.8528	2
12	11.099	11.099	11.100	14.798	526.930	1.0233	2
13	10.972	10.972	10.972	461.180	699.220	148.0200	2
14	12.987	10.972	10.972	984.170	532.590	2112.3000	3
15	11.515	10.972	10.973	0.426	751.110	13.6150	5
16	10.933	10.933	10.933	932.880	679.940	7.6170	3
17	10.979	10.933	10.933	886.470	915.010	108.9000	5
18	10.918	10.918	10.916	920.840	695.540	35.5550	2
19	12.987	10.918	10.917	499.940	987.670	3055.0000	2
20	10.904	10.904	10.904	991.980	602.400	1.2404	3
21	10.862	10.862	10.862	976.940	550.730	5.4100	3
22	12.987	10.862	10.862	490.200	3.470	29542.0000	4
23	12.987	10.862	10.862	570.960	2.757	5.9725	2
24	12.987	10.862	10.862	362.270	0.205	75395.0000	3
25	10.884	10.862	10.865	974.190	566.700	2.8880	3
26	10.830	10.830	10.831	857.470	494.620	28.3810	2

Πίνακας 7.5. (συνέχεια) Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον πολυωνυμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμ. Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογ.	C	γ	ε	Βαθμός Πολυων. r
27	11.026	10.830	10.831	959.650	983.610	1.4617	6
28	10.727	10.727	10.729	996.430	443.100	6.1137	2
29	10.758	10.727	10.728	966.170	348.410	1.7537	4
30	10.701	10.701	10.702	869.900	358.590	5.3731	3



Εικόνα 7.9. Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον πολυωνυμικό πυρήνα (Polynomial Kernel) Μηχανών Διανυσμάτων Υποστήριξης.

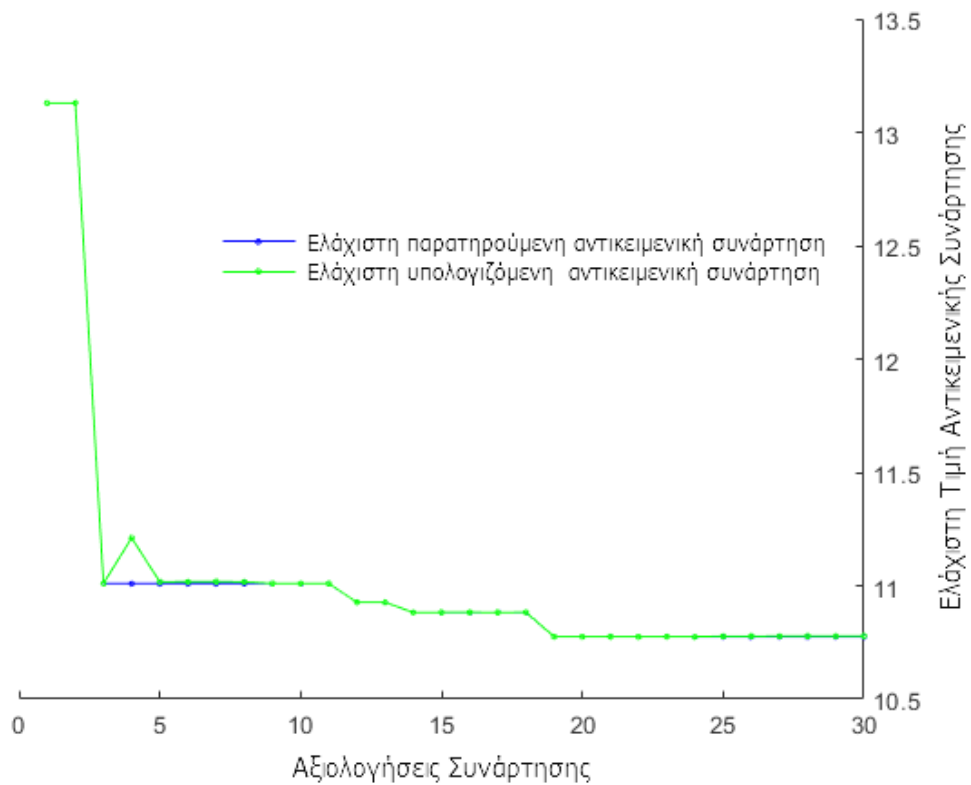
Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον γραμμικό πυρήνα παρουσιάζονται στον παρακάτω πίνακα.

Πίνακας 7.6. Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον γραμμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη Αλγορίθμου	Αντικειμενική Συνάρτηση	Βέλτιστη Παρατηρ.	Βέλτιστη Υπολογιζόμ.	C	ε
1	13.129	13.129	13.129	22.872	2408.8000
2	13.129	13.129	13.129	0.0811	5113.9000
3	11.010	11.010	11.010	0.154	0.8755
4	15.443	11.010	11.211	12.511	10.0150
5	11.028	11.010	11.017	0.079	0.8688
6	11.424	11.010	11.018	0.046	1.3405
7	13.129	11.010	11.018	9.509	81797.0000
8	17.231	11.010	11.017	10.425	276.6200
9	14.756	11.010	11.010	11.201	0.8184
10	11.579	11.010	11.010	0.097	6.3008
11	13.301	11.010	11.010	0.165	0.8212
12	10.927	10.927	10.927	0.087	5.3222
13	11.067	10.927	10.927	0.014	16.5980
14	10.882	10.882	10.882	0.002	2.9309
15	13.129	10.882	10.882	0.001	7913.4000
16	10.910	10.882	10.883	0.010	2.4389
17	13.999	10.882	10.882	0.001	0.8486
18	15.329	10.882	10.883	974.300	0.9144
19	10.775	10.775	10.775	0.011	14.9770
20	11.026	10.775	10.775	0.475	19.7150
21	10.847	10.775	10.775	0.020	2.0921
22	11.007	10.775	10.775	0.001	55.9740
23	12.510	10.775	10.775	0.119	887.9000
24	11.046	10.775	10.775	0.214	3.3692
25	10.891	10.775	10.777	0.003	9.4197

Πίνακας 7.6. (συνέχεια) Τα βέλτιστα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που βρέθηκαν για τον γραμμικό πυρήνα χρησιμοποιώντας ως κριτήριο την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Επανάληψη	Αντικειμενική	Βέλτιστη	Βέλτιστη	C	ϵ
Αλγόριθμου	Συνάρτηση	Παρατηρ.	Υπολογιζόμ.		
26	10.969	10.775	10.777	44.123	2.6088
27	10.822	10.775	10.778	0.006	12.8360
28	10.829	10.775	10.778	62.693	16.6820
29	11.011	10.775	10.778	9.987	18.6040
30	10.850	10.775	10.778	28.994	3.0874



Εικόνα 7.10. Η γραφική παράσταση της μεταβολής της ελάχιστης τιμής της αντικειμενικής συνάρτησης βάσει του αριθμού των αξιολογήσεων της συνάρτησης με τον αλγόριθμο βελτιστοποίησης Bayes όσον αφορά τον γραμμικό πυρήνα (Linear Kernel) Μηχανών Διανυσμάτων Υποστήριξης.

Όπως παρατηρούμε το καλύτερο μοντέλο των Μηχανών Διανυσμάτων Υποστήριξης (SVM) που βρέθηκε με τον αλγόριθμο Bayes είναι αυτό του πολυωνυμικού πυρήνα με τα χαρακτηριστικά υπερπαραμέτρων που αναπαρίστανται στον παρακάτω πίνακα.

Πίνακας 7.7. Τα βέλτιστα μοντέλα των Μηχανών Διανυσμάτων Υποστήριξης (SVM) που βρέθηκαν με τον αλγόριθμο Bayes για κάθε πυρήνα σύμφωνα με την ελάχιστη τιμή της αντικειμενικής συνάρτησης.

Πυρήνας Μοντέλου	Τιμή Αντικειμενικής Συνάρτησης
Γραμμικός	10.775
Πολυωνυμικός	10.701
Γκαουσιανός	11.279
Ακτινικός	11.361

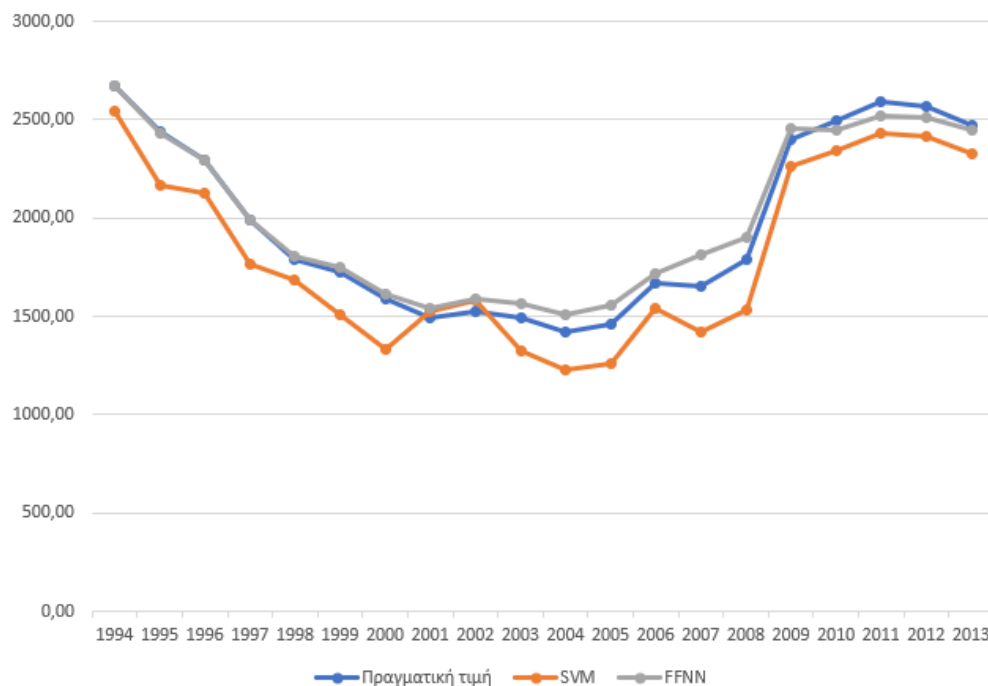
7.5.6 Σύγκριση Προβλέψεων Βέλτιστων Μοντέλων SVM και FFNN

Στην ενότητα αυτή, πραγματοποιείται σύγκριση των αποτελεσμάτων των προβλέψεων των βέλτιστων μοντέλων των Μηχανών Διανυσμάτων Υποστήριξης (SVM) και των νευρωνικών δικτύων εμπρόσθιας διάδοσης (FFNN). Όπως φαίνεται από τα αποτελέσματα, το βέλτιστο μοντέλο είναι αυτό των νευρωνικών δικτύων. Στον παρακάτω πίνακα, συγκρίνονται οι δύο προσεγγίσεις που εφαρμόστηκαν στην μελέτη αυτή.

Πίνακας 7.8. Συγκριτικός πίνακας που απεικονίζει, τα βελτιστοποιημένα μοντέλα των Μηχανών Διανυσμάτων Υποστήριξης (SVM) με τον αλγόριθμο Bayes και των νευρωνικών δικτύων εμπρόσθιας διάδοσης (FFNN) με τον αλγόριθμο PSO.

Μοντέλο	Μέθοδος Βελτιστοποίησης	Σφάλμα MSE	Σφάλμα RMSE
SVM-BAYES	Αλγόριθμος Bayes	7.5862	2.7543
FFNN-PSO	Αλγόριθμος PSO	0.0275	0.1659

Η παρακάτω εικόνα απεικονίζει τις προβλέψεις του βέλτιστου μοντέλου των Μηχανών Διανυσμάτων Υποστήριξης (πορτοκαλί χρώμα γραμμής), του βέλτιστου μοντέλου των τεχνητών νευρωνικών δικτύων (γκρι χρώμα γραμμής) και της πραγματικής τιμής (μπλε χρώμα γραμμής) για τα τελευταία 20 χρονικά βήματα (έτη) του δείγματος από το έτος 1994 έως το έτος 2013. Όπως φαίνεται από τα αποτελέσματα, το βέλτιστο μοντέλο είναι αυτό των νευρωνικών δικτύων.



Εικόνα 7.11. Συγκριτική γραφική παράσταση των προβλέψεων του βέλτιστου μοντέλου νευρωνικού δικτύου, του βέλτιστου μοντέλου SVM και της πραγματικής τιμής για τα τελευταία 20 χρονικά βήματα (έτη) του δείγματος.

7.6 Συμπεράσματα και Συζήτηση

Η εφαρμογή της τεχνητής νοημοσύνης έχει αυξηθεί σημαντικά, σε πολλά επιστημονικά πεδία, τις τελευταίες δεκαετίες με την ανάπτυξη νέων τεχνολογιών νευρωνικών δικτύων και τεχνικών μηχανικής μάθησης.

Η υιοθέτηση της τεχνητής νοημοσύνης στη δημόσια διοίκηση μπορεί να είναι πολύτιμη για τις αρχές, ειδικά σε προβλήματα που σχετίζονται με το προληπτικό μάνατζμεντ που απαιτεί την υλοποίηση τεχνικών πρόβλεψης.

Σε αυτή την έρευνα, η τεχνητή νοημοσύνη χρησιμοποιήθηκε για την ανάπτυξη μοντέλων Τεχνητών Νευρωνικών Δικτύων και Μηχανών Διανυσμάτων Υποστήριξης για την πρόβλεψη των επιπέδων ανεργίας. Εφαρμόστηκε ο τύπος των νευρωνικών δικτύων εμπρόσθιας διάδοσης, αφού θεωρείται ως ο πλέον κατάλληλος για προβλέψεις χρονοσειρών (Koskela et al., 1996; Tang and Fishwick, 1993). Αρκετοί κοινωνικοοικονομικοί παράγοντες που επηρεάζουν τα επίπεδα ανεργίας λήφθηκαν υπόψη για να αναπτυχθούν νευρωνικά μοντέλα πρόβλεψης, όπως η αύξηση του ΑΕΠ, η αύξηση των εξαγωγών αγαθών και υπηρεσιών, καθώς αυτοί οι παράγοντες επηρεάζουν τα επίπεδα ανεργίας σύμφωνα με τη βιβλιογραφία (Bayar 2014; Ozughalu and Ogwumike 2013).

Τα τελικά αποτελέσματα έδειξαν μια πολύ καλή ακρίβεια προβλέψεων των επιπέδων ανεργίας στο Ηνωμένο Βασίλειο. Το μοντέλο βέλτιστου τεχνητού νευρωνικού δικτύου αναπτύχθηκε με έρευνα διαφορετικών αρχιτεκτονικών του δικτύου. Εφαρμόζοντας την προτεινόμενη μεθοδολογία επετεύχθησαν καλύτερα αποτελέσματα πρόβλεψης σε σύγκριση με άλλες σχετικές έρευνες (Franses et al. 2004; Olmedo 2014). Αυτά τα αποτελέσματα μπορεί να είναι πολύ σημαντικά στη δημόσια διοίκηση και τους ενδιαφερόμενους φορείς ώστε να αυτοί να μπορούν να υιοθετήσουν προληπτικά μέτρα για την πρόληψη περαιτέρω αύξησης της ανεργίας που θα είχε αρνητικές επιπτώσεις στην κοινωνία και στην ποιότητα ζωής των πολιτών.

Σύμφωνα με την προτεινόμενη μεθοδολογία, που εφαρμόζεται στην παρούσα έρευνα, βελτιστοποιημένες τεχνικές τεχνητής νοημοσύνης χρησιμοποιούνται για την πρόβλεψη παραγόντων σημαντικών στην δημόσια διοίκηση, όπως στο κεφάλαιο αυτό της ανεργίας. Με το εκπαιδευμένο μοντέλο που κατασκευάστηκε, μοντελοποιήθηκε η σχέση των παραμέτρων εισόδου – εξόδου και μπορεί να χρησιμοποιηθεί, ως εργαλείο διοίκησης, από τις αρμόδιες αρχές, για να δοκιμαστούν διαφορετικά σενάρια με δεδομένα παραγόντων εισόδου του ‘σήμερα’ (τωρινό έτος) που επηρεάζουν την τιμή της ανεργίας (εξαρτημένη μεταβλητή) του ‘αύριο’, δηλαδή του επόμενου έτους. Το στοιχείο

αυτό, μπορεί να βοηθήσει αρκετά τις δημόσιες αρχές, αφού θα έχουν το χρόνο μέχρι να επέλθει το επόμενο χρονικό βήμα, για να λάβουν μέτρα στην κατεύθυνση της μείωσης της ανεργίας. Για παράδειγμα, η λήψη αποφάσεων από την Κεντρική Διοίκηση, για την αύξηση των θέσεων εργασίας, αύξηση επενδυτικών προγραμμάτων από την Ευρωπαϊκή Ένωση, κλπ.

ΚΕΦΑΛΑΙΟ 8: Τηλεπισκόπηση και Τεχνητή Νοημοσύνη στην Ταξινόμηση Δορυφορικών Εικόνων στη Δημόσια Διοίκηση

8.1 Περιγραφή της Έρευνας

Οι πιο συνηθισμένες μεθοδολογίες που έχουν προταθεί στην τηλεπισκόπηση για την ταξινόμηση δορυφορικών εικόνων αφορούσαν την χρήση στατιστικών ταξινομητών (Mills, 2011). Οι παραμετρικές μεθοδολογίες στηρίζονται στη γνώση της μορφής της κατανομής του πληθυσμού και το πρόβλημα ανάγεται σε πρόβλημα εκτίμησης των παραμέτρων του μοντέλου. Η συνηθέστερη παραμετρική μέθοδος που έχει χρησιμοποιηθεί είναι ο ταξινομητής της μέγιστης πιθανοφάνειας (Maximum Likelihood Classifier) (Strahler, 1980). Επίσης, άλλες συνηθισμένες τεχνικές ταξινόμησης δορυφορικών εικόνων είναι η ταξινόμηση με τον αλγόριθμο Bayes καθώς και η ταξινόμηση με την μεθοδολογία των εγγύτερων γειτόνων (nearest neighborhood) (Gorte and Stein, 1998). Το βασικό μειονέκτημα του ταξινομητή της μέγιστης πιθανοφάνειας, είναι ότι τα δεδομένα πρέπει να ακολουθούν την κανονική κατανομή (Strahler, 1980), που είναι δύσκολο σε προβλήματα τηλεπισκόπησης, ιδίως στην ταξινόμηση πολυφασματικών δορυφορικών εικόνων. Για το λόγο αυτό, στην παρούσα έρευνα, επιλέχθηκαν μη παραμετρικοί αλγόριθμοι ταξινόμησης βασισμένων στην τεχνητή νοημοσύνη και συγκεκριμένα οι αλγόριθμοι των Τεχνητών Νευρωνικών Δικτύων, των Μηχανών Διανυσμάτων Υποστήριξης και των Τυχαίων δασών και αναπτύχθηκε μία νέα υβριδική μεθοδολογία ταξινόμησης. Επιπλέον ένας άλλος παράγοντας επιλογής τους αποτελεί το ότι σύμφωνα με την βιβλιογραφία στους αλγόριθμους αυτούς επιδρά ελάχιστα έως καθόλου το φαινόμενο Hughes, που αλλιώς θα μείωνε την ακρίβεια

ταξινόμησης πολυφασματικών δορυφορικών εικόνων λόγω της ύπαρξης των πολλών διαστάσεων (Alonso et al., 2001; Hughes, 1968; Watanachaturaporn et al., 2004; Xia et al., 2015).

Η ανάπτυξη νέων τεχνολογιών τεχνητής νοημοσύνης, έχει επηρεάσει σημαντικά την επιστήμη της τηλεπισκόπησης και έδωσε λύση σε προβλήματα που δεν μπορούσαν να αντιμετωπιστούν με την κλασική στατιστική, όπως η προϋπόθεση να ακολουθεί ένα φαινόμενο μια συγκεκριμένη στατιστική κατανομή. Σε αυτή τη μελέτη, προτείνεται μία νέα υβριδική μεθοδολογία ταξινόμησης δορυφορικών εικόνων, συνδυάζοντας τις μεθοδολογίες τεχνητής νοημοσύνης, νευρωνικών δικτύων, Μηχανών Διανυσμάτων Υποστήριξης (Support Vector Machines - SVMs), Τυχαία Δάση (Random Forests). Η προτεινόμενη μεθοδολογία αποσκοπεί στην βελτιστοποίηση των μεθόδων ταξινόμησης, με το να συνδυάσει τα πλεονεκτήματα των παραπάνω μεθόδων για μία βελτιστοποιημένη ταξινόμηση των δορυφορικών εικόνων. Στο τέλος, τα αποτελέσματα της ταξινόμησης με την προτεινόμενη υβριδική μέθοδο, συγκρίνονται με αυτά των αυθεντικών αλγορίθμων Μηχανών Διανυσμάτων Υποστήριξης (Support Vector Machines - SVMs) και των Τυχαίων Δασών Δέντρων Απόφασης.

Οι δορυφορικές εικόνες ανακτήθηκαν για την περιοχή μελέτης που είναι η περιοχή της Μαγνησίας και στη συνέχεια, τα πολυφασματικά δεδομένα των εικονοστοιχείων της δορυφορικής εικόνας, εξήχθησαν και προεπεξεργάστηκαν ώστε να μπορούν να εφαρμοστούν τα μοντέλα μηχανικής μάθησης. Η παρούσα μελέτη ερευνά την εφαρμογή μοντέλων τεχνητής νοημοσύνης καθώς και την διερεύνηση της βέλτιστης τοπολογίας των μοντέλων αυτών για την επίτευξη καλύτερων προβλέψεων ταξινόμησης δορυφορικών εικόνων. Τα αποτελέσματα της έρευνας μπορούν να χρησιμοποιηθούν ως εργαλείο λήψης αποφάσεων στη δημόσια διοίκηση και το χωροταξικό σχεδιασμό.

8.2 Εισαγωγή

8.2.1 Ταξινόμηση Δορυφορικών Εικόνων

Η ταξινόμηση μίας δορυφορικής πολυφασματικής εικόνας, αφορά την διαδικασία της κατηγοριοποίησης των εικονοστοιχείων της σε τάξεις, που έχουν με κοινά χαρακτηριστικά, φασματικής απόκρισης (Μερτίκας, 1999). Η ταξινόμηση που διενεργείται βάσει της φασματικής υπογραφής του κάθε εικονοστοιχείου, ονομάζεται πολυφασματική ταξινόμηση (multispectral classification) (Καρτάλης και Φεΐδας, 2013). Οι ομαδοποιήσεις στο φασματικό χώρο, αντιστοιχούν σε μια γεωλογική ή φυσική γεωγραφική ενότητα στο χώρο με κοινά χαρακτηριστικά (Σκιάνης, κ.α., 2012). Στην πολυφασματική ταξινόμηση, κάθε εικονοστοιχείο εντάσσεται σε μια περιοχή ομαδοποίησης ή ταξινόμησης, με βάση κάποιο κριτήριο φασματικής υπογραφής. Μια περιοχή ομαδοποίησης μπορεί να αφορά έναν τύπο κάλυψης γης (π.χ. δασική βλάστηση) ή κάποιον άλλο τύπο οντότητας εδαφικού ή γεωλογικού χαρακτήρα όπως για παράδειγμα έναν τύπο πετρώματος (Sabins, 2007).

Υπάρχουν πολλών ειδών μέθοδοι ταξινόμησης δορυφορικών εικόνων, που βασίζονται σε διάφορες επιστήμες όπως η πληροφορική, η στατιστική, τα μαθηματικά, όπως είναι η θεωρία πιθανοτήτων, η θεωρία αλγορίθμων, η τεχνητή νοημοσύνη, η πολυπαραμετρική στατιστική, η ασαφής λογική, τα νευρωνικά δίκτυα, οι μηχανές διανυσμάτων υποστήριξης και τα τυχαία δάση δέντρων απόφασης.

8.2.2 Μηχανές Διανυσμάτων Υποστήριξης

Οι Μηχανές Διανυσμάτων Υποστήριξης (SVM) είναι ένα σύνολο αλγορίθμων μηχανικής μάθησης που εφαρμόζονται στην ταξινόμηση και την παλινδρόμηση. Η θεωρία του Μηχανών Διανυσμάτων Υποστήριξης προτάθηκε σχετικά πρόσφατα (Vapnik, 1995). Βασίζονται στον αλγόριθμο νευρωνικών δικτύων τύπου Perceptron (Βλαχάβας κ.α., 2006). Οι Μηχανές Διανυσμάτων Υποστήριξης χρησιμοποιούνται για τον εντοπισμό του βέλτιστου διαχωριστικού περιθωρίου για τον διαχωρισμό κατηγοριών σε ένα σύνολο δεδομένων. Τα δείγματα που βρίσκονται στο μεγιστοποιημένο περιθώριο

ονομάζονται διανύσματα υποστήριξης και το επίπεδο διχοτόμησης του περιθωρίου ονομάζεται διαχωριστικό υπερεπίπεδο.

Υπάρχουν διαφόρων ειδών πυρήνων που χρησιμοποιούνται στις Μηχανές Διανυσμάτων Υποστήριξης (SVM). Η συνάρτηση του πυρήνα μπορεί να θεωρηθεί ως μέτρο ομοιότητας μεταξύ των δειγμάτων x_i και x_j η οποία εφαρμόζεται στους ταξινομητές SVM για να εκτελούν διαχωρισμούς μεταξύ των κλάσεων ακόμη και όταν υπάρχουν πολύπλοκα όρια (Suykens and Vandewalle, 1999). Η συνάρτηση για τον γραμμικό πυρήνα (Linear Kernel) εκφράζεται από την σχέση:

$$K(x_i, x_j) = \langle x_i, x_j \rangle \quad (8.1)$$

Η συνάρτηση για τον πολωνυμικό πυρήνα (Polynomial Kernel) εκφράζεται από την σχέση:

$$K(x_i, x_j) = (\sigma(\gamma) * x_i^T x_j + r)^d, \quad \sigma > 0 \quad (8.2)$$

Όπου d αναπαριστά τον βαθμό του πολωνύμου (degree), το r την σταθερή παράμετρο, η συνάρτηση σ (σίγμα) εξαρτάται από την παράμετρο γ (γάμμα) που ελέγχει το πλάτος του πυρήνα. Η σχέση του γ με το σ εκφράζεται σύμφωνα με την εξίσωση:

$$\gamma = \frac{1}{\sigma^2} \quad (8.3)$$

Οι ακτινικοί πυρήνες RBF (Radial Basis Function) είναι τύποι πυρήνων όπου ένα μέτρο απόστασης εξομαλύνεται από μια ακτινική (εκθετική συνάρτηση). Στην ακτινωτή συνάρτηση πυρήνα εφαρμόζεται παρακάτω εξίσωση:

$$K(x_i, x_j) = e^{-\sigma(\gamma)\|x_i - x_j\|^2}, \quad \sigma > 0 \quad (8.4)$$

8.2.3 Τυχαία Δάση Δέντρων Απόφασης

Η τεχνική μηχανικής μάθησης των Τυχαίων Δασών είναι μια ομάδα μεγάλου αριθμού ταξινομητών - δέντρων αποφάσεων. Κάθε δέντρο αποφάσεως αναπτύσσεται ανάλογα με την τυχαία δειγματοληψία με αντικατάσταση των αρχικών δεδομένων εκπαίδευσης. Η κλάση μιας παρατήρησης που θα προβλεφθεί στην ταξινόμηση, καθορίζεται από την πλειοψηφία των ψήφων από το σύνολο των δέντρων αποφάσεων που αναπτύσσονται στο τυχαίο δάσος (Ghosh and Joshi, 2014).

Ο αλγόριθμος των τυχαίων δασών αποτελείται από δύο στάδια. Συγκεκριμένα, στο πρώτο στάδιο ακολουθούνται τα παρακάτω βήματα, για n δείγματα, m χαρακτηριστικών ή μεταβλητών για την δημιουργία τυχαίου δάσους, ενώ στο δεύτερο στάδιο πραγματοποιούνται οι προβλέψεις βάσει του δημιουργηθέντος μοντέλου (Liaw and Wiener, 2002). Συνοπτικά τα στάδια του αλγορίθμου περιγράφονται παρακάτω:

1ο στάδιο

1. Επέλεξε τυχαία χαρακτηριστικά K από τα συνολικά χαρακτηριστικά m όπου $k < m$
2. Μεταξύ των χαρακτηριστικών K , υπολόγισε τον κόμβο d χρησιμοποιώντας το καλύτερο σημείο διαίρεσης
3. Διαχώρισε τον κόμβο σε θυγατρικούς κόμβους χρησιμοποιώντας τον καλύτερο χωρισμό.
4. Επαναλάβετε τα βήματα 1 έως 3 έως ότου ο αριθμός των κόμβων γίνει "1"
5. Δημιούργησε το δάσος επαναλαμβάνοντας τα βήματα 1 έως 4 για " n " φορές για την δημιουργία " n " αριθμού δέντρων.

2ο στάδιο

1. Υπολόγισε τις ψήφους για κάθε στόχο που προβλέφθηκε.
2. Θεώρησε τον ψήφο της πλειοψηφίας (majority vote) του στόχου που προβλέφθηκε ως την τελική πρόβλεψη από τον αλγόριθμο τυχαίων δασών.

8.2 Σχετική Έρευνα

Η ταξινόμηση δορυφορικών εικόνων έχει χρησιμοποιηθεί σε διάφορες έρευνες στη δημόσια διοίκηση με σκοπό την διαχείριση των χρήσεων γης, των δασικών εκτάσεων, αλλά και την διαχείριση των υδάτων (Cheema and Bastiaanssen, 2010; He et al., 1998; Stefanov et al., 2001). Οι He et al. (1998), μελέτησαν την εφαρμογή ταξινομημένων δορυφορικών εικόνων και των γεωγραφικών συστημάτων πληροφοριών στην αξιολόγηση αστικών περιοχών κάλυψης πρασίνου. Οι δορυφορικές εικόνες ταξινομήθηκαν ώστε να μοντελοποιηθούν οι περιοχές με δασικές εκτάσεις. Επίσης, οι Stefanov et al. (2001), εφάρμοσαν τεχνικές ταξινόμησης καλύψεων-χρήσεων γης για την παρακολούθηση των ξηρών αστικών καλύψεων γης. Επίσης, οι Cheema and Bastiaanssen (2010), ερεύνησαν την ταξινόμηση χρήσεων και καλύψεων γης χρησιμοποιώντας δορυφορικά δεδομένα για την υποστήριξη της διαχείρισης των υδάτων.

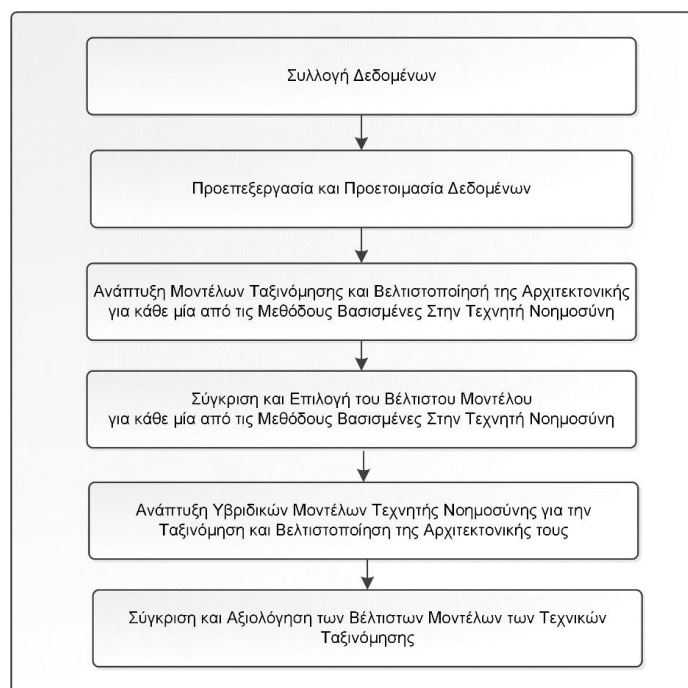
Η αλματώδης ανάπτυξη των τεχνολογιών της τεχνητής νοημοσύνης έχει επηρεάσει σημαντικά τη δημόσια διοίκηση και τον χωροταξικό σχεδιασμό. Διάφορες τεχνικές τηλεπισκόπησης σε συνδυασμό με μεθοδολογίες τεχνητής νοημοσύνης έχουν εφαρμοστεί για μία πιο βιώσιμη δημόσια διοίκηση και αστική ανάπτυξη (Gualtieri and Cromp, 2012; Rodriguez-Galiano et al., 2012). Οι Μηχανές Διανυσμάτων Υποστήριξης, τα νευρωνικά δίκτυα και τα Τυχαία Δάση, χρησιμοποιήθηκαν από διάφορους ερευνητές στην ταξινόμηση δορυφορικών εικόνων (Civco, 1993; Gualtieri and Cromp, 2012; Huang et al., 2002; Pal, 2005; Rodriguez-Galiano et al., 2012).

Σε μία μελέτη, οι Huang et al., (2002), αξιολόγησαν την εφαρμογή των Μηχανών Διανυσμάτων Υποστήριξης για την ταξινόμηση καλύψεως γης. Τα αποτελέσματα έδειξαν ότι ο αλγόριθμος των Μηχανών Διανυσμάτων Υποστήριξης έδειξε την μεγαλύτερη ακρίβεια και σταθερότητα συγκρινόμενος με άλλες μεθοδολογίες ταξινόμησης. Οι Gualtieri and Cromp (2012), μελέτησαν την εφαρμογή των Μηχανών Διανυσμάτων Υποστήριξης στην υπερφασματική τηλεπισκοπική ταξινόμηση. Τα τελικά αποτελέσματα έδειξαν πολύ καλή ακρίβεια στην υπερφασματική ταξινόμηση της δορυφορικής εικόνας. Επίσης, οι Rodriguez-Galiano et al. (2012), αξιολόγησαν την

αποτελεσματικότητα του αλγορίθμου των Τυχαίων Δασών στην ταξινόμηση καλύψεων γης, και στη συνέχεια σύγκριναν τα αποτελέσματα ταξινόμησης με αυτά ενός δέντρου ταξινόμησης. Τα αποτελέσματα έδειξαν ότι ο αλγόριθμος των Τυχαίων Δασών παράγαγε φανερά καλύτερα αποτελέσματα ταξινόμησης καλύψεων γης σε σύγκριση με το δέντρο ταξινόμησης. Ο Cincio, (1993), μελέτησε την εφαρμογή των τεχνητών νευρωνικών δικτύων στην ταξινόμηση καλύψεων γης και τα αποτελέσματα έδειξαν αρκετά καλή ακρίβεια.

8.3 Μεθοδολογία

Στην παρούσα έρευνα, η μεθοδολογία που ακολουθήθηκε για την ταξινόμηση της δορυφορικής εικόνας περιλαμβάνει τα στάδια: συλλογή δεδομένων, προεπεξεργασία και προετοιμασία δεδομένων, ανάπτυξη των μοντέλων τεχνητής νοημοσύνης, βελτιστοποίησή της αρχιτεκτονικής τους, ανάπτυξη υβριδικών μοντέλων, σύγκριση και αξιολόγηση των μοντέλων για την επιλογή του βέλτιστου. Η εικόνα 8.1, αναπαριστά μια επισκόπηση της ακολουθείσας μεθοδολογίας.



Εικόνα 8.1 Επισκόπηση της ακολουθείσας μεθοδολογίας για την ταξινόμηση της δορυφορικής εικόνας.

Αρχικά, ανακτήθηκε η δορυφορική εικόνα για την περιοχή μελέτης που είναι η περιοχή της Μαγνησίας καθώς και οι τάξεις χρήσεων γης CORINE 2012 που χρησιμοποιήθηκαν ως αναφορά στις ταξινομήσεις με τους αλγόριθμους τεχνητής νοημοσύνης. Στο επόμενο στάδιο, τα δεδομένα της δορυφορικής εικόνας εξήχθησαν και προεπεξεργάστηκαν και κανονικοποιήθηκαν στο διάστημα $[0,1]$, ώστε να μπορούν να χρησιμοποιηθούν ως διανύσματα εισόδου στα μοντέλα μηχανικής μάθησης των Μηχανών Διανυσμάτων Υποστήριξης και των Τυχαίων Δασών Δέντρων Απόφασης καθώς και στα προτεινόμενα υβριδικά μοντέλα. Χρησιμοποιήθηκε η κανονικοποίηση min-max (ελαχίστου – μεγίστου) στο διάστημα $[0,1]$, η οποία μετασχηματίζει γραμμικά τα δεδομένα, καθώς θεωρείται η καλύτερη και εξάγει ακριβέστερα αποτελέσματα από άλλες μεθόδους κανονικοποίησης, όπως η κανονικοποίηση z-σκορ και εκφράζεται από την ακόλουθη εξίσωση (Al Shalabi et al., 2006):

$$n_i = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (8.5)$$

Όπου:

n_i εκφράζει την i-οστή κανονικοποιημένη τιμή του δείγματος

x_i εκφράζει την i-οστή τιμή του δείγματος

$\min(x)$ εκφράζει την ελάχιστη τιμή του δείγματος

$\max(x)$ εκφράζει την μέγιστη τιμή του δείγματος

Για την ανάπτυξη των μοντέλων τεχνητών νευρωνικών δικτύων χρησιμοποιήθηκε το λογισμικό MATLAB, για την ανάπτυξη των μοντέλων Μηχανών Διανυσμάτων Υποστήριξης και Τυχαίων Δασών Δέντρων Απόφασης χρησιμοποιήθηκε η γλώσσα Python. Στην επόμενη φάση, αναπτύχθηκαν τα μοντέλα ταξινόμησης των αλγορίθμων τεχνητής νοημοσύνης και βελτιστοποιήθηκαν με την κατάλληλη ρύθμιση της αρχιτεκτονικής τους.

Στην επόμενη φάση, αναπτύχθηκαν τα μοντέλα ταξινόμησης των υβριδικών μοντέλων και βελτιστοποιήθηκαν με την κατάλληλη ρύθμιση της αρχιτεκτονικής τους. Στο τέλος αξιολογήθηκαν και συγκρίθηκαν τα βέλτιστα μοντέλα των τεχνικών μηχανικής μάθησης χρησιμοποιώντας διάφορες μετρικές, ώστε να επιλεγθεί το βέλτιστο.

8.4 Συλλογή και προεπεξεργασία δεδομένων

Στην έρευνα αυτή, χρησιμοποιήθηκαν ψηφιδωτά γεωχωρικά υπόβαθρα που καλύπτουν της περιοχή της Μαγνησίας. Η δορυφορική πολυφασματική εικόνα ανακτήθηκε από την ιστοσελίδα της Αμερικάνικης Γεωλογικής Υπηρεσίας⁵ (United States Geological Survey). Το αρχείο με τις τάξεις χρήσεων γης CORINE 2012 που χρησιμοποιήθηκε ως αναφορά στις ταξινομήσεις με τους αλγόριθμους μηχανικής μάθησης, ανακτήθηκε από την επίσημη ιστοσελίδα του προγράμματος της Ευρωπαϊκής Ένωσης Copernicus⁶. Πριν από την ανάλυση, πραγματοποιήθηκε προεπεξεργασία των δεδομένων της δορυφορικής εικόνας Landsat TM (Νοέμβριος 2011) με 7 φασματικά κανάλια, χωρικής διακριτικής ικανότητας 30 μ. χρησιμοποιώντας το ERDAS IMAGINE για την συνένωση των επτά καναλιών και του 8^{ου} καναλιού (σύννεφα) σε μία πολυφασματική εικόνα και για την εξαγωγή ενός υποσυνόλου της εικόνας που συμπεριλάμβανε την περιοχή μελέτης στην Μαγνησία. Επίσης, η προεπεξεργασία περιλαμβάνει και την δημιουργία εικόνων διαφόρων ψευδοχρωματικών συνθέσεων, καθώς και τη γεωμετρική διόρθωση και την ραδιομετρική ενίσχυση των δορυφορικών εικόνων. Στην συγκεκριμένη έρευνα, η δορυφορική εικόνα του Landsat που χρησιμοποιήθηκε ήταν διορθωμένη γεωμετρικά και ραδιομετρικά, όμως γίνεται η αναφορά αυτή διότι είναι ένα σημαντικό στάδιο στην προεπεξεργασία μία δορυφορικής εικόνας. Τα δεδομένα της δορυφορικής εικόνας εξήχθησαν και κανονικοποιήθηκαν πριν την εφαρμογή των μοντέλων ταξινόμησης. Στην επόμενη εικόνα, απεικονίζεται ένα

⁵ <https://www.usgs.gov>

⁶ <https://land.copernicus.eu>

τμήμα της περιοχής της Μαγνησίας σε ψευδοχρωματική σύνθεση των φασματικών καναλιών 2 (Blue), 3 (Green) και 4 (Red) του δορυφόρου Landsat 7.

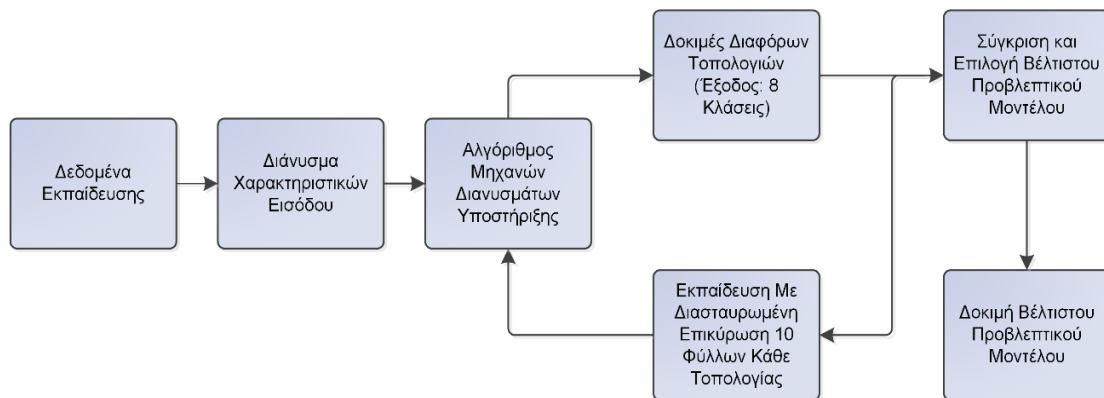


Εικόνα 8.2 Δορυφορική εικόνα Landsat 7 χωρικής ανάλυσης 30μ.

8.5 Ανάπτυξη Μοντέλων Μηχανών Διανυσμάτων Υποστήριξης

Οι μηχανές διανυσμάτων υποστήριξης έχουν αναφερθεί ότι παράγουν αποτελέσματα μεγαλύτερης ακρίβειας σε σχέση με τις παραδοσιακές προσεγγίσεις ταξινόμησης (Huang et al., 2002). Το αποτέλεσμα εξαρτάται από τον πυρήνα που χρησιμοποιήθηκε, την κατάλληλη επιλογή παραμέτρων για τον επιλεγμένο πυρήνα και τη μέθοδο που χρησιμοποιήθηκε για την παραγωγή των μηχανών διανυσμάτων υποστήριξης.

Η μεθοδολογία δημιουργίας του βέλτιστου μοντέλου πρόβλεψης Μηχανών Διανυσμάτων Υποστήριξης για την ταξινόμηση της δορυφορικής εικόνας αναπαρίσταται στην εικόνα 8.3.



Εικόνα 8.3 Μεθοδολογία δημιουργίας βέλτιστου μοντέλου πρόβλεψης Μηχανών Διανυσμάτων Υποστήριξης στην ταξινόμηση της δορυφορικής εικόνας.

Τα δεδομένα εκπαίδευσης εξήχθησαν από την προς ταξινόμηση δορυφορική εικόνα. Κάθε εικονοστοιχείο της εικόνας έχει 8 διαστάσεις που ανταποκρίνονται στις 8 μπάντες της δορυφορικής εικόνας. Λήφθηκε τυχαίο δείγμα 2000 εικονοστοιχείων, που χωρίστηκε σε τρία ανεξάρτητα μεταξύ τους μέρη. Το 70% του δείγματος αποτελεί το σύνολο εκπαίδευσης, το 15% το σύνολο επικύρωσης και το 15% το σύνολο δοκιμής.

Ο αριθμός των εικονοστοιχείων επιλέχθηκε σύμφωνα με την διεθνή βιβλιογραφία (Townshend et al., 1987; Watanachaturaporn et al., 2004) ώστε κάθε κλάση να περιέχει ικανό αριθμό εικονοστοιχείων κατά την ταξινόμηση σε κάθε υποσύνολο του δείγματος. Το σύνολο επικύρωσης θα αποτελείται από 300 εικονοστοιχεία, το σύνολο δοκιμής επίσης από 300 εικονοστοιχεία και το σύνολο εκπαίδευσης από 1400 εικονοστοιχεία. Ο ίδιος τρόπος επιλογής του δείγματος εφαρμόστηκε σε όλα τα μοντέλα τεχνητής νοημοσύνης που ερευνήθηκαν στην παρούσα μελέτη. Το διάνυσμα εισόδου αποτελείται από 2000 δείγματα των 8 παραμέτρων εισόδου. Η έξοδος αποτελείται από 8 κλάσεις. Εφαρμόστηκε ο αλγόριθμος των Μηχανών Διανυσμάτων Υποστήριξης και ερευνήθηκαν διάφορες τοπολογίες για να βρεθεί το βέλτιστο μοντέλο.

Αναπτύχθηκαν διάφορα μοντέλα SVM με εφαρμογή διαφόρων ειδών πυρήνων και σχετικών υπερπαραμέτρων που ερευνήθηκαν για την βελτιστοποίηση των αποτελεσμάτων. Χρησιμοποιήθηκε ο αλγόριθμος αναζήτησης πλέγματος (Grid Search),

όπως έχει περιγραφεί σε προηγούμενα κεφάλαια, για την βελτιστοποίηση των υπερπαραμέτρων. Ο λόγος που χρησιμοποιήθηκε η τεχνική αυτή και όχι για παράδειγμα η βελτιστοποίηση Bayes, είναι ότι η αναζήτηση είναι πλήρης και εγγυάται την εύρεση του ολικού βέλτιστου, ενώ άλλες βελτιστοποιήσεις, όπως η βελτιστοποίηση Bayes εμφανίζουν δυσκολίες σε προβλήματα υψηλών διαστάσεων (Mendoza et al., 2016; Swersky et al., 2014).

Στην αναζήτηση πλέγματος, οι υπερπαραμέτροι που ερευνήθηκαν ήταν: Η τιμή κόστους (C) σε όλους τους πυρήνες ερευνήθηκε για τις τιμές 10, 100, και από την τιμή 1000 έως 10000 με βήμα αύξησης 200. Η παράμετρος γάμμα (γ) των πυρήνων εκτός του γραμμικού (που δεν τον επηρεάζει) δοκιμάστηκε για τις τιμές από 1 έως 0.001 με βήμα 0.001. Για τον πολυωνυμικό πυρήνα ο βαθμός του πολυωνύμου που ερευνήθηκε ήταν από 1 έως 10 που θεωρούνται οι πιο συνήθεις. Η παράμετρος r που χρησιμοποιείται στην συνάρτηση του πολυωνυμικού πυρήνα ερευνήθηκε για τις τιμές 0, 0.1 έως 0.5 με βήμα 0.1 και 1 έως 5 με βήμα 1, και η παράμετρος ϵ για τις τιμές 0.1 έως 1 με βήμα 0.1 που θεωρούνται οι πιο συνήθεις (Gunn, 1998).

Χρησιμοποιήθηκε η διασταυρωμένη επικύρωση με 10 φύλλα (10-fold Cross-Validation) για την εύρεση του καλύτερου μοντέλου και την αποφυγή της υπερεκπαίδευσης (overfitting). Σύμφωνα με την μεθοδολογία διασταυρωμένης επικύρωσης με 10 φύλλα, το σύνολο εκπαίδευσης χωρίστηκε σε 10 υποσύνολα.

Για καθένα από τα 10 υποσύνολα χρησιμοποιήθηκαν τα 9 για την εκπαίδευση του μοντέλου και το ένα για την δοκιμή του μοντέλου. Το πλεονέκτημα της διασταυρωμένης επικύρωσης με 10 φύλλα σε σχέση με άλλα είδη επικύρωσης (όπως για παράδειγμα η τυχαία δειγματοληψία υποσυνόλων (Random Subsampling)) είναι ότι όλα τα δείγματα χρησιμοποιούνται κάποια στιγμή και στην διαδικασία της εκπαίδευσης και στην δοκιμή.

Το βέλτιστο μοντέλο επιλέχθηκε σύμφωνα με το μικρότερο σφάλμα διασταυρωμένης επικύρωσης με 10 φύλλα, και είναι αυτό με πυρήνα ακτινικό (radial), κόστος $C = 1000$, $\gamma = 0.7$ και $\epsilon = 0.1$ και σφάλμα 8.2%.

Ο επόμενος πίνακας, συνοψίζει τα μοντέλα που αναπτύχθηκαν για τους πυρήνες: γραμμικό, πολυωνυμικό και ακτινικό. Όπως παρατηρούμε από τα αποτελέσματα ο γραμμικός και ο πολυωνυμικός πυρήνας δεν είναι κατάλληλοι για αυτή την ταξινόμηση, καθώς παρουσιάζουν μεγάλο σφάλμα διασταυρωμένης επικύρωσης.

Πίνακας 8.1. Πίνακας σύγκρισης των καλύτερων μοντέλων κάθε πυρήνα των Μηχανών Διανυσμάτων Υποστήριξης βάσει του σφάλματος διασταυρωμένης επικύρωσης 10 φύλλων, για την ταξινόμηση της δορυφορικής εικόνας.

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΩΝ ΜΟΝΤΕΛΩΝ							
ΜΟΝΤΕΛΑ	Πυρήνας	Κόστος (C)	Γάμμα γ	Βαθμός Πολυωνύμου	Παράμετρος r	Παράμετρος epsilon	Σφάλμα % Διασταυρωμένης Επικύρωσης
Γραμμικά							
SVM1	Γραμμικός	10	-	-	-	0.2	85.4
SVM2	Γραμμικός	100	-	-	-	0.1	75.3
SVM3	Γραμμικός	1000	-	-	-	0.1	69.8
Ακτινικά							
SVM4	Ακτινικός	100	1.0	-	-	0.1	17.2
SVM5	Ακτινικός	1000	0.7			0.1	8.2
SVM6	Ακτινικός	10000	0.1	-	-	0.1	12.1
Πολυωνυμικά							
SVM7	Πολυωνυμικός	100	0.6	6	0.1	0.1	66.5
SVM8	Πολυωνυμικός	100	0.7	5	0.0	0.1	50.5
SVM9	Πολυωνυμικός	100	1.0	6	0.0	0.1	65.2

Στη συνέχεια, το βέλτιστο μοντέλο SVM5 δοκιμάστηκε στο σύνολο δοκιμής και η ακρίβεια πρόβλεψης βρέθηκε 91.8%. Στον παρακάτω πίνακα, παρατίθενται τα αποτελέσματα δοκιμής του βέλτιστου μοντέλου στο σύνολο δοκιμής, αλλά και τα

συγκριτικά αποτελέσματα των υπόλοιπων καλύτερων μοντέλων, που όπως φαίνεται διατηρείται η βέλτιστη ακρίβεια του μοντέλου και στο σετ δοκιμής.

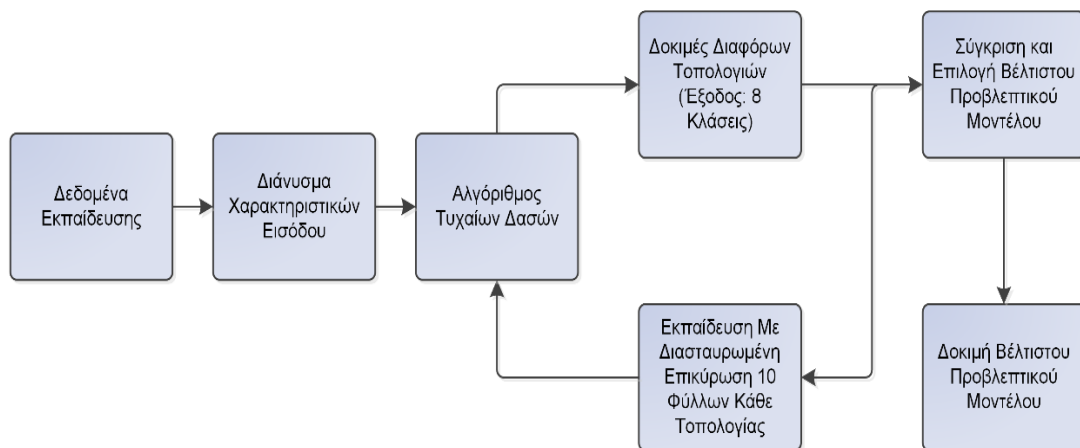
Πίνακας 8.2. Πίνακας σύγκρισης των καλύτερων μοντέλων κάθε πυρήνα των Μηχανών Διανυσμάτων Υποστήριξης στο σύνολο δοκιμής για την ταξινόμηση της δορυφορικής εικόνας.

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΩΝ ΜΟΝΤΕΛΩΝ							
ΜΟΝΤΕΛΑ	Πυρήνας	Κόστος (C)	Γάμμα γ	Βαθμός Πολυωνύμου	Παράμετρος r	Παράμετρος epsilon	Ακρίβεια % Δοκιμής Μοντέλου
Γραμμικά							
SVM1	Γραμμικός	10	-	-	-	0.2	15.5
SVM2	Γραμμικός	100	-	-	-	0.1	25.2
SVM3	Γραμμικός	1000	-	-	-	0.1	31.3
Ακτινικά							
SVM4	Ακτινικός	100	1	-	-	0.1	82.9
SVM5	Ακτινικός	1000	0.7			0.1	91.8
SVM6	Ακτινικός	10000	0.1	-	-	0.1	87.9
Πολυωνυμικά							
SVM7	Πολυωνυμικός	100	0.6	6	0.1	0.1	33.9
SVM8	Πολυωνυμικός	100	0.7	5	0.0	0.1	51.8
SVM9	Πολυωνυμικός	100	1	6	0.0	0.1	36.8

8.6 Ανάπτυξη μοντέλων με Τυχαία Δάση Δέντρων Απόφασης

Στην παρούσα μελέτη, ερευνήθηκε η ανάπτυξη μοντέλων Τυχαίων Δασών με δοκιμές διαφόρων δομών τυχαίων δασών για την βελτιστοποίηση των αποτελεσμάτων. Το σύνολο εκπαίδευσης ήταν το ίδιο με αυτό που χρησιμοποιήθηκε στην ταξινόμηση με Μηχανές Διανυσμάτων Υποστήριξης ώστε να είναι δυνατή η σύγκριση των δύο μοντέλων.

Τα δεδομένα εκπαίδευσης εξήχθησαν από την προς ταξινόμηση δορυφορική εικόνα. Κάθε εικονοστοιχείο της εικόνας έχει 8 διαστάσεις που ανταποκρίνονται στις μπάντες της δορυφορικής εικόνας. Για το σύνολο εκπαίδευσης λήφθηκε το ίδιο δείγμα των 2000 εικονοστοιχεία, όπως και στην εφαρμογή του μοντέλου Μηχανών Διανυσμάτων Υποστήριξης. Το διάνυσμα εισόδου αποτελείται από 2000 δείγματα των 8 παραμέτρων εισόδου που αντιστοιχούν στις 8 μπάντες της δορυφορικής εικόνας. Η έξοδος αποτελείται από 8 κλάσεις. Εφαρμόστηκε ο αλγόριθμος των Τυχαίων Δασών και ερευνήθηκαν διάφορες τοπολογίες για να βρεθεί το βέλτιστο μοντέλο. Η μεθοδολογία δημιουργίας του βέλτιστου μοντέλου πρόβλεψης Τυχαίων Δασών για την ταξινόμηση της δορυφορικής εικόνας αναπαρίσταται στην επόμενη εικόνα.



Εικόνα 8.4 Μεθοδολογία δημιουργίας βέλτιστου μοντέλου πρόβλεψης Τυχαίων Δασών στην ταξινόμηση της δορυφορικής εικόνας.

Επίσης, χρησιμοποιήθηκε η διασταυρωμένη επικύρωση με 10 φύλλα (10-fold Cross-Validation) για την εύρεση της καλύτερης τοπολογίας του μοντέλου. Σύμφωνα με

τον Breiman (1996), το μοντέλο των Τυχαίων Δασών δεν παρουσιάζει το πρόβλημα της υπερεκπαίδευσης, όμως στη βιβλιογραφία οι ερευνητές χρησιμοποιούν την διασταυρωμένη επικύρωση για την αξιολόγηση της καλύτερης τοπολογίας (Sun and Schulz, 2015). Σύμφωνα με την μεθοδολογία διασταυρωμένης επικύρωσης με 10 φύλλα, το σύνολο εκπαίδευσης χωρίστηκε σε 10 υποσύνολα. Για καθένα από τα 10 υποσύνολα χρησιμοποιήθηκαν τα 9 για την εκπαίδευση του μοντέλου και το ένα για την δοκιμή του μοντέλου.

Το βέλτιστο μοντέλο επιλέχθηκε σύμφωνα με το μικρότερο σφάλμα διασταυρωμένης επικύρωσης με 10 φύλλα βρέθηκε 8.1%. Ο παρακάτω πίνακας, συνοψίζει τα μοντέλα που αναπτύχθηκαν με διαφορετικό αριθμό τυχαίων δασών. Από τον πίνακα το βέλτιστο μοντέλο με το μικρότερο σφάλμα διασταυρωμένης επικύρωσης προκύπτει ότι είναι το RF6.

Πίνακας 8.3. Πίνακας σύγκρισης μοντέλων ταξινόμησης με Τυχαία Δάση βάσει του του σφάλματος διασταυρωμένης επικύρωσης 10 φύλλων.

MONTELO	Αριθμός Δέντρων Δάσους	Σφάλμα % Διασταυρωμένης Επικύρωσης 10-φύλλων
RF1	50	15.9
RF2	100	9,3
RF3	200	9.5
RF4	500	9,4
RF5	1000	9.3
RF6	2000	8.1
RF7	3000	9.5
RF8	4000	9.3
RF9	8000	9.35
RF10	15000	8.90

Στη συνέχεια, το βέλτιστο μοντέλο RF6 δοκιμάστηκε στο σύνολο δοκιμής και η ακρίβεια πρόβλεψης βρέθηκε 91.9%. Στον παρακάτω πίνακα, παρατίθενται τα αποτελέσματα δοκιμής του βέλτιστου μοντέλου στο σύνολο δοκιμής, αλλά και τα συγκριτικά αποτελέσματα των άλλων εννέα καλύτερων μοντέλων, που όπως φαίνεται διατηρείται η βέλτιστη ακρίβεια του RF6 και στο σετ δοκιμής.

Πίνακας 8.4 Πίνακας σύγκρισης μοντέλων ταξινόμησης δορυφορικής εικόνας με Τυχαία Δάση στο σύνολο δοκιμής.

MONTEΛΟ	Αριθμός Δέντρων Δάσους	Ακρίβεια % Δοκιμής Μοντέλου
RF1	50	80.6
RF2	100	86.0
RF3	200	88.6
RF4	500	90.2
RF5	1000	89.8
RF6	2000	91.9
RF7	3000	90.2
RF8	4000	90.6
RF9	8000	89.8
RF10	15000	91.2

Τα τυχαία δάση μπορούν να χρησιμοποιηθούν για να λύσουν προβλήματα παλινδρόμησης και ταξινόμησης με καλύτερη ακρίβεια σε σχέση με παραδοσιακές μεθόδους ταξινόμησης (Rodriguez-Galiano et al., 2012). Το μειονέκτημα είναι ότι χρειάζονται υψηλούς χρόνους και υπολογιστικούς πόρους, ιδιαίτερα όταν ο αριθμός των δέντρων είναι αρκετά υψηλός.

8.7 Το προτεινόμενο υβριδικό μοντέλο ταξινόμησης

Η προτεινόμενη μεθοδολογία Αποτελείται από 2 στάδια. Στο πρώτο στάδιο, προβαίνουμε σε εκτέλεση του αλγορίθμου τυχαίων δασών με την διαφορά ότι εισάγεται η εκτέλεση πρόσθετων υπολογισμών για τον υπολογισμό και την εξαγωγή του δείκτη σημαντικότητας Gini ως μέτρου της εντροπίας για να μοντελοποιηθεί το διάγραμμα σημαντικότητας των μεταβλητών εισόδου και να επιλεχθούν οι πιο σημαντικές μεταβλητές. Επίσης, υπολογίζεται ο πίνακας εγγύτητας, καταμετρώντας τον αριθμό των φορών που τα δύο δείγματα καταλήγουν στον ίδιο τερματικό κόμβο ενός δέντρου του τυχαίου δάσους, διαιρούμενο με τον αριθμό των δένδρων στο τυχαία δάσος.

Κατά το δεύτερο στάδιο, εφαρμόζεται η εκτέλεση Νευρωνικού Δικτύου Ταξινόμησης και του αλγορίθμου SVM με διάνυσμα εισόδου τον πίνακα εγγύτητας και ερευνάται η εύρεση της βέλτιστης τοπολογίας τους με την χρήση διασταυρούμενης επικύρωσης 10 φύλλων και στη συνέχεια το βέλτιστο μοντέλο εφαρμόζεται στην ταξινόμηση της δορυφορικής εικόνας. Τα βήματα εφαρμογής της προτεινόμενης μεθοδολογίας περιγράφονται συνοπτικά παρακάτω:

1. Εκτέλεση αλγορίθμου τυχαίων δασών με την προσθήκη εκτέλεσης των πρόσθετων υπολογισμών για την εξαγωγή του δείκτη σημαντικότητας Gini ως μέτρου της εντροπίας και του πίνακα εγγύτητας:
 - 1.1 Υπολογισμός δείκτη σημαντικότητας Gini ως υπολογιστική προσέγγιση της εντροπίας, για την εύρεση της σημαντικότητας των μεταβλητών εισόδου.
 - 1.2 Εξαγωγή των K περισσότερο σημαντικών μεταβλητών εισόδου από τις M (με $K \leq M$) μεταβλητές εισόδου βάσει της μείωσης του δείκτη σημαντικότητας Gini.
 - 1.3 Υπολογισμός του πίνακα εγγύτητας $N \times N$ για N δείγματα, K μεταβλητών ώστε να αποτελέσει διάνυσμα εισόδου για την αλγόριθμο SVM για τον βέλτιστο διαχωρισμό των τάξεων στον K -διάστατο χώρο.

2. Χρήση του πίνακα εγγύτητας φύλλων ως διάνυσμα εισόδου σε Πολυεπίπεδο Νευρωνικό Δίκτυο και σε Μηχανές Διανυσμάτων Υποστήριξης.
 - 2.1 Εύρεση της βέλτιστης τοπολογίας τους και δόμηση του βέλτιστου μοντέλου πρόβλεψης.
 - 2.2 Ταξινόμηση της δορυφορικής εικόνας βάσει του βέλτιστου μοντέλου.

8.7.1 Στάδιο 1ο: Υπολογισμός δείκτη Gini και πίνακα εγγύτητας

8.7.1.1 Στάδιο 1.1 Υπολογισμός δείκτη σημαντικότητας Gini ως μέτρου της εντροπίας

Επίσης, κατά την εκτέλεση του αλγορίθμου τυχαίων δασών ενσωματώνουμε μηχανισμό για την εκτέλεση πράξεων για τον υπολογισμό της μείωσης του δείκτη Gini impurity ώστε να μοντελοποιήσουμε το διάγραμμα σημαντικότητας των μεταβλητών.

Μια μειωμένη τιμή του δείκτη Gini impurity σημαίνει ότι μια συγκεκριμένη μεταβλητή πρόβλεψης παίζει μεγαλύτερο ρόλο στη ταξινόμηση των δεδομένων στις καθορισμένες κατηγορίες. Τα δεδομένα των δέντρων ταξινόμησης των τυχαίων δασών χωρίζονται σε μεμονωμένους κόμβους με βάση τις τιμές των μεταβλητών πρόβλεψης.

Ο δείκτης Gini impurity $i(\tau)$ για κάθε κόμβο τ σε T δέντρα τυχαίου δάσους αποτελεί μια υπολογιστική προσέγγιση στην μέτρηση της εντροπίας για το πόσο καλά χωρίζονται τα δείγματα στις δύο κλάσεις στον συγκεκριμένο κόμβο (Menze et al., 2009). Εκφράζεται μαθηματικά από την σχέση:

$$i(\tau) = 1 - p_1^2 - p_0^2 \quad (8.6)$$

Όπου p_k είναι το κλάσμα των n_k δειγμάτων στον κόμβο τ και εκφράζεται από τον μαθηματικό τύπο:

$$p_k = \frac{n_k}{n} \quad (8.7)$$

Η μείωση του δείκτη Δi που έχει ως αποτέλεσμα των διαχωρισμό σε δύο θυγατρικούς κόμβους τ_1 και τ_2 στο δέντρο με αντίστοιχα κλάσματα δειγμάτων:

$$p_1 = \frac{n_1}{n} \quad (8.8)$$

$$p_2 = \frac{n_2}{n} \quad (8.9)$$

και με κατώφλι t_θ στην μεταβλητή θ ορίζεται ως

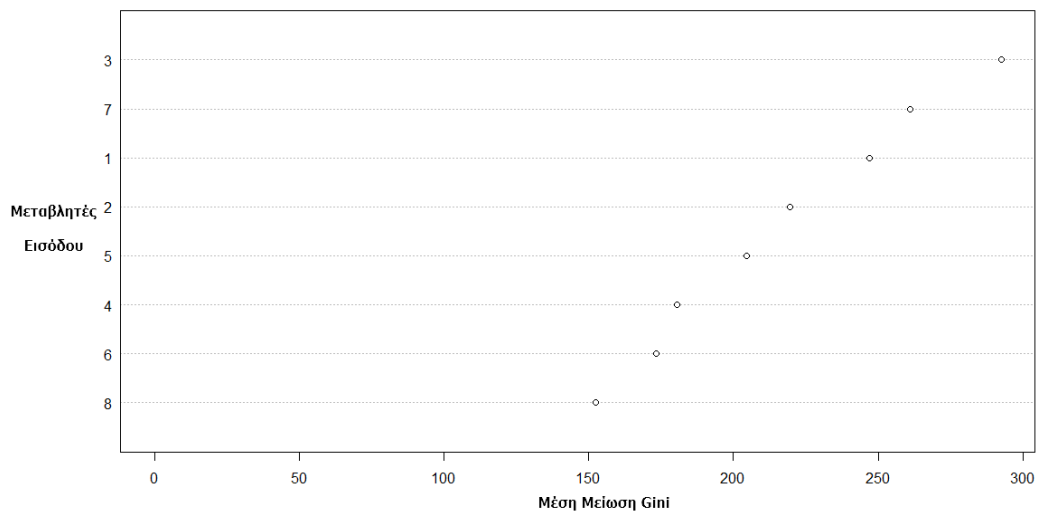
$$\Delta i(\tau) = i(\tau) - p_1 i(\tau_1) - p_2 i(\tau_2) \quad (8.10)$$

Μετά από την διάτρεξη όλων των μεταβλητών θ στον κόμβο και των πιθανών κατωφλιών t_θ καθορίζεται το Δi . Η μείωση του Gini impurity από τον βέλτιστο διαχωρισμό $\Delta i_\theta(\tau, T)$ αθροίζεται για όλους τους κόμβους τ σε όλα τα δέντρα T στο δάσος για όλες τις μεταβλητές θ . Ο δείκτης σημαντικότητας Gini I_G δείχνει πόσο συχνά ένα χαρακτηριστικό θ επιλέχθηκε για τον διαχωρισμό στον κόμβο και την συνολική διαχωριστική της αξία στο πρόβλημα ταξινόμησης και εκφράζεται μαθηματικά από την σχέση:

$$I_G(\theta) = \sum_T \sum_\tau \Delta i_\theta(\tau, T) \quad (8.11)$$

Στην παρακάτω εικόνα, απεικονίζεται το διάγραμμα σπουδαιότητας μεταβλητών εισόδου στον καθορισμό των κλάσεων μέσω του διαχωρισμού των κόμβων σε θυγατρικούς στα δέντρα στα Τυχαία Δάση. Όπως φαίνεται από την παρακάτω εικόνα την

μεγαλύτερη σπουδαιότητα στον καθορισμό των κλάσεων έχει η μεταβλητή 3 (φασματικό κανάλι 3) και μετά η μεταβλητές 7, 1 και 2. Η μικρότερη σημασία έχει η μεταβλητή 8 (φασματικό κανάλι 3).



Εικόνα 8.5 Διάγραμμα σπουδαιότητας μεταβλητών εισόδου στον καθορισμό των κλάσεων μέσω του διαχωρισμού των κόμβων σε θυγατρικούς στα δέντρα στα Τυχαία Δάση.

8.7.1.2 Στάδιο 1.2 Υπολογισμός του πίνακα εγγύτητας

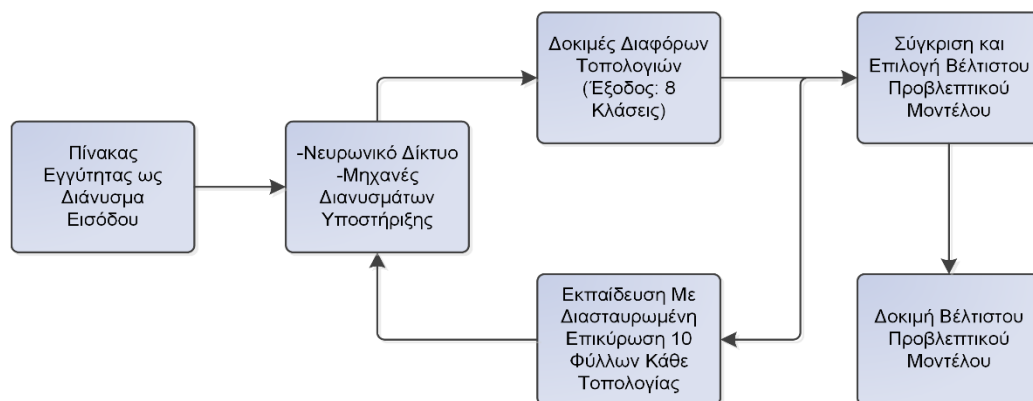
Σύμφωνα με το προτεινόμενο μοντέλο ταξινόμησης, πραγματοποιείται η εκτέλεση του μοντέλου των Τυχαίων Δασών με την διαφορά ότι κατά την εκτέλεση του αλγορίθμου τυχαίων δασών ενσωματώνουμε επίσης την εκτέλεση πράξεων για τον υπολογισμό του πίνακα εγγύτητας (proximity matrix) ανά δύο δείγματα του διανύσματος εισόδου, υπολογίζοντας τον αριθμό των φορών που τα δύο δείγματα καταλήγουν στον ίδιο τερματικό κόμβο ενός δέντρου του τυχαίου δάσους, διαιρούμενο με τον αριθμό των δένδρων στο τυχαίο δάσος.

Αυτό το μέτρο θα χρησιμοποιηθεί για τον βέλτιστο διαχωρισμό των κλάσεων στην προτεινόμενη μεθοδολογία. Κατά την προτεινόμενη μεθοδολογία, αφού υπολογιστεί ο πίνακας εγγύτητας, στη συνέχεια εφαρμόζεται ένα νευρωνικό δίκτυο ταξινόμησης και ο αλγόριθμος SVM με είσοδο τον πίνακα εγγύτητας φύλλων.

Επίσης, επειδή χρησιμοποιούνται 2000 δείγματα, ο πίνακας εισόδου θα είναι 2000 γραμμές επί 2000 στήλες. Επειδή είναι υπερβολικά μεγάλος, δεν μπορούμε να τον απεικονίσουμε στην παρούσα διατριβή.

8.7.2 Στάδιο 2ο: Τροφοδότηση Νευρωνικού Δικτύου και Μηχανών Διανυσμάτων Υποστήριξης με διάνυσμα εισόδου τον πίνακα εγγύτητας φύλλων

Στο στάδιο αυτό της προτεινόμενης μεθοδολογίας, ο πίνακας εγγύτητας που εξήχθη σε προηγούμενο στάδιο χρησιμοποιείται ως διάνυσμα εισόδου για την δημιουργία του βέλτιστου μοντέλου πρόβλεψης. Η μεθοδολογία για την ταξινόμηση της δορυφορικής εικόνας αναπαρίσταται στην εικόνα 9.6. Το διάνυσμα εισόδου έχει διαστάσεις 2000 επί 2000. Η έξοδος αποτελείται από 8 τάξεις. Τροφοδοτήθηκε νευρωνικό δίκτυο πολυεπίπεδου αισθητήρα ταξινόμησης πολυεπίπεδου αισθητήρα και μοντέλα Μηχανών Διανυσμάτων Υποστήριξης και ερευνήθηκαν διάφορες τοπολογίες για να βρεθεί το βέλτιστο μοντέλο.



Εικόνα 8.6 Το 2^ο στάδιο της προτεινόμενης μεθοδολογίας για την δημιουργίας βέλτιστου μοντέλου πρόβλεψης.

8.7.1 Μοντέλα Μηχανών Διανυσμάτων Υποστήριξης

Αναπτύχθηκαν διάφορα μοντέλα SVM με εφαρμογή διαφόρων ειδών πυρήνων και σχετικών παραμέτρων που ερευνήθηκαν για την βελτιστοποίηση των αποτελεσμάτων με τον ίδιο τρόπο που περιγράφηκε στην ενότητα ανάπτυξης μοντέλου Μηχανών Διανυσμάτων Υποστήριξης. Ερευνήθηκαν διάφορες τοπολογίες που αφορούσαν διαφορετικές τιμές κόστους (C) σε όλους τους πυρήνες και η παράμετρος γάμμα (γ) των

πυρήνων εκτός του γραμμικού (που δεν τον επηρεάζει) και ο βαθμός του πολωνύμου για τον πολωνυμικό πυρήνα (Gunn, 1998) με τον αλγόριθμο αναζήτησης πλέγματος. Χρησιμοποιήθηκε η διασταυρωμένη επικύρωση με 10 φύλλα (10-fold Cross-Validation) για την εύρεση του καλύτερου μοντέλου και την αποφυγή της υπερεκπαίδευσης (overfitting), όπως η μεθοδολογία αυτή περιγράφηκε στις προηγούμενες ενότητες.

Το βέλτιστο μοντέλο επιλέχθηκε σύμφωνα με το μικρότερο σφάλμα διασταυρωμένης επικύρωσης με 10 φύλλα, το οποίο ήταν στο ακτινικό μοντέλο με σφάλμα 5.7%, και αρχιτεκτονική: κόστος $C = 1000$, $\gamma = 0.01$ και παράμετρο $\epsilon = 0.1$. Ο παρακάτω πίνακας, συνοψίζει τα μοντέλα που αναπτύχθηκαν για τους πυρήνες: γραμμικό, πολωνυμικό και ακτινικό.

Πίνακας 8.5 Πίνακας σύγκρισης των καλύτερων μοντέλων κάθε πυρήνα των Μηχανών Διανυσμάτων Υποστήριξης του υβριδικού μοντέλου βάσει του σφάλματος διασταυρωμένης επικύρωσης 10 φύλλων, για την ταξινόμηση της δορυφορικής εικόνας.

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΩΝ ΜΟΝΤΕΛΩΝ							
ΜΟΝΤΕΛΑ	Πυρήνας	Κόστος (C)	Γάμμα γ	Βαθμός Πολωνύμου	Παράμετρος τ	Παράμετρος ϵ	Σφάλμα % Διασταυρωμένης Επικύρωσης 10-φύλλων
Γραμμικά							
H_SVM1	Γραμμικός	10	-	-	-	0.1	12.1
H_SVM2	Γραμμικός	100	-	-	-	0.1	12.4
H_SVM3	Γραμμικός	1000	-	-	-	0.1	13.01
Ακτινικά							
H_SVM4	Ακτινικός	1000	0.01	-	-	0.1	5.7
H_SVM5	Ακτινικός	1000	0.001	-	-	0.1	10.7
H_SVM6	Ακτινικός	10000	0.001	-	-	0.1	8.3
Πολωνυμικά							
H_SVM7	Πολωνυμικός	100	0.6	6	0.1	0.2	14.8
H_SVM8	Πολωνυμικός	100	0.7	6	0.0	0.1	9.9
H_SVM9	Πολωνυμικός	1000	0.8	3	0.01	0.1	11.0

Στη συνέχεια, το βέλτιστο μοντέλο της υβριδικής μεθοδολογίας H_SVM4 δοκιμάστηκε στο σύνολο δοκιμής και η ακρίβεια πρόβλεψης βρέθηκε 95.2% η οποία ήταν η καλύτερη από τα τρία γενικά μοντέλα που αναπτύχθηκαν: Τυχαίων Δασών, Μηχανών Διανυσμάτων Υποστήριξης και το Υβριδικό μοντέλο. Στον παρακάτω πίνακα, παρατίθενται τα αποτελέσματα δοκιμής του βέλτιστου μοντέλου στο σύνολο δοκιμής, αλλά και τα συγκριτικά αποτελέσματα των υπόλοιπων καλύτερων μοντέλων, που όπως φαίνεται ότι διατηρείται η βέλτιστη ακρίβεια του μοντέλου και στο σετ δοκιμής.

Πίνακας 8.6 Πίνακας σύγκρισης των καλύτερων μοντέλων κάθε πυρήνα των Μηχανών Διανυσμάτων Υποστήριξης του υβριδικού μοντέλου βάσει του σφάλματος διασταυρωμένης επικύρωσης 10 φύλλων.

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΩΝ ΜΟΝΤΕΛΩΝ							
ΜΟΝΤΕΛΑ	Πυρήνας	Κόστος (C)	Γάμμα γ	Βαθμός Πολυωνύμου	Παράμετρος τ	Παράμετρος epsilon	Ακρίβεια % Δοκιμής Μοντέλου
Γραμμικά							
H_SVM1	Γραμμικός	10	-	-	-	0.1	91.6
H_SVM2	Γραμμικός	100	-	-	-	0.1	92.2
H_SVM3	Γραμμικός	1000	-	-	-	0.1	92.4
Ακτινικά							
H_SVM4	Ακτινικός	1000	0.01	-	-	0.1	95.2
H_SVM5	Ακτινικός	1000	0.001			0.1	91.0
H_SVM6	Ακτινικός	10000	0.001	-	-	0.1	94.1
Πολυωνυμικά							
H_SVM7	Πολυωνυμικός	100	0.6	6	0.1	0.2	88.1
H_SVM8	Πολυωνυμικός	100	0.7	6	0.0	0.1	92.4
H_SVM9	Πολυωνυμικός	1000	0.8	3	0.01	0.1	88.4

8.7.2 Νευρωνικό Δίκτυο Πολυεπίπεδου Αισθητήρα

Αναπτύχθηκαν διάφορα μοντέλα νευρωνικών δικτύων με εφαρμογή διαφόρων τοπολογιών. Επειδή, το διάνυσμα εισόδου αποτελείται από 4.000.000 στοιχεία (2000 επί 2000 πίνακας) χρησιμοποιήθηκε ως αλγόριθμος εκπαίδευσης ο αλγόριθμος Κλιμακούμενων Συζυγών Κλίσεων (Scaled Conjugate Gradient – SCG), που απαιτεί λιγότερη μνήμη εν συγκρίσει με άλλους αλγορίθμους εκπαίδευσης όπως ο Levenberg Marquardt και ο αλγόριθμος Bayesian Regularization (BR) (Møller, 1993).

Τα δεδομένα εισόδου χωρίστηκαν σε τρία διαφορετικά μέρη. Το 70% των δεδομένων χρησιμοποιήθηκε ως σύνολο εκπαίδευσης, το 15% για το σύνολο επικύρωσης και 15% για το σύνολο δοκιμής. Το σύνολο δεδομένων εκπαίδευσης χρησιμοποιήθηκε για να εκπαιδευτεί το νευρωνικό δίκτυο. Το σύνολο επικύρωσης χρησιμοποιήθηκε για την αξιολόγηση των μοντέλων τεχνητών νευρωνικών δικτύων και την αποφυγή της υπερεκπαίδευσης (overfitting).

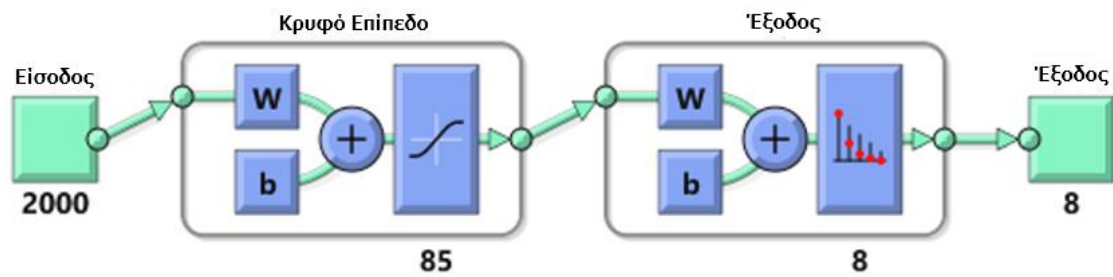
Δοκιμάστηκαν διάφορες τοπολογίες συνάρτησης μεταφοράς κρυφού επιπέδου και επιπέδου εξόδου και οι τοπολογίες συνοψίζονται στον επόμενο πίνακα. Το βέλτιστο μοντέλο επιλέχθηκε σύμφωνα με το μικρότερο σφάλμα επικύρωσης.

Ο παρακάτω πίνακας αναπαριστά τα καλύτερα μοντέλα του νευρωνικού δικτύου του υβριδικού μοντέλου βάσει του σφάλματος επικύρωσης καθώς και ο αριθμός των νευρώνων στο κρυφό επίπεδο και τις συναρτήσεις μεταφοράς του κρυφού επιπέδου και του επιπέδου εξόδου.

Πίνακας 8.7 Πίνακας σύγκρισης των καλύτερων μοντέλων του νευρωνικού δικτύου του υβριδικού μοντέλου βάσει του σφάλματος επικύρωσης.

MONTEΛΑ	Αλγόριθμος Εκπαίδευσης	Αριθμός νευρώνων στο κρυφό επίπεδο	Συνάρτηση Μεταφοράς Κρυφού Επιπέδου	Συνάρτηση Μεταφοράς Επιπέδου Εξόδου	Σφάλμα % Επικύρωσης
H_NEURAL1	SCG	15	Λογιστική – σιγμοειδής	Softmax	92.3
H_NEURAL2	SCG	23	Εφαπτομενική – σιγμοειδής	Softmax	91.2
H_NEURAL3	SCG	38	Εφαπτομενική – σιγμοειδής	Softmax	89.5
H_NEURAL4	SCG	42	Εφαπτομενική – σιγμοειδής	Softmax	92.4
H_NEURAL5	SCG	56	Εφαπτομενική – σιγμοειδής	Softmax	90.7
H_NEURAL6	SCG	67	Εφαπτομενική – σιγμοειδής	Softmax	89.3
H_NEURAL7	SCG	72	Εφαπτομενική – σιγμοειδής	Softmax	93.6
H_NEURAL8	SCG	78	Εφαπτομενική – σιγμοειδής	Softmax	94.1
H_NEURAL9	SCG	85	Εφαπτομενική – σιγμοειδής	Softmax	96.7
H_NEURAL10	SCG	95	Εφαπτομενική – σιγμοειδής	Softmax	92.7

Στη συνέχεια, το βέλτιστο μοντέλο της υβριδικής μεθοδολογίας H_NEURAL9 δοκιμάστηκε στο σύνολο δοκιμής και η ακρίβεια πρόβλεψης βρέθηκε 96.7% η οποία ήταν η καλύτερη από όλα τα άλλα μοντέλα που αναπτύχθηκαν. Στην εικόνα 9.7 αναπαρίστανται η τοπολογία του βέλτιστου νευρωνικού μοντέλου.

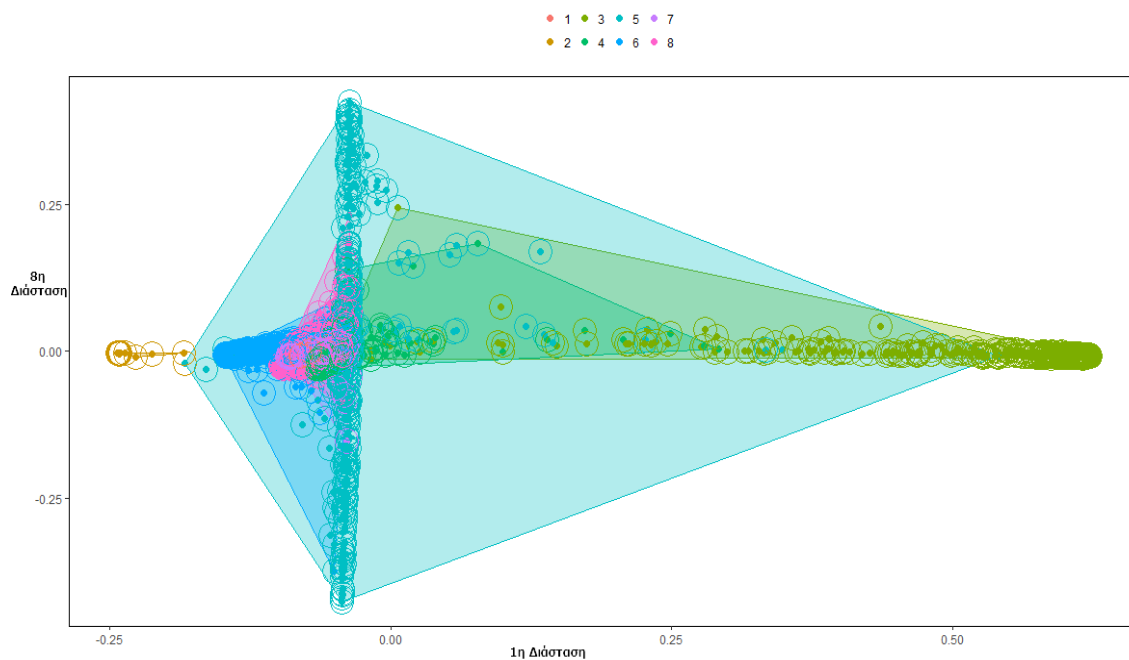


Εικόνα 8.7 Διάγραμμα που αναπαριστά την αρχιτεκτονική του βέλτιστου τεχνητού νευρωνικού δικτύου ταξινόμησης που αποτελείται από 85 νευρώνες στο κρυφό επίπεδο και την εφαπτομενική – σιγμοειδής ως συνάρτηση μεταφοράς, ενώ στο επίπεδο εξόδου υλοποιεί την Softmax ως συνάρτηση μεταφοράς.

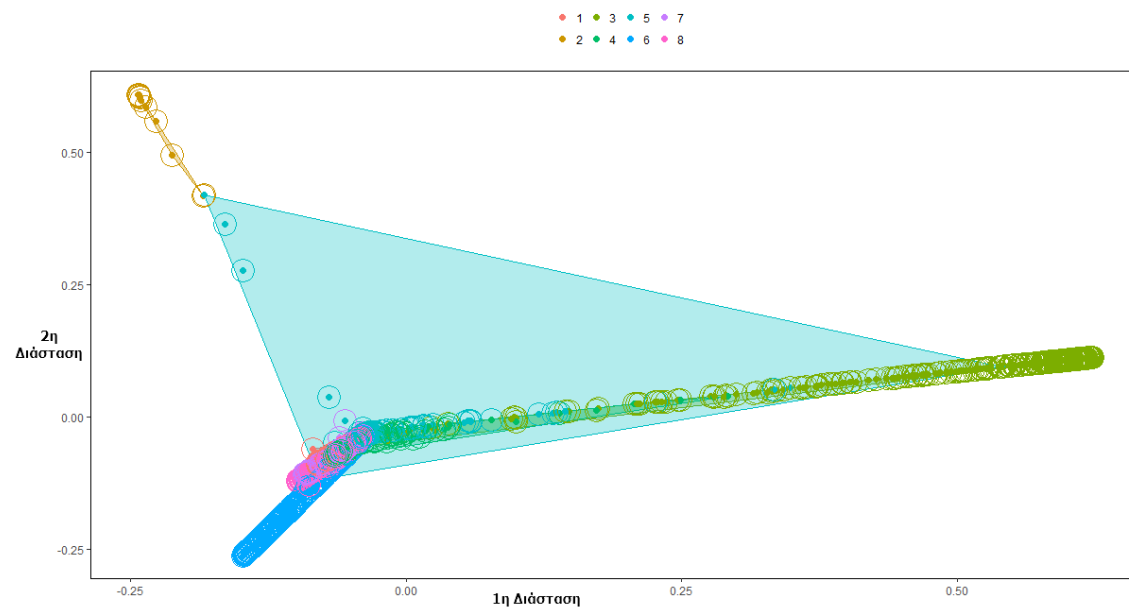
8.7.3 Αναπαράσταση ομαδοποίησης

Θα αναπαραστήσουμε γραφικά τον διαχωρισμό των τάξεων βάσει του βέλτιστου μοντέλου που ήταν το νευρωνικό δίκτυο με είσοδο τον πίνακα εγγύτητας. Επειδή υπάρχουν 8 μεταβλητές εισόδου, η εγγύτητα των σημείων για την ομαδοποίηση υπολογίζεται με βάση τον 8-διάστατο χώρο.

Επειδή γραφικά δεν μπορεί να απεικονιστεί ο 8-διάστατος χώρος, ενδεικτικά στις επόμενες εικόνες απεικονίζονται δισδιάστατα διαγράμματα μεταξύ συγκεκριμένων διαστάσεων όπου απεικονίζεται πιο εμφανώς ο διαχωρισμός των ομάδων με διαφορετικό χρώμα.

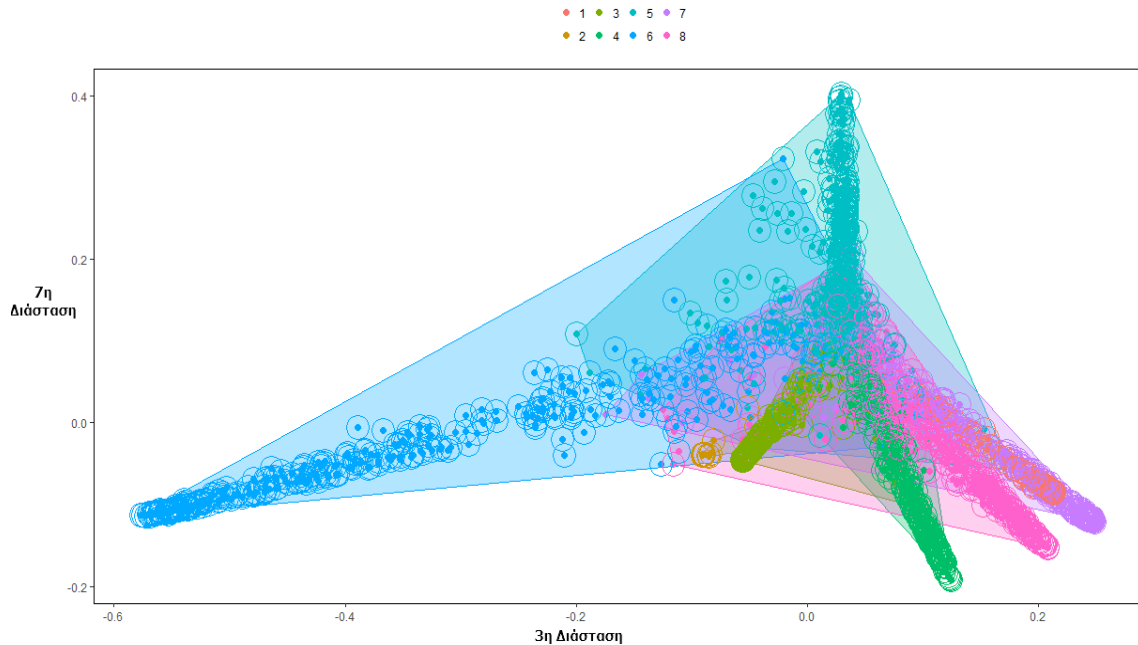


Εικόνα 8.8 Διάγραμμα μεταξύ της 1ης και της 8ης διάστασης που αναπαριστά τις ομάδες με βάση την εγγύτητα φύλλων στις 8 διαστάσεις.



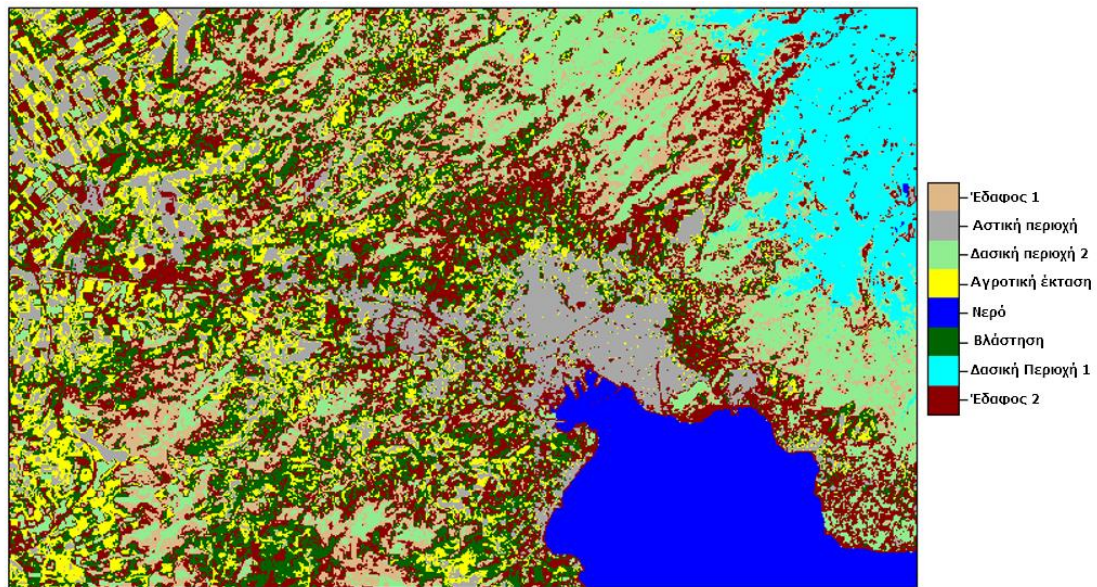
Εικόνα 8.9 Διάγραμμα μεταξύ της 1ης και της 2ης διάστασης που αναπαριστά τις ομάδες με βάση την εγγύτητα φύλλων στις 8 διαστάσεις.

Όπως παρατηρούμε στην επόμενη εικόνα, όταν αναπαραστήσουμε τις ομάδες του πίνακα εγγύτητας σε δισδιάστατη απεικόνιση των δυο πιο σημαντικών μεταβλητών της 3^{ης} και 7^{ης} σύμφωνα με το διάγραμμα μείωσης του δείκτη Gini impurity ο διαχωρισμός των ομάδων παρατηρείται εμφανέστερα.



Εικόνα 8.10 Διάγραμμα μεταξύ της 3ης και της 7ης διάστασης που αναπαριστά τις ομάδες στις 8 διαστάσεις.

Στην εικόνα 9.10, απεικονίζεται η τελική ταξινόμηση της δορυφορικής εικόνας σύμφωνα με το βέλτιστο υβριδικό μοντέλο.



Εικόνα 8.11 Ταξινόμηση της δορυφορικής εικόνας σύμφωνα με το βέλτιστο υβριδικό μοντέλο.

8.8 Μετρικές αξιολόγησης ταξινόμησης αλγορίθμων τεχνητής νοημοσύνης

Στην ενότητα αυτή, αναφέρονται οι μετρικές αξιολόγησης που χρησιμοποιήθηκαν για την μέτρηση της επίδοσης αλγορίθμων τεχνητής νοημοσύνης στην ταξινόμηση της δορυφορικής εικόνας.

8.8.1 Ακρίβεια (precision)

Η μετρική της ακρίβειας (precision) εκτιμά το κατά πόσο είναι ακριβείς οι προβλέψεις του αλγορίθμου στην ταξινόμηση. Αν συμβολίσουμε ως TP (True Positive) τις αληθώς θετικές περιπτώσεις και FP (False Positive) τις ψευδώς θετικές περιπτώσεις, τότε η ακρίβεια ορίζεται σαν:

$$\text{Ακρίβεια} = \frac{TP}{TP + FP} \quad (8.12)$$

8.8.2 Ανάκληση (Recall)

Η μετρική της ανάκλησης (Recall) που αναφέρεται και ως ευαισθησία (sensitivity) αξιολογεί την πληρότητα των αποτελεσμάτων μιας ταξινόμησης. Αν συμβολίσουμε ως TP (True Positive) τις αληθώς θετικές περιπτώσεις και FN (False Negative) τις ψευδώς αρνητικές περιπτώσεις, τότε η Ανάκληση (Recall) υπολογίζεται από την ακόλουθη εξίσωση:

$$\text{Ανάκληση} = \frac{TP}{TP + FN} \quad (8.13)$$

8.8.3 Ειδίκευση (Specificity)

Η μετρική της ειδίκευσης (specificity) μετρά την ικανότητα μίας τεχνικής να ταξινομήσει σωστά ότι τα δείγματα δεν ανήκουν σε μία κλάση (Parikh et al., 2008). Αν συμβολίσουμε ως TN (True Negative) τις αληθώς αρνητικές περιπτώσεις και FP (False Positive) τις ψευδώς θετικές περιπτώσεις, τότε η Ειδίκευση υπολογίζεται από την ακόλουθη εξίσωση:

$$\text{Ειδίκευση} = \frac{TN}{FP + TN} \quad (8.14)$$

8.8.4 F-measure

Η μετρική F-Measure αξιολογεί την επίδοση μιας ταξινόμησης, συνδυάζοντας την ακρίβεια και την ανάκληση (Parikh et al., 2008). Η μετρική F-Measure αποτελεί τον αρμονικό μέσο όρο της ακρίβειας και της ανάκλησης, και υπολογίζεται από την ακόλουθη εξίσωση:

$$F\ measure = 2 * \frac{Ακρίβεια * Ανάκληση}{Ακρίβεια + Ανάκληση} \quad (8.15)$$

Από τις εξισώσεις (8.12), (8.13) και (8.14) προκύπτει η εξίσωση:

$$F\ measure = 2 * \frac{TP}{2 * TP + FP + FN} \quad (8.16)$$

8.8.5 Συντελεστής Kappa

Ο συντελεστής Kappa μετρά την συμφωνία μεταξύ των ταξινομηθέντων και πραγματικών τιμών. Η τιμή 1 του δείκτη εκφράζει τέλεια συμφωνία μεταξύ των ταξινομηθέντων και πραγματικών τιμών. Ο Kappa coefficient υπολογίζεται από τον κάτωθι μαθηματικό τύπο:

$$K = \frac{N \sum_{i=1}^n m_{i,i} \sum_{i=1}^n (G_i C_i)}{N^2 - \sum_{i=1}^n (G_i C_i)} \quad (8.17)$$

Όπου:

- i συμβολίζει τον αριθμό της τάξης.
- N συμβολίζει τον συνολικό αριθμό των ταξινομημένων τιμών των τάξεων.
- $m_{i,i}$ αναπαριστά τον αριθμό των τιμών που ανήκουν στην πραγματική τάξη i και έχουν ταξινομηθεί στην τάξη i .
- C_i εκφράζει τον συνολικό αριθμό των προβλεφθέντων τιμών που ανήκουν στην τάξη i .
- G_i εκφράζει τον συνολικό αριθμό των πραγματικών τιμών που ανήκουν στην τάξη i .

8.9 Σύγκριση μοντέλων με τις μετρικές ταξινόμησης – Συζήτηση

Για την αξιολόγηση και σύγκριση των μοντέλων κατηγοριοποίησης που εφαρμόστηκαν, χρησιμοποιήθηκαν οι μετρικές της ακρίβειας, της ανάκλησης, της ειδικότητας και του μέτρου F-measure. Η αξιολόγηση των μοντέλων κατηγοριοποίησης βασίζεται στους αριθμούς των περιπτώσεων του συνόλου δοκιμής που προβλέφθηκαν σωστά από το κάθε μοντέλο για κάθε κλάση. Οι αριθμοί αυτοί αναπαρίστανται σε έναν πίνακα σύγχυσης (confusion matrix). Οι γραμμές του πίνακα σύγχυσης αντιπροσωπεύουν τις προβλεφθείσες κλάσεις εξόδου του κάθε μοντέλου, και οι στήλες τις πραγματικές τιμές. Ο πίνακας σύγχυσης στην ταξινόμηση της δορυφορικής εικόνας με το βέλτιστο προτεινόμενο υβριδικό μοντέλο νευρωνικού δικτύου παρατίθεται παρακάτω.

Πίνακας 8.8 Πίνακας σύγχυσης στην ταξινόμηση της δορυφορικής εικόνας με το βέλτιστο προτεινόμενο υβριδικό μοντέλο νευρωνικού δικτύου.

		Κλάσεις Ταξινόμησης							
Προβλέψεις Κλάσεων		1	2	3	4	5	6	7	8
	1	19	0	0	0	0	0	0	0
	2	0	29	1	0	1	0	0	0
	3	0	0	49	0	0	0	0	0
	4	0	0	1	31	1	0	0	0
	5	0	1	0	0	41	1	0	0
	6	0	0	0	0	1	70	1	0
	7	0	0	0	0	0	0	38	2
	8	0	0	0	0	0	0	0	15

Οι μετρικές για τα βέλτιστα μοντέλα ταξινόμησης για κάθε μεθοδολογία μηχανικής μάθησης αναπαρίστανται στον επόμενο πίνακα.

Πίνακας 8.9 Πίνακας μετρικών για την σύγκριση των μοντέλων τεχνητής νοημοσύνης.

Μετρικές	1° Προτεινόμενο (Βέλτιστο) Υβριδικό Μοντέλο (Νευρωνικού Δικτύου – Τυχαίων Δασών)	2° Προτεινόμενο Υβριδικό Μοντέλο (Μηχανών Διανυσμάτων Υποστήριξης – Τυχαίων Δασών)	3° Μοντέλο Βέλτιστο Μοντέλο Τυχαίων Δασών Δέντρων Απόφασης	4° Μοντέλο Βέλτιστο Μοντέλο Μηχανών Διανυσμάτων Υποστήριξης
Ακρίβεια	0.9669	0.9519	0.9057	0.9267
Ανάκληση	0.9669	0.9519	0.9057	0.9267
Ειδίκευση	0.9953	0.9987	0.9867	0.9944
F-measure	0.9669	0.9519	0.9057	0.9267
Kappa	0.961	0.951	0.886	0.917

Όπως παρατηρούμε από τον πίνακα των μετρικών στις μετρικές: ακρίβεια, ανάκληση, F-measure, το πρώτο προτεινόμενο υβριδικό μοντέλο Νευρωνικού Δικτύου – Τυχαίων Δασών εμφανίζει τις βέλτιστες μετρικές, σε σχέση με τα άλλα μοντέλα, ενώ το δεύτερο προτεινόμενο υβριδικό μοντέλο (Μηχανών Διανυσμάτων Υποστήριξης – Τυχαίων Δασών) εμφανίζει καλύτερες μετρικές σε σχέση με τα 3° και 4° συμβατικά μοντέλα των Τυχαίων Δασών Δέντρων Απόφασης και των Μηχανών Διανυσμάτων Υποστήριξης αντιστοίχως. Όσον αφορά την μετρική της ειδίκευσης (specificity), η οποία μετρά την ικανότητα μίας τεχνικής να ταξινομήσει σωστά ότι τα δείγματα δεν ανήκουν σε μία κλάση (Parikh et al., 2008), τα 1° και 2° προτεινόμενα μοντέλα υπερέχουν ελαφρά αφού τα μοντέλα εμφανίζουν παρόμοιες τιμές.

Επίσης, σημειώνεται ότι για την ανάπτυξη του συμβατικού μοντέλου ταξινόμησης με Τυχαία Δάση, δεν είναι αναγκαίος ο υπολογισμός του πίνακα εγγύτητας, ούτε του δείκτη Gini. Οι παράγοντες αυτοί δεν επηρεάζουν την απόδοση του αλγορίθμου των

Τυχαίων Δασών, υπολογίστηκαν μόνο για να χρησιμοποιηθούν κατά την ανάπτυξη των προτεινόμενων υβριδικών μοντέλων ώστε να βελτιώσουν την προβλεψιμότητα.

8.10 Συμπεράσματα

Διάφορες έρευνες έχουν δείξει την αποτελεσματικότητα των Μηχανών Διανυσμάτων Υποστήριξης, των νευρωνικών δικτύων αλλά και των Τυχαίων Δασών και στην ταξινόμηση δορυφορικών εικόνων (Huang et al., 2002; Rodriguez-Galiano et al., 2012) και για το λόγο αυτό επιλέχθηκαν στην παρούσα έρευνα ώστε να συγκριθούν, αλλά και για να επιχειρηθεί η δόμηση ενός υβριδικού μοντέλου που θα βελτιώνει την ακρίβεια της ταξινόμησης. Τα αποτελέσματα της ταξινόμησης του υβριδικού μοντέλου δείχνουν βελτιωμένη ακρίβεια στην ταξινόμηση της δορυφορικής εικόνας συγκρινόμενα με άλλες μελέτες ταξινόμησης δορυφορικών εικόνων (Civco, 1993; Huang et al., 2002; Rodriguez-Galiano et al., 2012). Το πλεονέκτημα του υβριδικού μοντέλου είναι ότι αξιοποιεί τα πλεονεκτήματα επιτυχημένων αλγορίθμων τεχνητής νοημοσύνης στην ταξινόμηση της δορυφορικής εικόνας για την βελτιστοποίηση των προβλέψεων.

Η ταξινόμηση των δορυφορικών εικόνων έχει πλείστες εφαρμογές στη δημόσια διοίκηση, όπως στην χαρτογράφηση της βλάστησης με χρήση δεικτών όπως για παράδειγμα ο NDVI, στην αποτύπωση των δασικών περιοχών, στην εκτίμηση διαχρονικά της αστικής ανάπτυξης (Whitlock et al., 2003).

ΚΕΦΑΛΑΙΟ 9: Συμπεράσματα - Συζήτηση

9.1 Συμπεράσματα - Συζήτηση

Στην παρούσα διατριβή, προτείνεται μία νέα μεθοδολογία διοίκησης δημοσίων υπηρεσιών, βασισμένη σε βελτιστοποιημένες μεθόδους τεχνητής νοημοσύνης, σε συνδυασμό με μεθόδους γεωγραφικών συστημάτων πληροφοριών και τηλεπισκόπησης. Επίσης, στην παρούσα διατριβή, προτείνονται νέες υβριδικές τεχνικές παλινδρόμησης αλλά και ταξινόμησης που αναπτύχθηκαν, βασισμένες σε τροποποιημένους αλγορίθμους τεχνητής νοημοσύνης, οι οποίες παρήγαγαν αποτελέσματα βελτιωμένης ακρίβειας σε σύγκριση με τους συμβατικούς αλγορίθμους μηχανικής μάθησης που δοκιμάστηκαν. Επιπλέον, χρησιμοποιήθηκαν τεχνικές βελτιστοποίησης που ανήκουν στον κλάδο της αυτοματοποιημένης μηχανικής μάθησης AutoML (Automated Machine Learning), (Feurer et al., 2015; Guyon et al., 2016; Mendoza et al., 2016), όπως ο αλγόριθμος βελτιστοποίησης Bayes, και ο αλγόριθμος Βελτιστοποίησης Σωματιδίων Σμήνους. Στην αυτοματοποιημένη μηχανική μάθηση AutoML, οι μηχανές μαθαίνουν και βελτιστοποιούνται από μόνες τους, το οποίο είναι πολύ σημαντικό, ιδίως στην περίπτωση της επέκτασης της προτεινόμενης μεθοδολογίας στην παρούσα διατριβή, σε ένα αυτοματοποιημένο σύστημα τεχνητής νοημοσύνης στην δημόσια διοίκηση.

Οι μεθοδολογίες που εφαρμόστηκαν αφορούσαν πολλαπλούς τομείς της δημόσιας και δικαστικής διοίκησης και συγκεκριμένα προβλέψεις παραγόντων σημαντικών στην δημόσια και δικαστική διοίκηση όπως: ποσοτική πρόβλεψη της εγκληματικότητας, χωρική πρόβλεψη της εγκληματικότητας με την χρήση των γεωγραφικών συστημάτων πληροφοριών σε συνδυασμό με μεθόδους τεχνητής νοημοσύνης, πρόβλεψη περιβαλλοντικών δεικτών (π.χ. όζον), προβλέψεις κοινωνικοοικονομικών παραγόντων όπως η ανεργία.

Το σύστημα που προτείνεται για την δημόσια διοίκηση, βασισμένο στη τεχνητή νοημοσύνη, εντάσσεται στην λεγόμενη «προβλεπτική διοίκηση» (predictive

management) που έχει εμφανιστεί ως όρος στην βιβλιογραφία τα τελευταία χρόνια (Kim et al., 2012; Wang et al., 2018), αλλά δεν έχει ακόμα διερευνηθεί πλήρως ως προς τον τρόπο λειτουργίας και εφαρμογής της, ενώ ο αριθμός των βιβλιογραφικών αναφορών είναι πολύ περιορισμένος και επιπροσθέτως, δεν έχει μελετηθεί ακόμα στον τομέα της δημόσιας και δικαστικής διοίκησης. Συγκεκριμένα, σε μία σχετική μελέτη, οι Garnier et al. (2014), ερεύνησαν μία υπολογιστική τεχνική χαμηλού κόστους για το προβλεπτικό μάνατζμεντ της θέρμανσης σε κτίρια. Επίσης, σε μία άλλη έρευνα, οι Kim et al. (2012), μελέτησαν την εφαρμογή της μηχανικής μάθησης στο προβλεπτικό μάνατζμεντ της βιομάζας από φυτοπλαγκτόν, χρησιμοποιώντας έναν γενετικό αλγόριθμο βασισμένο στα αναδρομικά ή ανατροφοδοτούμενα τεχνητά νευρωνικά δίκτυα (Recurrent Neural Networks). Επίσης, ο Houry, (2018), μελέτησε την δημιουργία ενός προβλεπτικού εργαλείου για την χρήση του στο προβλεπτικό μάνατζμεντ σπουδαστών που εγκαταλείπουν τις σπουδές τους σε εξ αποστάσεως προπτυχιακά προγράμματα.

Στην παρούσα έρευνα, μελετήθηκε η χρήση της τεχνητής νοημοσύνης στη δημιουργία μοντέλων πρόβλεψης παραγόντων σημαντικών στη δημόσια διοίκηση, σε περιβαλλοντικό και κοινωνικο-οικονομικό επίπεδο όπως η εγκληματικότητα (σημαντική στη λήψη αποφάσεων σε δικαστικές και αστυνομικές αρχές), περιβαλλοντικοί παράγοντες ρύπανσης (όπως πχ NO₂, O₃, PM₁₀), κοινωνικοοικονομικοί παράγοντες όπως η ανεργία. Η πρόβλεψη περιβαλλοντικών παραγόντων ρύπανσης όπως για παράδειγμα του διοξειδίου του αζώτου NO₂, του όζοντος O₃, και του PM₁₀, βοηθάει στην προληπτική διοίκηση, με την κατάλληλη λήψη μέτρων από τις δημόσιες αρχές για την προστασία του περιβάλλοντος και των πολιτών.

Οι προβλέψεις πραγματοποιήθηκαν με την χρήση βελτιστοποιημένων νευρωνικών δικτύων, και άλλων τεχνικών τεχνητής νοημοσύνης, με την δημιουργία, χρήση και αξιολόγηση των μοντέλων πρόβλεψης και την τελική επιλογή του βέλτιστου μοντέλου πρόβλεψης. Πραγματοποιήθηκε η μελέτη της χρήσης διαφορετικών αλγορίθμων εκπαίδευσης των νευρωνικών δικτύων και διαφορετικών συναρτήσεων ενεργοποίησης των κρυφών επιπέδων. Τα αποτελέσματα έδειξαν αρκετά καλύτερη ακρίβεια πρόβλεψης συγκρινόμενα με άλλες μελέτες (Franses et al. 2004; Graf-Jaccottet and Jaunin, 1998; Olmedo 2014).

Επίσης, στην παρούσα έρευνα, ερευνήθηκε η χρήση των γεωγραφικών συστημάτων πληροφοριών σε συνδυασμό με μεθόδους τεχνητής νοημοσύνης για την χωρική πρόβλεψη της εγκληματικότητας. Πραγματοποιήθηκε η δημιουργία τεχνητών νευρωνικών δικτύων με την κατάλληλη τοπολογία, ώστε να είναι δυνατή η πρόβλεψη των περιοχών υψηλής συγκέντρωσης της εγκληματικότητας (θερμών σημείων). Διερευνήθηκε η δημιουργία μοντέλων νευρωνικών δικτύων με την χρήση τεχνικών αυτόματης μηχανικής μάθησης AutoML για την βελτιστοποίησή τους, αλλά και ερευνήθηκε ο προσδιορισμός των κατάλληλων υπερπαραμέτρων με την μέθοδο αναζήτησης πλέγματος: η τοπολογία των νευρώνων, οι συναρτήσεις ενεργοποίησης, οι αλγόριθμοι εκπαίδευσης, η δομή τους (πολυεπίπεδοι αισθητήρες, κλπ), για την βελτιστοποίηση της προβλεψιμότητας των νευρωνικών δικτύων.

Επιπλέον, στην παρούσα διατριβή, χρησιμοποιήθηκαν και ταξινομήθηκαν δεδομένα τηλεπισκόπησης και συγκεκριμένα δεδομένα εικονοστοιχείων δορυφορικής εικόνας. Στην παρούσα έρευνα, για την ταξινόμηση της πολυφασματικής δορυφορικής εικόνας, επιλέχθηκαν οι αλγόριθμοι των Τεχνητών Νευρωνικών Δικτύων, των Μηχανών Διανυσμάτων Υποστήριξης και των Τυχαίων δασών δεδομένου ότι οι πιο συνηθισμένες μεθοδολογίες ταξινόμησης, όπως ο ταξινομητής της μέγιστης πιθανοφάνειας (Maximum Likelihood Classifier), πρέπει να ακολουθούν την κανονική κατανομή (Strahler, 1980), που είναι δύσκολο σε προβλήματα τηλεπισκόπησης, ιδίως στην ταξινόμηση πολυφασματικών δορυφορικών εικόνων. Επιπλέον, σύμφωνα με την βιβλιογραφία στους αλγόριθμους αυτούς επιδρά ελάχιστα έως καθόλου το φαινόμενο Hughes, που αλλιώς θα μείωνε την ακρίβεια ταξινόμησης λόγω της ύπαρξης των πολλών διαστάσεων (Alonso et al., 2001; Hughes, 1968; Watanachaturaporn et al., 2004). Αναπτύχθηκαν τα μοντέλα τεχνητής νοημοσύνης, ώστε ερευνώντας διάφορες τοπολογίες νευρώνων, κρυφών επιπέδων, συναρτήσεων ενεργοποίησης (νευρωνικά δίκτυα), συναρτήσεων πυρήνα (Μηχανές Διανυσμάτων Υποστήριξης), αριθμός δέντρων απόφασης (Τυχαία Δάση Δέντρων Απόφασης), να δομηθεί το βέλτιστο μοντέλο πρόβλεψης. Επίσης, αναπτύχθηκε μία νέα υβριδική μεθοδολογία ταξινόμησης που έδειξε καλύτερα αποτελέσματα ταξινόμησης, σύμφωνα με τις μετρικές που αξιολογήθηκαν τα μοντέλα.

Στον επόμενο πίνακα, παρουσιάζονται συνοπτικά, οι μέθοδοι που μελετήθηκαν στην παρούσα διατριβή και σε ποιες δημόσιες υπηρεσίες μπορούν να αξιοποιηθούν.

Πίνακας 9.1 Πίνακας σύνοψης των μεθόδων τεχνητής νοημοσύνης που μελετήθηκαν και σε ποιες δημόσιες υπηρεσίες μπορούν να αξιοποιηθούν.

Παράγοντας	Μέθοδοι	Αρμόδιες Δημόσιες Υπηρεσίες
Εγκληματικότητα	Τεχνητά Νευρωνικά Δίκτυα, Νοημοσύνη Σμήνους Σωματιδίων, Γεωγραφικά Συστήματα Πληροφοριών, Μηχανές Διανυσμάτων Υποστήριξης	Αστυνομικές Αρχές, Δικαστικές Αρχές, Εκπαιδευτικοί Οργανισμοί, Ασφάλεια Συγκοινωνιών
Περιβαλλοντικοί παράγοντες	Τεχνητά Νευρωνικά Δίκτυα, Μηχανές Διανυσμάτων Υποστήριξης, Νοημοσύνη Σμήνους Σωματιδίων	Περιβαλλοντικές Αρχές, Περιβαλλοντικοί Οργανισμοί
Κοινωνικοοικονομικοί παράγοντες	Τεχνητά Νευρωνικά Δίκτυα, Μηχανές Διανυσμάτων Υποστήριξης, Νοημοσύνη Σμήνους Σωματιδίων	Υπηρεσίες Κοινωνικο- οικονομικού σχεδιασμού
Κάλυψη γης - ταξινόμηση πολυφασματικών δορυφορικών εικόνων	Τηλεπισκόπηση, Μηχανές Διανυσμάτων Υποστήριξης, Τυχαία Δάση Δέντρων Απόφασης, Τεχνητά Νευρωνικά Δίκτυα	Αξιολόγηση αστικών δομών, πρασίνου, χρήσεων γης

Το πλεονέκτημα της προτεινόμενης μεθοδολογίας είναι ότι τα χρησιμοποιούμενα μοντέλα βασισμένα στην Τεχνητή Νοημοσύνη μπορούν να μοντελοποιήσουν και μη γραμμικές σχέσεις εισόδου – εξόδου (Svozil et al., 1997). Είναι μη παραμετρικά μοντέλα, ενώ οι περισσότερες στατιστικές μέθοδοι χρειάζονται υψηλότερο βαθμό στατιστικής ανάλυσης.

Τα νευρωνικά δίκτυα χρησιμοποιούνται ευρέως στην επίλυση διαφόρων προβλημάτων ταξινόμησης και παλινδρόμησης μπορούν να ανιχνεύσουν έμμεσα πολύπλοκες μη γραμμικές σχέσεις μεταξύ ανεξάρτητων και εξαρτημένων μεταβλητών (Tu, 1996). Τα μειονεκτήματα των νευρωνικών δικτύων και γενικά των τεχνικών μηχανικής μάθησης αφορούν την αυξημένη απαίτηση υπολογιστικών πόρων και το ότι είναι επιρρεπή στο φαινόμενο της υπερ-εκπαίδευσης, το οποίο όμως μπορεί να αντιμετωπιστεί κατάλληλα, όπως στην παρούσα διατριβή με την χρήση διαφόρων μεθόδων επικύρωσης όπως η διασταυρωμένη επικύρωση K-φύλλων.

Επίσης, χρησιμοποιούνται πολλαπλές μέθοδοι τεχνητής νοημοσύνης συνδυαζόμενες με μεθόδους γεωγραφικών συστημάτων πληροφοριών και τηλεπισκόπησης στη δημόσια διοίκηση. Επιπλέον, πραγματοποιήθηκε χαρτογράφηση των κυριότερων παραγόντων που επηρεάζουν την δικαστική και δημόσια διοίκηση και στην συνέχεια εφαρμογή πολλαπλών μεθοδολογιών για την δημιουργία μοντέλων ποσοτικής αλλά και χωρικής τους πρόβλεψης.

Στον επόμενο πίνακα, παρουσιάζονται οι τομείς της δημόσιας διοίκησης, τα είδη αποφάσεων, όπου μπορεί να εφαρμοστεί η προτεινόμενη μεθοδολογία, σύμφωνα με την προβλεπτική διοίκηση, καθώς και οι σχετικές βιβλιογραφικές αναφορές.

Πίνακας 9.2. Οι τομείς της δημόσιας διοίκησης, τα είδη αποφάσεων, όπου μπορεί να εφαρμοστεί η προτεινόμενη μεθοδολογία και η σχετική βιβλιογραφία.

Τομέας Δημόσιας Διοίκησης	Είδος Απόφασης/Λήψη Μέτρων	Βιβλιογραφικές Αναφορές
Δικαστικές και Αστυνομικές Αρχές	Αποφάσεις επιχειρησιακές κατά της εγκληματικότητας, σχεδιασμός περιπολιών αστυνομικών δυνάμεων αποφάσεις βέλτιστης χωρικής κατανομής αστυνομικών δυνάμεων και προβλεπτικής αστυνόμευσης	Camacho - Collados and Liberatore (2015), Camacho - Collados et al. (2015), Tambe et al. (2016), Perry (2013)

Πίνακας 9.2. (συνέχεια) Οι τομείς της δημόσιας διοίκησης, τα είδη αποφάσεων, όπου που μπορεί να εφαρμοστεί η προτεινόμενη μεθοδολογία και η σχετική βιβλιογραφία.

Τομέας Δημόσιας Διοίκησης	Είδος Απόφασης/Λήψη Μέτρων	Βιβλιογραφικές Αναφορές
Τουρισμός, Συγκοινωνίες	Λήψη μέτρων πρόληψης εγκληματικότητας σε τουριστικές περιοχές, σε περιοχές συγκοινωνιών με αυξημένη πληθυσμιακή πυκνότητα, αποφάσεις εκτίμησης ρίσκου περιοχών	Ferreira and Harmse (2000), Jenkins (2001), Montolio and Planells - Struse, (2016), Kouziokas (2017c)
Εκπαίδευση	Λήψη αποφάσεων προστασίας των εκπαιδευτικών μονάδων και των μαθητών από φαινόμενα εγκληματικότητας	Hemphill et al. (2015), Holt et al. (2016)
Διοίκηση με τη Συμμετοχή Πολιτών	Περιπολίες πολιτών βάσει των προβλέψεων εγκληματικότητας	Junior et al. (2017), Zimring, (2011)
Αστική Διοίκηση - Έξυπνες Πόλεις	Εκτίμηση και μοντελοποίηση χωρικού ρίσκου στην πόλη, αποφάσεις δημιουργίας συστημάτων προειδοποίησης για τους πολίτες	Caplan and Kennedy (2016), Gorr and Lee (2015)
Κεντρική Δημόσια Διοίκηση	Λήψη αποφάσεων για τον περιορισμό εγκληματικότητας, εφαρμογή πολιτικών e-government, ψήφιση νόμων, υιοθέτηση πολιτικών αποφάσεων και νομοθέτηση σχετικών διατάξεων	Ku and Leroy (2014), Stolz (2015), Worrall (2014)

Επίσης, το πλεονέκτημα της προτεινόμενης μεθοδολογίας που αναπτύχθηκε σε σύγκριση με άλλες έρευνες (Caplan and Kennedy, 2016; Chainey and Ratcliffe, 2013; Gorr and Lee, 2015; Ki, 2013; Ku and Leroy, 2014; Leipnik and Albert, 2003; Palocsay et al., 2000), είναι ότι οι δημόσιοι φορείς μπορούν να λάβουν υπόψη και να αξιοποιήσουν

προβλέψεις, χωρικές ή μη, παραγόντων σημαντικών στη λήψη αποφάσεων, στην δημόσια και δικαστική διοίκηση, αλλά και στον αστικό χωρικό σχεδιασμό και το περιβαλλοντικό μάνατζμεντ, αφού βασίζεται στην έννοια της προβλεπτικής διοίκησης, δηλαδή την αντιμετώπιση σήμερα, των προβλημάτων του αύριο.

Η πρόβλεψη αφορά το επόμενο χρονικό βήμα. Δηλαδή, με δεδομένα παραγόντων εισόδου του ‘σήμερα’, που επηρεάζουν την τιμή της εξόδου (εξαρτημένη μεταβλητή), μπορεί να προβλεφθεί η τιμή του ‘αύριο’ της εξαρτημένης μεταβλητής. Το στοιχείο αυτό, μπορεί να βοηθήσει αρκετά τις δημόσιες και δικαστικές αρχές, αφού θα έχουν το χρόνο μέχρι να επέλθει το επόμενο χρονικό βήμα, για να λάβουν τα κατάλληλα προληπτικά μέτρα.

Όσον αφορά τους δυνητικούς μελλοντικούς κινδύνους της εφαρμογής της πλήρως αυτοματοποιημένης τεχνητής νοημοσύνης στη δημόσια διοίκηση και ηλεκτρονική διακυβέρνηση, αυτοί επικεντρώνονται στον κίνδυνο αντικατάστασης του ανθρώπου από τις μηχανές, καθώς και στη δημιουργία της λεγόμενης «τεχνολογικής ανεργίας», η οποία είναι μία σημαντική επίπτωση από την αντικατάσταση του ανθρώπινου δυναμικού από ευφυή συστήματα. Ταυτόχρονα όμως θα δημιουργηθεί και η ανάγκη για νέες θέσεις και εξειδικεύσεις ανθρώπινου δυναμικού κατά την εφαρμογή της τεχνητής νοημοσύνης στη δημόσια διοίκηση.

Το ότι οι μηχανές και τα ευφυή συστήματα βασισμένα στην τεχνητή νοημοσύνη, έχουν ήδη εισβάλει σε διάφορες τομείς της δημόσιας διοίκησης, όπως στο δικαστικό σύστημα, αποδεινύεται και με το παράδειγμα της δημιουργίας του πρώτου ρομποτικού δικηγόρου – bot, που εφαρμόζεται ήδη σε πολλές πολιτείες των Ηνωμένων Πολιτειών Αμερικής (Dale 2016).

Επίσης, ένας άλλος μελλοντικός κίνδυνος από την υιοθέτηση ευφυών συστημάτων τεχνητής νοημοσύνης, αφορά τα τυχόν κενά στην ασφάλεια τους, που θα μπορούσαν να επιτρέψουν σε διάφορους κακόβουλους χρήστες να αποκτήσουν πρόσβαση στις λειτουργίες τους, και κατ’ επέκταση σε όλα τα δεδομένα που διαχειρίζονται, και έτσι να απειληθεί η δημόσια ασφάλεια και η ιδιωτικότητα των πολιτών. Τέτοιου είδους

προβλήματα μπορούν να αντιμετωπιστούν με διάφορους τρόπους, όπως η κατάλληλη συντήρησή των συστημάτων και η εφαρμογή προηγμένων πολιτικών ασφαλείας.

9.2 Περιορισμοί – Μελλοντικές Επεκτάσεις

Οι περιορισμοί της προτεινόμενης μεθοδολογίας αφορούν τα χωρικά δεδομένα των περιστατικών εγκληματικότητας που πρέπει να είναι διαθέσιμα για να διενεργηθούν οι χωρικές αναλύσεις, αλλά και τον αριθμό των περιστατικών εγκλημάτων δεδομένου ότι η ανάλυση θερμών σημείων για να εφαρμοστεί, απαιτεί έναν ελάχιστο αριθμό συμβάντων.

Επιπλέον, η κατανομή των παρατηρήσεων στο χώρο πρέπει να τείνει προς το ομαδοποιημένο χωρικό πρότυπο, ώστε να μπορούν να εντοπιστούν θερμά σημεία, με την ανάλυση θερμών σημείων.

Επίσης, ένας άλλος περιορισμός αφορά το πλήθος των παρατηρήσεων στο διάνυσμα εισόδου των μοντέλων τεχνητής νοημοσύνης, αφού πρέπει να υπάρχει ικανός αριθμός δεδομένων για να είναι δυνατή η εκπαίδευση και η βελτιστοποίηση των μοντέλων τεχνητής νοημοσύνης.

Οι περιορισμοί εφαρμογής της προτεινόμενης μεθοδολογίας στις συγκεκριμένες μελέτες περιπτώσεις, αφορούν το ότι η ανάπτυξη των μοντέλων τεχνητής νοημοσύνης με βελτιστοποιημένες τεχνικές AutoML πραγματοποιήθηκαν για τις συγκεκριμένες τοποθεσίες, το συγκεκριμένο χρονικό διάστημα και για τις συγκεκριμένες παραμέτρους εισόδου. Οι μελλοντικές προτάσεις για επέκταση της έρευνας αυτής, αφορούν την ανάπτυξη και τον έλεγχο παρόμοιων μοντέλων πρόβλεψης για τη διερεύνηση του βαθμού χωρικής ή ποσοτικής προβλεψιμότητας διάφορων τύπων παραγόντων διοίκησης σε άλλες πόλεις και περιοχές του κόσμου, εφαρμόζοντας την προτεινόμενη μεθοδολογία.

Όπως αναλυτικά αναφέρεται σε κάθε αντίστοιχο κεφάλαιο, επιλέχθηκαν οι συγκεκριμένες περιοχές για τις μελέτες περίπτωσης εφαρμογής της προτεινόμενης μεθοδολογίας, λόγω της διαθεσιμότητας μεγαλύτερου αριθμού δεδομένων σε βάθος χρόνου όλων των παραγόντων εισόδου που μελετήθηκαν, ώστε ο αριθμός των παρατηρήσεων να είναι ικανός για την εκπαίδευση και βελτιστοποίηση των μοντέλων τεχνητής νοημοσύνης.

Επίσης, όσον αφορά τις μελλοντικές επεκτάσεις της παρούσας έρευνας, όπως προαναφέρθηκε, η μεθοδολογία βασίστηκε σε τεχνικές AutoML, που σημαίνει ότι είναι πιο εύκολα επεκτάσιμη σε ένα αυτόματο σύστημα τεχνητής νοημοσύνης.

Επίσης, τα μοντέλα που δημιουργήθηκαν με το λογισμικό Matlab και μοντελοποιήθηκαν με το Simulink, μια επέκταση του Matlab, είναι επαναχρησιμοποιήσιμα (reusable), αντικειμενοστραφή, αλλά και διαλειτουργικά, αφού το Matlab μπορεί να συνεργάζεται με προγράμματα γραμμένα σε άλλες γλώσσες προγραμματισμού, όπως οι Java, python, C και C++ (Dabney and Harman, 2004).

Δεδομένου ότι έχει μοντελοποιηθεί η σχέση των παραμέτρων εισόδου με την παράμετρο εξόδου, και τα εκπαιδευμένα μοντέλα έχουν αποθηκευτεί, θα μπορούν να επαναχρησιμοποιηθούν ως εργαλείο διοίκησης, από τις αρμόδιες αρχές, για τις συγκεκριμένες μελέτες περίπτωσης. Για παράδειγμα, με την χρήση των μοντέλων, μπορεί να ερευνηθεί για συγκεκριμένες τιμές εισόδου, ποια θα είναι η έξοδος (εγκληματικότητα, όζον, ανεργία, κλπ) του επόμενου χρονικού βήματος, ώστε να ληφθούν οι κατάλληλες αποφάσεις των αρχών.

Μια άλλη μελλοντική επέκταση του συστήματος αφορά την δημιουργία ενός τεχνητού εγκεφάλου που θα «σκέφτεται» σύμφωνα με τα δομημένα μοντέλα τεχνητής νοημοσύνης και θα πραγματοποιεί σε πραγματικό χρόνο προβλέψεις του επόμενου χρονικού βήματος, παραγόντων σημαντικών στη δημόσια διοίκηση, και θα αλληλεπιδρά και θα ενημερώνει άλλα έξυπνα συστήματα, όπως: Ευφυή συστήματα ρύθμισης κυκλοφορίας (π.χ. για την μείωση της ατμοσφαιρικής ρύπανσης, κλπ.), συστήματα προειδοποίησης κινδύνου για τους πολίτες που θα αφορούν για παράδειγμα την εγκληματικότητα, παράγοντες ρύπανσης και άλλους παράγοντες, και συστήματα αστυνομικών και δικαστικών αρχών για την πρόβλεψη περιοχών αυξημένης έντασης της εγκληματικότητας.

ΚΕΦΑΛΑΙΟ 10: ΒΙΒΛΙΟΓΡΑΦΙΑ

10.1 Ελληνική βιβλιογραφία

Αποστολάκης, Ι., Λούκης, Ε. και Χάλαρης, Ι. (2008) *Η Ηλεκτρονική Δημόσια Διοίκηση-Οργάνωση, Τεχνολογία & εφαρμογές*. Αθήνα, Παπαζήση.

Αστάρας, Θ., Οικονομίδης, Δ. and Μουρατίδης, Α. (2011) *Ψηφιακή Χαρτογραφία και Γεωγραφικά Συστήματα Πληροφοριών*. Αθήνα: Δίσιγμα.

Βλαχάβας, Ι., Κεφαλός, Π., Βασιλειάδης, Ν., Κόκκορας, Φ. και Σακελλαρίου, Η. (2006) *Τεχνητή νοημοσύνη, Γ Έκδοση*. Γκιούρδας.

Γεωργούλη, Α. (2015) *Τεχνητή νοημοσύνη*. Αθήνα: Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών.

Διαμαντάρας, Κ. (2007) *Τεχνητά νευρωνικά δίκτυα. Κλειδάριθμος*. Αθήνα.

Ευελπίδου, Ν. και Αντωνίου, Β. (2015) *Γεωγραφικά συστήματα πληροφοριών*. Αθήνα: Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών.

Κάβουρας, Μ., Δάρρα, Α., Κονταξάκη, Σ., και Τομαή, Ε. (2016) *Επιστήμη Γεωγραφικής Πληροφορίας-Αρχές και Τεχνολογίες*. Αθήνα: Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών.

Καϊμάρης, Δ. (2014) *Τηλεπισκόπηση*. Αθήνα: Ζήτη.

- Καμπουρλάζος, Β. και Παπακώστας, Γ. (2015) *Εισαγωγή στην Υπολογιστική Νοημοσύνη*. Αθήνα: Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών.
- Καρτάλης, Κ. και Φείδας, Χ. (2013) *Αρχές και Εφαρμογές Δορυφορικής Τηλεπισκόπησης*. Εκδόσεις Τζιόλα.
- Κουτσόπουλος, Κ. (2005) *Γεωγραφικά συστήματα πληροφοριών και ανάλυση χώρου*. Δίσιγμα, Αθήνα.
- Κουτσόπουλος, Κ. και Ανδρουλακάκης, Ν. (2012) *Γεωγραφικά συστήματα πληροφοριών με ArcGis 10*. Αθήνα: Παπασωτηρίου.
- Μερτίκας, Σ. Π. (2006) *Τηλεπισκόπηση και Ψηφιακή Ανάλυση Εικόνας*. Εκδόσεις ΙΩΝ.
- Περάκης, Κ. (1999) *Σημειώσεις μαθήματος: Φωτοερμηνεία και τηλεπισκόπηση*, Βόλος: Πανεπιστημιακές Εκδόσεις Θεσσαλίας.
- Περάκης, Κ., Μωυσιάδης, Α. και Φαρασλής, Ι. (2015) *Η τηλεπισκόπηση σε 13 ενότητες*. Αθήνα: Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών.
- Σκιάνης, Γ., Νικολακόπουλος, Κ. και Βαϊόπουλος, Δ. (2012) *Τηλεπισκόπηση: Αρχές, Ψηφιακή Ανάλυση Εικόνας, Εφαρμογές*. Εκδοτικός Όμιλος Ίων.
- Τζελεπίδης, Π. (2001) *Νοήμονα Συστήματα*. Αθήνα: Σταμούλη ΑΕ.
- Χαλκιάς, Χ. και Γκούσια, Μ. (2015) *Γεωγραφική ανάλυση με την αξιοποίηση της γεωπληροφορικής*. Αθήνα: Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών.

10.2 Ξενόγλωσση βιβλιογραφία

Abadinsky, H. (2012) *Organized crime*. Cengage Learning.

Abdul-Wahab, S., Bouhamra, W., Ettouney, H., Sowerby, B. and Crittenden, B. D. (2000) 'Analysis of ozone pollution in the Shuaiba industrial area in Kuwait', *International journal of environmental studies*, **57** (2) 207-224.

Adams, C., Muir, S. and Hoque, Z. (2014) 'Measurement of sustainability performance in the public sector', *Sustainability Accounting, Management and Policy Journal*, **5** (1) 46-67.

Al Shalabi, L., Shaaban, Z. and Kasasbeh, B. (2006) 'Data mining: A preprocessing engine', *Journal of Computer Science*, **2** (9) 735-739.

Alonso, M. C., Malpica, J. A. and de Agirre, A. M. (2001). 'Consequences of the Hughes phenomenon on some classification techniques'. *Proceedings of the ASPRS 2001 annual conference*, 1-5.

Altindag, D. T. (2012) 'Crime and unemployment: Evidence from Europe', *International review of Law and Economics*, **32** (1) 145-157.

Alwee, R., Hj Shamsuddin, S. M. and Sallehuddin, R. (2013) 'Hybrid support vector regression and autoregressive integrated moving average models improved by particle swarm optimization for property crime rates forecasting with economic indicators', *The Scientific World Journal*, 2013. doi: <http://dx.doi.org/10.1155/2013/951475>.

Alwee, R., Shamsuddin, S. M. H. and Sallehuddin, R. (2013) 'Economic indicators selection for property crime rates using Grey Relational Analysis and Support Vector Regression'. *Proceedings of the International Conference on Systems, Control, Signal Processing and Informatics*, 16-19.

- Andrei, N. (2007) 'Scaled conjugate gradient algorithms for unconstrained optimization', *Computational Optimization and Applications*, **38** (3) 401-416.
- Andresen, M. A. (2006) 'Crime measures and the spatial analysis of criminal activity', *British Journal of criminology*, **46** (2) 258-285.
- Antenucci, J. C., Brown, K., Croswell, P. L., Kevany, M. J. and Archer, H. (1991) *Geographic Information Systems: a guide to the technology*.
- Azadeh, A., Babazadeh, R. and Asadzadeh, S. (2013) 'Optimum estimation and forecasting of renewable energy consumption by artificial neural networks', *Renewable and Sustainable Energy Reviews*, **27** 605-612.
- Azapagic, A. and Perdan, S. (2000) 'Indicators of sustainable development for industry: a general framework', *Process Safety and Environmental Protection*, **78** (4) 243-261.
- Baldi, P. (1995) 'Gradient descent learning algorithm overview: A general dynamical systems perspective', *IEEE Transactions on Neural Networks*, **6** (1) 182-195.
- Balestrassi, P. P., Popova, E., Paiva, A. d. and Lima, J. M. (2009) 'Design of experiments on neural network's training for nonlinear time series forecasting', *Neurocomputing*, **72** (4-6) 1160-1178.
- Basheer, I. and Hajmeer, M. (2000) 'Artificial neural networks: fundamentals, computing, design, and application', *Journal of microbiological methods*, **43** (1) 3-31.
- Basnayake, W., Attygalle, D. T., Liyanage, L. and Nandalal, K. 'Grid search based parameter tuning of dynamic neural network to forecast daily reservoir inflow'. *Proceedings of the International Conference on Computational Modeling and Simulation (ICCMS 2017), 17-19 May, 2017, Colombo, Sri Lanka*, 291-295.

- Bayar, Y. (2014) 'Effects of economic growth, export and foreign direct investment inflows on unemployment in Turkey', *Investment Management and Financial Innovations*, **11** (2) 20-27.
- Bayram, H., Sapsford, R. J., Abdelaziz, M. M. and Khair, O. A. (2001) 'Effect of ozone and nitrogen dioxide on the release of proinflammatory mediators from bronchial epithelial cells of nonatopic nonasthmatic subjects and atopic asthmatic patients in vitro', *Journal of Allergy and Clinical Immunology*, **107** (2) 287-294.
- Beck, C. and McCue, C. (2009) 'Predictive policing: what can we learn from Wal-Mart and Amazon about fighting crime in a recession?', *Police Chief*, **76** (11) 18.
- Benjamini, Y. and Hochberg, Y. (1995) 'Controlling the false discovery rate: a practical and powerful approach to multiple testing', *Journal of the royal statistical society. Series B (Methodological)*, 289-300.
- Bergstra, J. S., Bardenet, R., Bengio, Y. and Kégl, B. (2011) 'Algorithms for hyperparameter optimization'. *Advances in neural information processing systems*, 2546-2554.
- Bertiz Armenico, R., Vinluan Albert, A., Laureta Marc, P. and Paltad Anthony, B. (2015) 'Forecasting Model for Criminality in Barangay Commonwealth, Quezon City, Philippines using Data Mining Techniques', *International Journal of Conceptions on Computing and Information Technology*, **3** 28-33.
- Blum, C. and Li, X. (2008) 'Swarm intelligence in optimization', *Swarm Intelligence*: Springer, 43-85.
- Blum, E. K. and Li, L. K. (1991) 'Approximation theory and feedforward networks', *Neural networks*, **4** (4) 511-515.
- Bogomolov, A., Lepri, B., Staiano, J., Oliver, N., Pianesi, F. and Pentland, A. 'Once Upon a Crime: Towards Crime Prediction from Demographics and Mobile Data'. *Proceedings of the 16th International Conference on Multimodal Interaction*,

- Istanbul, Turkey, 2014. 2663254: ACM, 427-434.
- Bonabeau, E., Marco, D. d. R. D. F., Dorigo, M., Théraulaz, G. and Theraulaz, G. (1999) *Swarm intelligence: from natural to artificial systems*. Oxford university press.
- Breiman, L. (1996) 'Bagging predictors', *Machine learning*, **24** (2) 123-140.
- Brochu, E., Cora, V. M. and De Freitas, N. (2010) 'A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning', *arXiv preprint arXiv:1012.2599*.
- Brown, M. A. (1982) 'Modelling the spatial distribution of suburban crime', *Economic Geography*, **58** (3) 247-261.
- Burrough, P. A. (1992) 'Are GIS data structures too simple minded?', *Computers & Geosciences*, **18** (4) 395-400.
- Bursik, R. J. (1988) 'Social disorganization and theories of crime and delinquency: Problems and prospects', *Criminology*, **26** (4) 519-552.
- Calvert, J. G. (1976) 'Test of the theory of ozone generation in Los Angeles atmosphere', *Environmental science & technology*, **10** (3), 248-256.
- Camacho-Collados, M. and Liberatore, F. (2015) 'A decision support system for predictive police patrolling', *Decision Support Systems*, **75** 25-37.
- Camacho-Collados, M., Liberatore, F. and Angulo, J. M. (2015) 'A multi-criteria Police Districting Problem for the efficient and effective design of patrol sector', *European Journal of Operational Research*, **246** (2) 674-684.
- Carslaw, D. C., Beevers, S. D. and Bell, M. C. (2007) 'Risks of exceeding the hourly EU limit value for nitrogen dioxide resulting from increased road transport emissions of primary nitrogen dioxide', *Atmospheric Environment*, **41** (10) 2073-2082. doi: <https://doi.org/10.1016/j.atmosenv.2006.10.074>.

- Caplan, J. M. and Kennedy, L. W. (2016) *Risk terrain modeling: Crime prediction and risk reduction*. Univ of California Press.
- Chainey, S. and Ratcliffe, J. (2013) *GIS and crime mapping*. John Wiley & Sons.
- Chainey, S., Reid, S. and Stuart, N. (2002) *When is a hotspot a hotspot? A procedure for creating statistically robust hotspot maps of crime*. Taylor & Francis, London, England.
- Chainey, S., Tompson, L. and Uhlig, S. (2008) 'The utility of hotspot mapping for predicting spatial patterns of crime', *Security Journal*, **21** (1-2) 4-28.
- Cheema, M. and Bastiaanssen, W. G. (2010) 'Land use and land cover classification in the irrigated Indus Basin using growth phenology information from satellite data to support water management analysis', *Agricultural water management*, **97** (10) 1541-1552.
- Chen, C.-I. (2008) 'Application of the novel nonlinear grey Bernoulli model for forecasting unemployment rate', *Chaos, Solitons & Fractals*, **37** (1) 278-287.
- Chen, H., Chung, W., Xu, J. J., Wang, G., Qin, Y. and Chau, M. (2004) 'Crime data mining: a general framework and some examples', *computer*, **37** (4) 50-56.
- Chen, R., Cai, J., Meng, X., Kim, H., Honda, Y., Guo, Y. L., Samoli, E., Yang, X. and Kan, H. (2014) 'Ozone and daily mortality rate in 21 cities of East Asia: how does season modify the association?', *American journal of epidemiology*, **180** (7) 729-736.
- Civco, D. L. (1993) 'Artificial neural networks for land-cover classification and mapping', *International Journal of Geographical Information Science*, **7** (2) 173-186.
- Clark, P. J. and Evans, F. C. (1954) 'Distance to nearest neighbor as a measure of spatial relationships in populations', *Ecology*, **35** (4) 445-453.

- Clerc, M. and Kennedy, J. (2002) 'The particle swarm-explosion, stability, and convergence in a multidimensional complex space', *IEEE transactions on Evolutionary Computation*, **6** (1) 58-73.
- Cohen, J., Gorr, W. L. and Olligschlaeger, A. M. (2007) 'Leading indicators and spatial interactions: A crime-forecasting model for proactive police deployment', *Geographical Analysis*, **39** (1) 105-127.
- Coleman, T. F. and Li, Y. (1994) 'On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds', *Mathematical programming*, **67** (1-3) 189-224.
- Corcoran, J. J., Wilson, I. D. and Ware, J. A. (2003) 'Predicting the geo-temporal variations of crime and disorder', *International Journal of Forecasting*, **19** (4) 623-634.
- Cortés, U., Sánchez-Marrè, M., Ceccaroni, L., R-Roda, I. and Poch, M. (2000) 'Artificial intelligence and environmental decision support systems', *Applied intelligence*, **13** (1) 77-91.
- Cristianini, N., Shawe-Taylor, J., Elisseeff, A. and Kandola, J. S. (2002) 'On kernel-target alignment'. *Advances in neural information processing systems*, 367-373.
- Crouse, D. L., Peters, P. A., Hystad, P., Brook, J. R., van Donkelaar, A., Martin, R. V., Villeneuve, P. J., Jerrett, M., Goldberg, M. S. and Pope III, C. A. (2015) 'Ambient PM_{2.5}, O₃, and NO₂ exposures and associations with mortality over 16 years of follow-up in the Canadian Census Health and Environment Cohort (CanCHEC)', *Environmental health perspectives*, **123** (11) 1180.
- Cuinica, L. G., Abreu, I. and da Silva, J. E. (2014) 'Effect of air pollutant NO₂ on *Betula pendula*, *Ostrya carpinifolia* and *Carpinus betulus* pollen fertility and human allergenicity', *Environmental Pollution*, **186** 50-55.

- Dabney, J. B. and Harman, T. L. (2004) *Mastering simulink*. Pearson.
- Dale, R. (2016) 'The return of the chatbots', *Natural Language Engineering*, **22** (5), 811-817.
- De Bondt, W. and Vermeulen, G. (2009) 'Esperanto for EU crime statistics. Towards common European offence definitions in an EU-level offence classification system', *Gof's Research Paper Series*, **2** 87-124.
- Diaz-Chavez, R. (2014) 'Indicators for socio-economic sustainability assessment', *Socio-economic impacts of bioenergy production*: Springer, 17-37.
- Directive, H. (1992) 'Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora', *Official Journal of the European Union*, **206** 7-50.
- Directorate, O. E. (2008) OECD Key Environmental Indicators. Paris, France: Organisation for Economic Development and Cooperation.
- Dubois, E., Schurrer, C. and Velicogna, M. (2013) *The functioning of judicial systems and the situation of the economy in the European Union Member States. Compiled Report*: Conseil de l'Europe.
- Dumitru, C. and Maria, V. (2013) 'Advantages and Disadvantages of Using Neural Networks for Predictions', *Ovidius University Annals, Series Economic Sciences*, **13** (1) 444-449.
- Dunleavy, P., Margetts, H., Bastow, S. and Tinkler, J. (2006) 'New public management is dead—long live digital-era governance', *Journal of public administration research and theory*, **16** (3) 467-494.
- Eberhart, R. and Kennedy, J. 'A new optimizer using particle swarm theory'. *Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium on*: IEEE, 39-43.

- Eberhart, R. C., Shi, Y. and Kennedy, J. (2001) *Swarm intelligence*. Elsevier.
- EEA, E. (2005) *core set of indicators, Guide*: EEA Technical report.
- Elwood, S. (2011) 'Geographic Information Science: Visualization, visual methods, and the geoweb', *Progress in Human Geography*, **35** (3) 401-408.
- Evans, D., Demeter, A., GAJDOŠ, P. and Halada, L. (2013) 'Adapting environmental conservation legislation for an enlarged European Union: experience from the Habitats Directive', *Environmental Conservation*, **40** (2) 97-107.
- Ferentinou, M. and Chalkias, C. (2013) 'Mapping mass movement susceptibility across Greece with GIS, ANN and statistical methods', *Landslide Science and Practice*: Springer, 321-327.
- Ferreira, S. and Harmse, A. (2000) 'Crime and tourism in South Africa: International tourists perception and risk', *South African Geographical Journal*, **82** (2) 80-85.
- Feurer, M., Klein, A., Eggensperger, K., Springenberg, J., Blum, M. and Hutter, F. (2015) 'Efficient and robust automated machine learning'. *Advances in Neural Information Processing Systems*, 2962-2970.
- Fourie, P. and Groenwold, A. A. (2002) 'The particle swarm optimization algorithm in size and shape optimization', *Structural and Multidisciplinary Optimization*, **23** (4) 259-267.
- Franses, P. H., Paap, R. and Vroomen, B. (2004) 'Forecasting unemployment using an autoregression with censored latent effects parameters', *International Journal of Forecasting*, **20** (2) 255-271.
- Garnier, A., Eynard, J., Caussanel, M. and Grieu, S. (2014) 'Low computational cost technique for predictive management of thermal comfort in non-residential buildings', *Journal of Process Control*, **24** (6) 750-762.

- Getis, A. and Ord, J. K. (1992) 'The analysis of spatial association by use of distance statistics', *Geographical analysis*, **24** (3) 189-206.
- Gillani, S. Y. M., Rehman, H. U. and Gill, A. R. (2009) 'Unemployment, poverty, inflation and crime nexus: Cointegration and causality analysis of Pakistan', *Pakistan Economic and Social Review*, 79-98.
- Goddard, S., Harms, S. K., Reichenbach, S. E., Tadesse, T. and Waltman, W. J. (2003) 'Geospatial decision support for drought risk management', *Communications of the ACM*, **46** (1) 35-37.
- Gorr, W. L. and Lee, Y. (2015) 'Early warning system for temporary crime hot spots', *Journal of Quantitative Criminology*, **31** (1) 25-47.
- Gorte, B. and Stein, A. (1998) 'Bayesian classification and class area estimation of satellite images using stratification', *IEEE Transactions on Geoscience and Remote Sensing*, **36** (3) 803-812.
- Graf-Jaccottet, M. and Jaunin, M. H. (1998) 'Predictive models for ground ozone and nitrogen dioxide time series', *Environmetrics*, **9** (4) 393-406.
- Grossberg, S. (1988) 'Nonlinear neural networks: Principles, mechanisms, and architectures', *Neural networks*, **1** (1) 17-61.
- Gualtieri, J. A. and Crompt, R. F. 'Support vector machines for hyperspectral remote sensing classification'. *27th AIPR Workshop: Advances in Computer-Assisted Recognition*: International Society for Optics and Photonics, 221-233.
- Gunn, S. R. (1998) 'Support vector machines for classification and regression', *ISIS technical report*, **14** (1) 5-16.
- Guyon, I., Chaabane, I., Escalante, H. J., Escalera, S., Jajetic, D., Lloyd, J. R., Macià, N., Ray, B., Romaszko, L. and Sebag, M. (2016) 'A brief review of the ChaLearn

- AutoML challenge: any-time any-dataset learning without human intervention'. *Workshop on Automatic Machine Learning*, 21-30.
- Hagan, M. T. and Menhaj, M. B. (1994) 'Training feedforward networks with the Marquardt algorithm', *IEEE transactions on Neural Networks*, **5** (6) 989-993.
- Hassan, R., Cohanin, B., de Weck, O. and Venter, G. (2005) 'A Comparison of Particle Swarm Optimization and the Genetic Algorithm', *46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Co-located Conferences*: American Institute of Aeronautics and Astronautics, 1-13.
- Haykin, S. (1994) *Neural networks: a comprehensive foundation*. Prentice Hall PTR.
- He, H. S., Mladenoff, D. J., Radeloff, V. C. and Crow, T. R. (1998) 'Integration of GIS data and classified satellite imagery for regional forest assessment', *Ecological Applications*, **8** (4) 1072-1083.
- Hebb, D. O. 1949. *The organization of behavior: A neurophysiological approach*. Wiley.
- Helbich, M., Hagenauer, J., Leitner, M. and Edwards, R. (2013) 'Exploration of unstructured narrative crime reports: an unsupervised neural network and point pattern analysis approach', *Cartography and Geographic Information Science*, **40** (4) 326-336.
- Hemphill, S. A., Tollit, M., Kotevski, A. and Heerde, J. A. (2015) 'Predictors of traditional and cyber-bullying victimization: A longitudinal study of Australian secondary school students', *Journal of interpersonal violence*, **30** (15) 2567-2590.
- Hirschfield, A. (2008) 'The multi-faceted nature of crime', *Built Environment*, **34** (1) 5-20.
- Hirschfield, A. and Bowers, K. (2014) *Mapping and analysing crime data: Lessons from research and practice*. CRC Press.

- Hirschfield, A., Brown, P. and Todd, P. (1995) 'GIS and the analysis of spatially-referenced crime data: Experiences in Merseyside, UK', *International Journal of Geographical Information Systems*, **9** (2) 191-210.
- Holt, T. J., Fitzgerald, S., Bossler, A. M., Chee, G. and Ng, E. (2016) 'Assessing the risk factors of cyber and mobile phone bullying victimization in a nationally representative sample of Singapore youth', *International journal of offender therapy and comparative criminology*, **60** (5) 598-615.
- Hooghe, M., Vanhoutte, B., Hardyns, W. and Bircan, T. (2010) 'Unemployment, inequality, poverty and crime: spatial distribution patterns of criminal acts in belgium, 2001–06', *The British Journal of Criminology*, **51** (1) 1-20.
- Hornik, K. (1991) 'Approximation capabilities of multilayer feedforward networks', *Neural networks*, **4** (2) 251-257.
- Houry, S.A. (2018) 'Generic Early Warning Signals for Program Student Withdrawals: A Complexity Perspective Based on Critical Transitions and Fractals', *Computer and Systems Engineering*, **12** (1) 92.
- Huang, C., Davis, L. and Townshend, J. (2002) 'An assessment of support vector machines for land cover classification', *International Journal of remote sensing*, **23** (4) 725-749.
- Hughes, G. (1968) 'On the mean accuracy of statistical pattern recognizers', *IEEE transactions on information theory*, **14** (1) 55-63.
- Jain, A. K., Mao, J. and Mohiuddin, K. M. (1996) 'Artificial neural networks: A tutorial', *Computer*, **29** (3) 31-44.
- Jenkins, B. M. (2001) 'Protecting Public Surface Transportation Against Terrorism and Serious Crime: Continuing Research on Best Security Practices', MTI Report 01-07.

- Jones, D. R., Schonlau, M. and Welch, W. J. (1998) 'Efficient global optimization of expensive black-box functions', *Journal of Global optimization*, **13** (4) 455-492.
- Junior, A. A., Cacho, N., Thome, A. C., Medeiros, A. and Borges, J. 'A predictive policing application to support patrol planning in smart cities'. *Smart Cities Conference (ISC2), 2017 International: IEEE*, 1-6.
- Kalantzi, E. G., Makris, D., Duquenne, M. N., Kaklamani, S., Stapountzis, H. and Gourgoulanis, K. I. (2011) 'Air pollutants and morbidity of cardiopulmonary diseases in a semi-urban Greek peninsula', *Atmospheric Environment*, **45** (39) 7121-7126.
- Kanevski, M., Parkin, R., Pozdnukhov, A., Timonin, V., Maignan, M., Demyanov, V. and Canu, S. (2004) 'Environmental data mining and modeling based on machine learning algorithms and geostatistics', *Environmental Modelling & Software*, **19** (9) 845-855.
- Karathanasopoulos, A., Sermpinis, G., Stasinakis, C. and Theofilatos, K. (2015) 'Forecasting US unemployment with radial basis neural networks, Kalman filters and support vector regressions', *Computational economics*, 1-19.
- Kennedy, L. W. and Van Brunschot, E. G. (2009) *The risk in crime*. Rowman & Littlefield.
- Kennedy, L. W., Caplan, J. M. and Piza, E. (2011) 'Risk clusters, hotspots, and spatial intelligence: risk terrain modeling as an algorithm for police resource allocation strategies', *Journal of Quantitative Criminology*, **27** (3) 339-362.
- Ki, J. (2013) 'Developing a geospatial web-GIS system for landscape and urban planning', *International Journal of Digital Earth*, **6** (6) 580-588.
- Kim, D., Jeong, K., McKay, R., Chon, T. and Joo, G. (2012) 'Machine learning for predictive management: short and long term prediction of phytoplankton biomass

- using genetic algorithm based recurrent neural networks', *International Journal of Environmental Research*, **6** (1) 95-108.
- Kleemans, E. R., Soudijn, M. R. and Weenink, A. W. (2012) 'Organized crime, situational crime prevention and routine activity theory', *Trends in Organized Crime*, **15** (2-3) 87-92.
- Kohavi, R. (1995) 'A study of cross-validation and bootstrap for accuracy estimation and model selection'. *Ijcai*: Montreal, Canada, 1137-1145.
- Koskela, T., Lehtokangas, M., Saarinen, J. and Kaski, K. 'Time series prediction with multilayer perceptron, FIR and Elman neural networks'. *Proceedings of the World Congress on Neural Networks*: Citeseer, 491-496.
- Kouziokas, G. N. (2016a) 'An Information System for Monitoring Environmental Indicators in Public Management for Sustainable Development'. *11th International Scientific Conference eRA-11, at Piraeus, Greece*: Piraeus University of Applied Sciences, 16-23.
- Kouziokas, G. N. (2016b) 'Artificial intelligence and crime prediction in public management of transportation safety in urban environment', *3rd Conference on Sustainable Urban Mobility, Volos, Greece*: University of Thessaly, 534-539.
- Kouziokas, G. N. (2016c) 'Geospatial Based Information System Development in Public Administration for Sustainable Development and Planning in Urban Environment', *European Journal of Sustainable Development*, **5** (4) 347-352. doi: <http://dx.doi.org/10.14207/ejsd.2016.v5n4p347>.
- Kouziokas, G. N. (2016d) 'Technology-based management of environmental organizations using an Environmental Management Information System (EMIS): Design and development', *Environmental Technology & Innovation*, **5** 106-116 doi:<http://dx.doi.org/10.1016/j.eti.2016.01.006>

- Kouziokas, G. N. (2017a) 'An Information System for Judicial and Public Administration Using Artificial Intelligence and Geospatial Data'. *Proceedings of the 21st Pan-Hellenic Conference on Informatics*, Larissa, Greece, 2017. 3139402: ACM, 1-2, doi:http://dx.doi.org/10.1145/3139367.3139402.
- Kouziokas, G.N. (2017b) 'Applying a GIS-Based System and Crime Hotspot Analysis in Public Management and Place Marketing and Branding for Sustainable Urban Development and Planning', in *Proceedings of the 2nd Panhellenic Place Marketing and Place Branding Conference*, March 2017, Greece, University of Thessaly.
- Kouziokas, G.N. (2017c) 'Artificial Intelligence Based Crime Forecasting in Public Administration by Implementing a Feedforward Multi-layer Perceptron information', in *16th International Conference on Artificial Intelligence and Law, – VIII Workshop on Artificial Intelligence and the Complexity of Legal Systems*, June 2017, London, UK.
- Kouziokas, G.N. (2017d) 'Crime Indicators-Based Information System in Public and Judicial Management' in Bratengeyer, E., Evans, D. and Orange, G., eds., *Aristotelian Thinking Impact on the Technological Evolution and Social Progress*, IAFeS ed., Vienna, Austria: IAFeS International Association for eScience, 5 225-232.
- Kouziokas, G.N. (2017e) 'Decision Support System Based on Indicators and Artificial Intelligence in Public and Environmental Management', in *Sixth International Conference on Environmental Management, Engineering, Planning and Economics (CEMEPE) and the SECOTOX Conference*, Thessaloniki, Greece, Grafima Publications, 649-657.
- Kouziokas, G.N. (2017f) 'Feedforward Neural Network Model in Unemployment Rates Forecasting in Public Administration', in *Proceedings of the 6th International Symposium and 28th National Conference on Operational Research: OR in the*

digital era - ICT challenges, Thessaloniki, Greece, June 2017, University of Macedonia, 168-172.

Kouziokas, G.N. (2017g) 'Growth Domestic Product Forecasting by Implementing a Feedforward Multilayer Perceptron', in *Proceedings of the 3rd International Conference on EconoPhysics, September 2017, Volos, Greece, University of Thessaly*.

Kouziokas, G. N. (2017h) 'Machine Learning Technique in Time Series Prediction of Gross Domestic Product'. *Proceedings of the 21st Pan-Hellenic Conference on Informatics*, Larissa, Greece. 3139443: ACM, 1-2, doi:<http://dx.doi.org/10.1145/3139367.3139443>.

Kouziokas, G. N. (2017i) 'The application of artificial intelligence in public administration for forecasting high crime risk transportation areas in urban environment', *Transportation Research Procedia*, **24** 467-473, doi: <https://doi.org/10.1016/j.trpro.2017.05.083>.

Kouziokas, G.N. (2018a) 'Neural network-based road accident forecasting in transportation and public management', in *Proceedings of the 4th Conference on Sustainable Urban Mobility*, Skiathos, May 2018, University of Thessaly, 811-816.

Kouziokas, G.N. (2018b) 'Unemployment Prediction in UK by Using a Feedforward Multilayer Perceptron', *Operational Research in Business and Economics, Springer*, 65-74, doi: https://doi.org/10.1007/978-3-319-95666-4_5.

Kouziokas, G. N. and Perakis, K. (2017) 'Decision Support System Based on Artificial Intelligence, GIS and Remote Sensing for Sustainable Public and Judicial Management', *European Journal of Sustainable Development*, **6** (3) 397-404, doi: <https://dx.doi.org/10.14207/ejsd.2017.v6n3p397>.

Kouziokas, G. N., Chatzigeorgiou, A. and Perakis, K. (2016) 'Predicting Environmental Data in Public Management by Using Artificial Intelligence'. *Proceedings of the*

11th International Scientific Conference eRA-11, Piraeus, Greece, September: Piraeus University of Applied Sciences, 39-46.

Kouziokas, G. N., Chatzigeorgiou, A. and Perakis, K. (2017a) 'Applying Levenberg Marquardt Algorithm in Feedforward Neural Network Models for Predicting Crime in Public Management', in Christara, C. & Vavalis, M. (eds.) *Computational and Informational Sciences and Engineering*. Volos: University of Thessaly, **4** 61-70.

Kouziokas, G.N., Chatzigeorgiou, A. and Perakis, K. (2017b) 'Artificial Intelligence and Regression in Predicting Ground Water Levels in Public Administration', *European Water*, **57** 361-366.

Kouziokas, G.N., Chatzigeorgiou, A. and Perakis, K. (2017c) 'Artificial neural Networks in Forecasting Key Air Pollutant Factors in Public and Environmental Management', in Bratengeyer, E., Evans, D. and Orange, G., eds., *Aristotelian Thinking Impact on the Technological Evolution and Social Progress* IAFeS ed., Vienna, Austria: IAFeS International Association for eScience, **5** 243-250.

Kouziokas, G.N., Chatzigeorgiou, A. and Perakis, K. (2017d) 'Artificial Intelligence in Atmospheric PM10 Forecasting in Public and Environmental Management', in *Proceedings of the Sixth International Conference on Environmental Management, Engineering, Planning and Economics (CEMEPE) and the SECOTOX Conference*, Thessaloniki, Greece, Grafima Publications, 606-614.

Kouziokas, G.N., Chatzigeorgiou, A. and Perakis, K. (2017e) 'Neural network based methodology for predicting energy consumption in public and environmental management', in *Proceedings of the 6th International Symposium and 28th National Conference on Operational Research: OR in the digital era - ICT challenges*, Thessaloniki, Greece, June 2017, University of Macedonia, 140-145.

Kouziokas, G.N., Chatzigeorgiou, A. and Perakis, K. (2017f) 'Public e-Governance and City Marketing Based on Indicators Information System in Urban Environment', in

Proceedings of the 2nd Panhellenic Place Marketing and Place Branding Conference, March 2017, Greece, University of Thessaly.

- Kouziokas, G.N., Chatzigeorgiou, A. and Perakis, K. (2018) 'Final Energy Consumption Forecasting by Applying Artificial Intelligence Models', *Operational Research in Business and Economics*, Springer 1-10, doi: https://doi.org/10.1007/978-3-319-95666-4_1.
- Ku, C.-H. and Leroy, G. (2014) 'A decision support system: Automated crime report analysis and classification for e-government', *Government Information Quarterly*, **31** (4) 534-544.
- Leipnik, M. R. and Albert, D. P. (2003) *GIS in law enforcement: Implementation issues and case studies*. CRC Press.
- Leitner, M. (2013) *Crime modeling and mapping using geospatial technologies*. Springer Science & Business Media.
- Leondes, C. T. (2002) *Intelligent Systems: Technology and Applications, Six Volume Set*. CRC Press.
- Levine, N. (2006) 'Crime mapping and the Crimestat program', *Geographical analysis*, **38** (1) 41-56.
- Levine, N. (2008) 'The “Hottest” part of a hotspot: comments on “The utility of hotspot mapping for predicting spatial patterns of crime”', *Security journal*, **21** (4) 295-302, doi: <http://dx.doi.org/10.1057/sj.2008.5>.
- Levitt, S. D. (2001) 'Alternative strategies for identifying the link between unemployment and crime', *Journal of quantitative criminology*, **17** (4) 377-390.
- Liaw, A. and Wiener, M. (2002) 'Classification and regression by randomForest', *R news*, **2** (3) 18-22.

- Liu, H. 'On the levenberg-marquardt training method for feed-forward neural networks'. *Natural Computation (ICNC), 2010 Sixth International Conference on Natural Computation: IEEE*, 456-460, doi: <http://dx.doi.org/10.1109/ICNC.2010.5583151>.
- Lourakis, M. and Argyros, A. A. 'Is Levenberg-Marquardt the most efficient optimization algorithm for implementing bundle adjustment?'. *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on: IEEE*, 1526-1531, doi:<https://doi.org/10.1109/ICCV.2005.128>.
- Lourakis, M. I. A. (2005) 'A brief description of the Levenberg-Marquardt algorithm implemented by levmar', *Foundation of Research and Technology*, **4** 1-6.
- Ma, W., Wang, Y. and Dong, N. 'Study on stock price prediction based on BP neural network'. *Emergency Management and Management Sciences (ICEMMS), 2010 IEEE International Conference on: IEEE*, 57-60. doi:<http://dx.doi.org/10.1109/ICEMMS.2010.5563502>.
- Marquardt, D. W. (1963) 'An algorithm for least-squares estimation of nonlinear parameters', *Journal of the society for Industrial and Applied Mathematics*, **11** (2) 431-441.
- Martinez-Cantin, R. (2014) 'Bayesopt: A bayesian optimization library for nonlinear optimization, experimental design and bandits', *The Journal of Machine Learning Research*, **15** (1) 3735-3739.
- McCue, C. (2014) *Data mining and predictive analysis: Intelligence gathering and crime analysis*. Butterworth-Heinemann.
- McCulloch, W. S. and Pitts, W. (1943) 'A logical calculus of the ideas immanent in nervous activity', *The bulletin of mathematical biophysics*, **5** (4) 115-133.
- Melick, M. D. (2003) 'The relationship between crime and unemployment', *The Park Place Economist*, **11** (1) 29-36.

- Mendoza, H., Klein, A., Feurer, M., Springenberg, J. T. and Hutter, F. (2016) 'Towards automatically-tuned neural networks'. *Workshop on Automatic Machine Learning*, 2016, 58-65.
- Menze, B. H., Kelm, B. M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W. and Hamprecht, F. A. (2009) 'A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data', *BMC bioinformatics*, 10(1), 213.
- Metaxas, T. (2010) 'Cities competition, place marketing and economic development in South Europe: The Barcelona case as FDI destination', *Theoretical and empirical researches in urban management*, **5** (14) 5-19.
- Metaxiotis, K., Ergazakis, K., Samouilidis, E. and Psarras, J. (2003) 'Decision support through knowledge management: the role of the artificial intelligence', *Information Management & Computer Security*, **11** (5) 216-221.
- Mills, P. (2011) 'Efficient statistical classification of satellite measurements', *International Journal of Remote Sensing*, **32** (21) 6109-6132.
- Mirowsky, J. E., Dailey, L. A. and Devlin, R. B. (2016) 'Differential expression of pro-inflammatory and oxidative stress mediators induced by nitrogen dioxide and ozone in primary human bronchial epithelial cells', *Inhalation toxicology*, **28** (8) 374-382, doi: <https://dx.doi.org/10.1080%2F08958378.2016.1185199>.
- Mockus, J. (1994) 'Application of Bayesian approach to numerical methods of global and stochastic optimization', *Journal of Global Optimization*, **4** (4) 347-365.
- Mockus, J. and Mockus, L. (1991) 'Bayesian approach to global optimization and application to multiobjective and constrained problems', *Journal of Optimization Theory and Applications*, **70** (1) 157-172.
- Møller, M. F. (1993) 'A scaled conjugate gradient algorithm for fast supervised learning', *Neural networks*, **6** (4) 525-533.

- Montolio, D. and Planells-Struse, S. (2016) 'Does tourism boost criminal activity? Evidence from a top touristic country', *Crime & Delinquency*, **62** (12) 1597-1623.
- Moré, J. J. (1978) 'The Levenberg-Marquardt algorithm: implementation and theory', *Numerical analysis*: Springer, 105-116.
- Müller, K.-R., Smola, A., Rätsch, G., Schölkopf, B., Kohlmorgen, J. and Vapnik, V. (1999) 'Using support vector machines for time series prediction', *Advances in kernel methods—support vector learning*, 243-254.
- Nakaya, T. and Yano, K. (2010) 'Visualising Crime Clusters in a Space-time Cube: An Exploratory Data-analysis Approach Using Space-time Kernel Density Estimation and Scan Statistics', *Transactions in GIS*, **14** (3) 223-239.
- Ngia, L. S. and Sjoberg, J. (2000) 'Efficient training of neural nets for nonlinear adaptive filtering using a recursive Levenberg-Marquardt algorithm', *IEEE Transactions on Signal Processing*, **48** (7) 1915-1927.
- Olligschlaeger, A. M. (1997) 'Artificial neural networks and crime mapping', *Crime mapping and crime prevention*, 313-348.
- Olmedo, E. (2014) 'Forecasting spanish unemployment using near neighbour and neural net techniques', *Computational Economics*, **43** (2) 183-197.
- Openshaw, S. and Openshaw, C. (1997) *Artificial intelligence in geography*. New York, USA: John Wiley & Sons.
- Ord, J. K. and Getis, A. (1995) 'Local spatial autocorrelation statistics: distributional issues and an application', *Geographical analysis*, **27** (4) 286-306.
- Ou, C. and Lin, W. 'Comparison between PSO and GA for parameters optimization of PID controller'. *Mechatronics and Automation, Proceedings of the 2006 IEEE International Conference on*: IEEE, 2471-2475.

- Ozughalu, U. M. and Ogwumike, F. O. (2013) 'Can Economic Growth, Foreign Direct Investment and Exports Provide the Desired Panacea to the Problem of Unemployment in Nigeria?', *Journal of Economics and Sustainable Development*, **4** (1) 36-51.
- Pal, M. (2005) 'Random forest classifier for remote sensing classification', *International Journal of Remote Sensing*, **26** (1) 217-222.
- Palocsay, S. W., Wang, P. and Brookshire, R. G. (2000) 'Predicting criminal recidivism using neural networks', *Socio-Economic Planning Sciences*, **34** (4) 271-284, doi: [https://doi.org/10.1016/S0038-0121\(00\)00003-3](https://doi.org/10.1016/S0038-0121(00)00003-3).
- Parikh, R., Mathai, A., Parikh, S., Sekhar, G. C. and Thomas, R. (2008) 'Understanding and using sensitivity, specificity and predictive values', *Indian journal of ophthalmology*, **56** (1) 45, doi: <http://dx.doi.org/10.4103/0301-4738.37595>.
- Pattyn, M. and Wouters, P. (2008) 'Prioritizing Crime Problems in Belgium, According to Strategic Police Planning: Developing the National Police Security Picture for Belgium by Means of a Multi-Criteria Decision-Making Model', *The Handbook of Knowledge-Based Policing: Current Conceptions and Future Directions*, 341-362.
- Paulsen, D. J. and Robinson, M. B. (2004) *Spatial aspects of crime: Theory and practice*. Pearson A and B.
- Perry, W. L. (2013) *Predictive policing: The role of crime forecasting in law enforcement operations*. CA: Rand Corporation.
- Pino-Mejías, R., Jiménez-Gamero, M.-D., Cubiles-de-la-Vega, M.-D. and Pascual-Acosta, A. (2008) 'Reduced bootstrap aggregating of learning algorithms', *Pattern recognition letters*, **29** (3) 265-271.
- Potter, G. W. and Kappeler, V. E. (2006) *Constructing crime: Perspectives on making news and social problems*.

- Prasad, A. M., Iverson, L. R. and Liaw, A. (2006) 'Newer classification and regression tree techniques: bagging and random forests for ecological prediction', *Ecosystems*, **9** (2) 181-199.
- Ratcliffe, J. H. (2004) 'The hotspot matrix: A framework for the spatio-temporal targeting of crime reduction', *Police practice and research*, **5** (1) 5-23.
- Ratcliffe, J. H. and McCullagh, M. J. (1999) 'Hotbeds of crime and the search for spatial accuracy', *Journal of geographical systems*, **1** (4) 385-398.
- Riedmiller, M. and Braun, H. 'A direct adaptive method for faster backpropagation learning: The RPROP algorithm'. *Neural Networks, 1993., IEEE International Conference on: IEEE*, 586-591.
- Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M. and Rigol-Sanchez, J. P. (2012) 'An assessment of the effectiveness of a random forest classifier for land-cover classification', *ISPRS Journal of Photogrammetry and Remote Sensing*, **67** 93-104.
- Sabins, F. F. (2007) *Remote sensing: principles and applications*. Waveland Press.
- Scott, L. M. and Janikas, M. V. (2010) 'Spatial statistics in ArcGIS', *Handbook of applied spatial analysis*: Springer, 27-41.
- Shahriari, B., Swersky, K., Wang, Z., Adams, R. P. and De Freitas, N. (2016) 'Taking the human out of the loop: A review of bayesian optimization', *Proceedings of the IEEE*, **104** (1) 148-175.
- Shawe-Taylor, J., Williams, C. K., Cristianini, N. and Kandola, J. (2005) 'On the eigenspectrum of the Gram matrix and the generalization error of kernel-PCA', *IEEE Transactions on Information Theory*, **51** (7) 2510-2522.
- Shawe-Taylor, J., Williams, C., Cristianini, N. and Kandola, J. (2002) 'On the eigenspectrum of the Gram matrix and its relationship to the operator

- eigenspectrum'. *International Conference on Algorithmic Learning Theory*: Springer, 23-40.
- Shi, Y. and Eberhart, R. 'A modified particle swarm optimizer'. *Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on*: IEEE, 69-73.
- Shoosmith, G. L. (2013) 'Space–time autoregressive models and forecasting national, regional and state crime rates', *International journal of forecasting*, **29** (1) 191-201.
- Snoek, J., Larochelle, H. and Adams, R. P. (2012) 'Practical bayesian optimization of machine learning algorithms'. *Advances in neural information processing systems*, 2951-2959.
- Stefanov, W. L., Ramsey, M. S. and Christensen, P. R. (2001) 'Monitoring urban land cover change: An expert system approach to land cover classification of semiarid to arid urban centers', *Remote sensing of Environment*, **77** (2) 173-185.
- Steihaug, T. (1983) 'The conjugate gradient method and trust regions in large scale optimization', *SIAM Journal on Numerical Analysis*, **20** (3) 626-637.
- Stolz, B. A. (2015) 'The growth of federal criminal justice policy making: the role of US civil rights legislation', *Criminal justice policy review*, **26** (5) 463-487.
- Strahler, A. H. (1980) 'The use of prior probabilities in maximum likelihood classification of remotely sensed data', *Remote sensing of Environment*, **10** (2) 135-163, doi: [https://doi.org/10.1016/0034-4257\(80\)90011-5](https://doi.org/10.1016/0034-4257(80)90011-5).
- Sun, L. and Schulz, K. (2015) 'The improvement of land cover classification by thermal remote sensing', *Remote Sensing*, **7** (7) 8368-8390, doi: <https://doi.org/10.3390/rs70708368>.

- Suykens, J. A. and Vandewalle, J. (1999) 'Least squares support vector machine classifiers', *Neural processing letters*, **9** (3) 293-300, doi: <https://doi.org/10.1023/A:1018628609742>.
- Svozil, D., Kvasnicka, V. and Pospichal, J. (1997) 'Introduction to multi-layer feed-forward neural networks', *Chemometrics and intelligent laboratory systems*, 39 (1) 43-62.
- Swersky, K., Duvenaud, D., Snoek, J., Hutter, F. and Osborne, M. A. (2014) 'Raiders of the lost architecture: Kernels for Bayesian optimization in conditional parameter spaces', *arXiv preprint arXiv:1409.4011*.
- Tambe, M., Pence, K. and Speer, P. (2016) 'Optimal Allocation of Police Patrol Resources Using a Continuous-Time Crime Model'. *Decision and Game Theory for Security: 7th International Conference, GameSec 2016, New York, NY, USA, November 2-4, 2016, Proceedings*: Springer, 139.
- Tang, Z. and Fishwick, P. A. (1993) 'Feedforward neural nets as models for time series forecasting', *ORSA journal on computing*, **5** (4) 374-385.
- Tao, Y., Huang, W., Huang, X., Zhong, L., Lu, S.-E., Li, Y., Dai, L., Zhang, Y. and Zhu, T. (2012) 'Estimated acute effects of ambient ozone and nitrogen dioxide on mortality in the Pearl River Delta of southern China', *Environmental health perspectives*, **120** (3) 393.
- Townshend, J. R., Justice, C. and Kalb, V. (1987) 'Characterization and classification of South American land cover types using satellite data', *International Journal of Remote Sensing*, **8** (8) 1189-1207.
- Tu, J. V. (1996) 'Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes', *Journal of clinical epidemiology*, **49** (11) 1225-1231.

- Turing, A. M. (1948) 'Intelligent machinery, a heretical theory', *The Turing test: Verbal behavior as the hallmark of intelligence*, 105.
- U.K. Office for National Statistics (2018) 'Data and Statistics about the U.K.' www.ons.gov.uk
- U.S. Bureau of Justice Statistics (2018) 'Data and Statistics about the U.S.' www.bjs.gov.
- U.S. Bureau of Labor Statistics (2018) 'Data and Statistics about the U.S.' www.bls.gov.
- U.S. Environmental Protection Agency (2018) 'Data and Statistics about the U.S.' www.epa.gov.
- Vapnik, V. (1995) *The Nature of Statistical Learning Theory*. Springer-Verlag, New York.
- Vapnik, V. (1998) 'The support vector method of function estimation', *Nonlinear Modeling*: Springer, 55-85.
- Varotsos, C. (2005) 'Modern computational techniques for environmental data; application to the global ozone layer', *Computational Science-ICCS 2005*, Springer Berlin Heidelberg, 43-68.
- Ventura, S. J. (1995) 'The use of geographic information systems in local government', *Public Administration Review*, 461-467.
- Voženílek, V. 'Artificial intelligence and GIS: mutual meeting and passing'. *Intelligent Networking and Collaborative Systems, 2009. INCOS'09. International Conference on: IEEE*, 279-284.
- Wang, B., Wang, D., Chan, C., Yin, R. and Black, D. (2018) 'Predictive Management of Electric Vehicles in a Community Microgrid', *arXiv preprint arXiv:1802.01512*.
- Wang, D., Ding, W., Lo, H., Morabito, M., Chen, P., Salazar, J. and Stepinski, T. (2013) 'Understanding the spatial distribution of crime based on its related variables using

- geospatial discriminative patterns', *Computers, Environment and Urban Systems*, **39** 93-106.
- Watanachaturaporn, P., Varshney, P. K. and Arora, M. K. (2004). 'Evaluation of factors affecting support vector machines for hyperspectral classification'. *the American Society for Photogrammetry & Remote Sensing (ASPRS) 2004 Annual Conference, Denver, CO*.
- Weizenbaum, J. (1966) 'ELIZA—a computer program for the study of natural language communication between man and machine', *Communications of the ACM*, **9** (1) 36-45.
- Whitlock, C., Shafer, S. L. and Marlon, J. (2003) 'The role of climate and vegetation change in shaping past and future fire regimes in the northwestern US and the implications for ecosystem management', *Forest ecology and management*, **178** (1-2) 5-21.
- Widrow, B. and Hoff, M. E. (1960) *Adaptive switching circuits*: In 1960 IRE WESCON convention record, 96–104. New York: IRE. Reprinted in Anderson, A. & E. Rosenfeld (eds) 1988. *Neurocomputing—foundations of research*. Cambridge, MA: MIT Press.
- Wiig, K. M. (2002) 'Knowledge management in public administration', *Journal of knowledge management*, **6** (3) 224-239.
- Wilamowski, B. M. and Irwin, J. D. (2016) *Intelligent systems*. CRC Press.
- Williams, M., Atkinson, R., Anderson, H. and Kelly, F. (2014) 'Associations between daily mortality in London and combined oxidant capacity, ozone and nitrogen dioxide', *Air Quality, Atmosphere & Health*, **7** (4) 407-414.
- Worrall, J. L. (2014) *Crime control in America: What works?:* Pearson Higher Ed.

- Xia, J., Liao, W., Chanussot, J., Du, P., Song, G. and Philips, W. (2015) 'Improving random forest with ensemble of features and semisupervised feature extraction', *IEEE Geoscience and Remote Sensing Letters*, **12** (7) 1471-1475.
- Xu, S., An, X., Qiao, X., Zhu, L. and Li, L. (2013) 'Multi-output least-squares support vector regression machines', *Pattern Recognition Letters*, **34** (9) 1078-1084.
- Yildiz, R., Ocal, O. and Yildirim, E. (2013) 'The Effects of Unemployment, Income and Education on Crime: Evidence from Individual Data', *International Journal of Economic Perspectives*, **7** (2) 32.
- Zhang, G. P., Patuwo, B. E. and Hu, M. Y. (2001) 'A simulation study of artificial neural networks for nonlinear time-series forecasting', *Computers & Operations Research*, **28** (4) 381-396.
- Zhang, G., Patuwo, B. E. and Hu, M. Y. (1998) 'Forecasting with artificial neural networks:: The state of the art', *International journal of forecasting*, **14** (1) 35-62.
- Zhao, S. J. and Xu, Y. M. 'Levenberg-Marquardt algorithm for nonlinear principal component analysis neural network through inputs training'. *Intelligent Control and Automation, 2004. WCICA 2004. Fifth World Congress on: IEEE*, 3278-3281.
- Zhou, G., Lin, J. and Ma, X. (2014) 'A Web-Based GIS for Crime Mapping and Decision Support', *Forensic GIS*: Springer, 221-243.
- Zimring, F. E. (2011) *The city that became safe: New York's lessons for urban crime and its control*. Oxford University Press.