ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΒΙΟΪΑΤΡΙΚΗ

Ερμηνεύσιμη Ομαδοποίηση με Αξιοποίηση Νευρωνικών Δικτύων

Φοιτήτρια: Μουλακάκη Ελευθερία
Επιβλέπων Καθηγητής: Τασουλής Σωτήριος
Μέλη Επιτροπής: Δελήμπασης Κωνσταντίνος,Καχαρούντας Αθανάσιος , Τασουλής Σωτήριος

Ημερομηνία: Ιούλιος 2025

Περιεχόμενα

6 Ιουλίου 2025

Περίληψη

Μια πρόσφατη τάση στη μηχανική μάθηση ήταν ο εμπλουτισμός των μαθησιακών μοντέλων με την ικανότητα να εξηγούν τις δικές τους προβλέψεις. Το αναδυόμενο πεδίο της εξηγήσιμης τεχνητής νοημοσύνης (XAI) έχει μέχρι στιγμής επικεντρωθεί κυρίως στην εποπτευόμενη μάθηση, ιδίως στους ταξινομητές βαθιών νευρωνικών δικτύων. Σε πολλά πρακτικά προβλήματα, ωστόσο, οι πληροφορίες της ετικέτας δεν δίνονται και ο στόχος είναι να ανακαλυφθεί η υποκείμενη δομή των δεδομένων, για παράδειγμα, τις συστάδες τους. Ενώ υπάρχουν ισχυρές μέθοδοι για την εξαγωγή της δομής συμπλέγματος σε δεδομένα, συνήθως δεν απαντούν στην ερώτηση γιατί ένα συγκεκριμένο σημείο δεδομένων έχει εκχωρηθεί σε ένα δεδομένο σύμπλεγμα. Προτείνεται ένα νέο πλαίσιο που μπορεί, να εξηγήσει τις εκχωρήσεις συμπλέγματος ως προς τα χαρακτηριστικά εισόδου με αποτελεσματικό και αξιόπιστο τρόπο. Βασίζεται στη νέα αντίληψη ότι τα μοντέλα ομαδοποίησης μπορούν να ξαναγραφτούν ως νευρωνικά δίκτυα —ή να «νευρονοποιηθούν». Οι προβλέψεις συμπλέγματος των ληφθέντων δικτύων μπορούν στη συνέχεια να αποδοθούν γρήγορα και με ακρίβεια στα χαρακτηριστικά εισόδου.

1 Εισαγωγή.

Η ομαδοποίηση αποτελεί μια πολύ σημαντική μέθοδο μη επιβλεπόμενης μάθησης που έχει ως κύριο στόχο να εξάγει συμπεράσματα σχετικά με τη δομή και τις εγγενείς ετερογένειες των αρχικών δεδομένων.Στη μη επιβλεπόμενη μάθηση διαθέτουμε μόνο τα δεδομένα εισόδου , δεν διαθέτουμε ετικέτες για τα αρχικά δεδομένα και με τη χρήση κατάλληλων αλγορίθμων προσπαθούμε να ανακαλύψουμε μοτίβα και ομοιότητες στα ίδια τα δεδομένα με σκοπό την τοποθέτηση τους σε ομάδες.Η πληροφορία αυτή που εξάγεται απο τη διαδικασία της ομαδοποίησης είναι σημαντική για να οδηγηθούμε σε συμπεράσματα σχετικά με την δομή και τη σχέση των δεδομένων σε διάφορους τομείς από την έκφραση γονιδίων και τη σύνθεση των οικοσυστημάτων έως και την ανάλυση δεδομένων κειμένου.Μια βασική τεχνική ομαδοποίησης που θα αναλυθεί παρακάτω είναι ο αλγόριθμος K-μέσων , ο οποίος βασίζεται στον υπολογισμό των αποστάσεων μεταξύ των δεδομένων και των αντίστοιχων κέντρων των ομάδων που θέλουμε να τοποθετήσουμε τα δεδομένα . Η τοποθέτηση των δεδομένων σε ομάδες γίνεται με βάση την απόσταση τους απο το αντίστοιχο κέντρο . Με την ολοκλήρωση του αλγορίθμου θα έχουμε τις ομάδες με τις τιμές των κέντρων τους καθώς και τις ετικέτες των δεδομένων για την ομάδα στην οποία τοποθετήθηκαν. Για τη διαδικασία της ¨νευρονοποίησης' χρειαζόμαστε τα κέντρα των ομάδων ώστε να υπολογιστούν τα βάρη και οι μεροληπτικοί όροι του νευρωνικού δικτύου που θα αναπτυχθεί.Το νευρωνικό δίκτυο θα αναπαριστά μια συγκεκριμένη ομαδοποίηση σε συγκεκριμένα δεδομένα και για συγκεκριμένο αριθμό ομάδων , για τον λόγο αυτόν τα βάρη και η πόλωση

που υπολογίζεται παραμένουν σταθερά ,δηλαδή δεν υπάρχει εκπαίδευση.Σκοπός είναι μέσω της διαδικασίας αυτής να υλοποιήσουμε μια τεχνική LRP ώστε μέσω των επιπέδων του νευρωνικού δικτύου να αποδοθεί η συνάφεια , δηλαδή πόσο μεγάλη συμβολή είχε κάθε χαρακτηριστικό μίας εγγραφής στην τοποθέτησή της σε μια συγκεκριμένη ομάδα.Με τον τρόπο αυτό δίνεται μια εξήγηση στο γιατί ένα σημείο τοποθετήθηκε σε μια συγκεκριμένη ομάδα.Για να επιτευχθεί αυτό εκτελούνται δύο βήματα :

1. Δημιουργία του νευρωνικού δικτύου που θα αναπαριστά την ομαδοποίηση .

2. Οπισθοδιάδοση της συνάφειας μέσω των επιπέδων του νευρωνικού δικτύου και της διαδικασίας LRP.

Μέχρι στιγμής η έρευνα για τις μεθόδους επεξήγησης έχει επικεντρωθεί στη περίπτωση της εποπτευόμενης μάθησης.Οι μέθοδοι που βασίζονται στη διαβάθμιση , τις τοπικές διαταραχές και τις υποκατάστατες συναρτήσεις δεν κάνουν συγκεκριμένες υποθέσεις σχετικά με τη δομή του μοντέλου.Ορισμένες τεχνικές ερμηνείας ομαδοποίησης έχουν βασιστεί σε αντίστοιχα δέντρα απόφασης όπου το δέντρο εκπαιδεύεται να προσεγγίζει τον αλγόριθμο κ-μέσων και η ανάθεση σε ομάδες ερμηνεύεται χρησιμοποιώντας τεχνικές XAI ειδικές για δέντρα αποφάσεων.Με αυτή την προσέγγιση ο χρήστης θα πρέπει να αντισταθμίσει το αποτέλεσμα του αρχικού μοντέλου με την επεξήγηση.Στη τεχνική αυτή σκοπός δεν είναι τόσο η βέλτιωση ενός βασικού μοντέλου ομαδοποίησης με τη χρήση νευρωνικών δικτύων όσο στο να καταστούν οι υπάρχοντες δημοφιλείς αλγόριθμοι ομαδοποίησης εξηγήσιμοι.Η εργασία ενισχύει την επικύρωση των μοντέλων ομαδοποίησης με την παραγωγή ερμηνεύσιμων απο τον άνθρωπο δεδομένων,γεγονός κρίσιμο για τον προσδιορισμό για το αν οι αναθέσεις σε ομάδες υποστηρίζονται απο σημαντικά χαρακτηριστικά ή από αυτά που ο χρήστης θεωρεί τεχνουργήματα.

2 Χρήση της ερμηνεύσιμης τεχνητής νοημοσύνης σε σχετικά πεδία.

Ο τομέας της ερμηνεύσιμης τεχνητής νοημοσύνης έχει πολλές εφαρμογές σε ζητήματα που αφορούν προβλήματα της καθημερινότητας καθώς η εξήγηση στο γιατί ένας αλγόριθμος πήρε μια συγκεκριμένη απόφαση , στο γιατί πρόβλεψε μια συγκεκριμένη τιμή είναι κρίσιμη για τη κατανόηση του προβλήματος από τον χρήστη αλλά και από ανθρώπους οι οποίοι δεν σχετίζονται με τον κλάδο.Έτσι η διαφάνεια που δημιουργείται για τη λειτουργία ενός μοντέλου και στην αιτιολόγιση της απόφασης που λήφθηκε οδηγεί στη δημιουργία εμπιστοσύνης ανάμεσα στους χρήστες αλλά και στους ενδιαφερόμενους κάθε κλάδου ώστε η τεχνητή νοημοσύνη να μην λειτουργεί ως ένα μαύρο κουτί αλλά εγαλείο που μπορούμε να κατανοήσουμε και να εμπιστευτούμε.Τέτοιοι τομείς που μπορεί να βρεί εφαρμογή μια τέτοια τεχνική είναι :

1. Στην ιατρική και στη διάγνωση ασθενειών όπως για παράδειγμα πώς ένας αλγόριθμος εντόπισε έναν καρκινικό όγκο ή γιατί έχει προταθεί ένα συγκεκριμένο φάρμακο.

2. Στον χρηματοοικονομικό και τραπεζικό τομέα , για παράδειγμα γιατί απορρίφθηκε ένα δάνειο ή γιατί μία συναλλαγή χαρακτηρίζεται ύποπτη.

3. Σε νομικές και δικαστικές υποθέσεις , ένας αλγόριθμος που αναλύει συμβόλαια πρέπει να εξηγεί τις νομικές επιπτώσεις ή στη πρόταση μίας ποινής να υπάρχει διαφάνεια και δικαιοσύνη.

4. Στη Κυβερνοασφάλεια και ανίχνευση απειλών , για παράδειγμα γιατί ένα email θεωρείται phishing ή πως ένας αλγόριθμος εντόπισε μια εισβολή.

5. Στο μάρκετινγκ δίνει εξήγηση στο πώς ένας αλγόριθμος κατηγοριοποίησε έναν πελάτη σε μια συγκεκριμένη ομάδα.

6. Στη κλιματική αλλαγή και στο περιβάλλον , δίνει απάντηση στο πώς προβλέπεται η κλιματική αλλαγή ή στο γιατί προτείνει συγκεκριμένες πολιτικές μείωσης ρύπων.

2.1 Χρησιμότητα της NEON τεχνικής στη παρούσα εργασία.

Με τη μεθοδολογία που θα προταθεί παρακάτω στη συγκεκριμένη εργασία στο σύνολο δεδομένων BBC dataset πετυχαίνουμε την εξής λειτουργία :

- Εύρεση λέξεων που έχουν συμβάλει στην ομαδοποίηση των κειμένων με μια επιπλέον μέθοδο , τη διαδικασία LRP και οχι αποκλειστικά με τη καταμέτρηση συχνότητας λέξεων κάθε ομάδας.

- Εξάγουμε μέσω των λέξεων που προκύπτουν απο τη συγκεκριμένη διαδικασία κατηγορίες θεμάτων για κάθε ομάδα και κατέπέκταση για κάθε κείμενο που ανήκει σε αυτή. Δηλαδή αποτελεί μια εναλλακτική μέθοδο topic modeling με πιο συγκεκριμένες λέξεις για την περιγραφή κάθε θέματος.

Ουσιαστικά με τη νευρονοποίηση της ομαδοποίησης K-μέσων και ύστερα με την εφαρμογή της τεχνικής LRP για την εύρεση της συνάφειας κάθε χαρακτηριστικού σε δεδομένα κειμένου διαμορφώνουμε έναν διαφορετικό τρόπο περιγραφής του θέματος κάθε ομάδας χωρίς να ακολουθούμε τις τεχνικές των αλγορίθμων topic modeling. Επίσης για την περιγραφή των ομάδων δεν χρησιμοποιούμε τη συνηθισμένη διαδικασία καταμέτρησης συχνότητας των λέξεων που απαρτίζουν κάθε ομάδα αλλά με βάση αποκλειστικά τη συνάφεια που προκύπτει σε κάθε λέξη κάθε κειμένου κάθε ομάδας εξάγουμε συμπεράσματα για το είδος των λέξεων που καθόρισαν τη τοποθέτηση του συγκεκριμένου κειμένου στη συγκεκριμένη ομάδα. Όπως αναφέρεται και στο αρχικό επιστημονικό άρθρο το οποίο ακολουθείται αυστηρά , η συγκεκριμένη τεχνική δεν στοχεύει στη βελτίωση της ποιότητας της ομαδοποίησης των δεδομένων , στη περίπτωση αυτή των κειμένων , αλλά προτείνει μια διαφορετική μέθοδο για να εξηγηθεί η δομή των συστάδων , η λογική πίσω απο τον σχηματισμό τους καθώς και η εξαγωγή θεμάτων για κάθε συστάδα. Παρατηρείται ότι οι λέξεις που προκύπτουν για κάθε συστάδα είναι αρκετά σχετικές με το θέμα το οποίο επικρατεί στα κείμενα καθώς παρατηρούνται ορολογίες πιο εξειδικευμένες .

3 Σχετικές έρευνες.

Έως τώρα οι έρευνες σχετικά με τις τεχνικές εξήγησης στη τεχνητή νοημοσύνη έχουν επικεντρωθεί κυρίως στην εποπτευόμενη μάθηση. Οι μέθοδοι των κλίσεων [9],[1] , στις

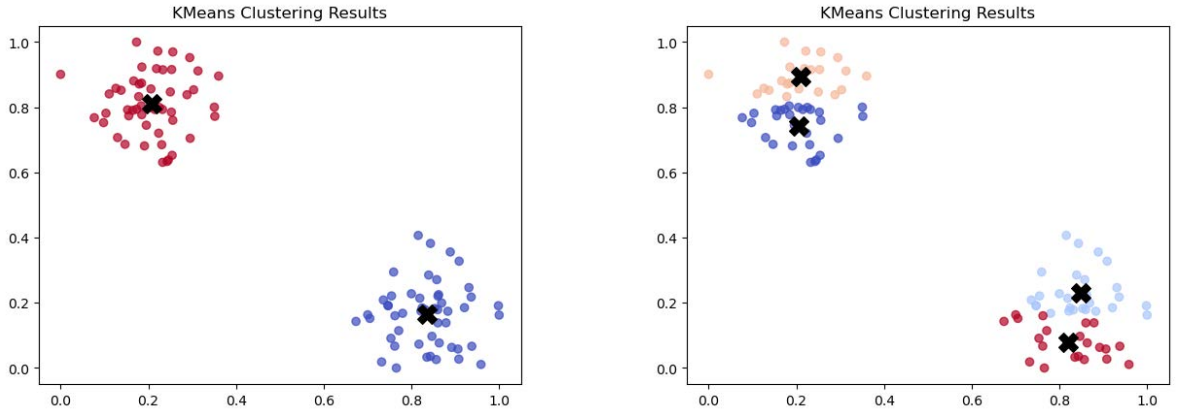
τοπικές διαταραχές (local perturbations) [6],[5] και οι υποκατάστατες συναρτήσεις (surrogate functions) [8] εφαρμόζονται σε ένα ευρύ φάσμα ταξινομητών επομένως δεν κάνουν υποθέσεις σχετικά με τη δομή του μοντέλου.Άλλες μέθοδοι αναπαριστούν τον ταξινομητή ως νευρωνικό δίκτυο και εφαρμόζεται οπισθοδιάδοση προκειμένου να παραχθούν εξηγήσεις . Άλλες έρευνες έχουν εφαρμόσει τη μέθοδο αυτή σε τύπους μοντέλων όπως τα SVM μίας κατηγορίας [7] ,LSTM δίκτυα ή νευρωνικά δίκτυα γραφημάτων.Επίσης μέθοδοι όπως η LIME [8] και οι εξηγήσεις Shapley [4] που κινούνται σε αυτή τη κατεύθυνση δεν δίνουν πληροφορίες σχετικά με το υποκείμενο σύνολο δεδομένων. Τονίζεται ότι μέθοδοι ερμηνείας σε συστάδες έχουν μέχρι στιγμής εφαρμοστεί σε δέντρα αποφάσεων που προσεγγίζουν όσο περισσότερο γίνεται την ομαδοποίηση κ-μέσων [3].Έρευνες σχετικά με τη σύνδεση ενός αλγόριθμου ομαδοποίησης και των νευρωνικών δικτύων έχουν ξανά γίνει όπως με το μοντέλο k-meansNet [2]το οποίο δημιουργεί ένα νευρωνικό δίκτυο ώστε να προσομοιώνει ένα μοντέλο ομαδοποίησης . Η διαφορά είναι ότι αυτοί οι μέθοδοι στοχεύουν στο να πλησιάσουν όσο περισσότερο γίνεται το αποτέλεσμα της ομαδοποίησης ενώ στη συγκεκριμένη εργασία στόχος είναι η ερμηνευσιμότητα της ομαδοποίησης .

4 Απαραίτητες γνώσεις.

4.1 Ο αλγόριθμος K-μέσων.

Η ομαδοποίηση K-μέσων είναι μια απλή προσέγγιση για τον διαχωρισμό ενός συνόλου δεδομένων σε K διακριτές μη επικαλυπτόμενες συστάδες . Για να ξεκινήσει η διαδικασία της ομαδοποίησης πρέπει πρώτα να καθορίσουμε τον επιθυμητό αριθμό των συστάδων K, τότε ο αλγόριθμος θα αντιστοιχίσει κάθε σημείο σε ακριβώς μια απο τις συστάδες K.Ορίζοντας αρχικά αυθαίρετα τις τιμές των κέντρων των ομάδων που ορίσαμε και στη συνέχεια υπολογίζει την ευκλείδια απόσταση κάθε σημείου από κάθε κέντρο.Η τοποθέτηση του σημείου σε ομάδα γίνεται με βάση την απόσταση , δηλαδή επιλέγεται η μικρότερη απο αυτές τις αποστάσεις και το σημείο τοποθετείται στην ομάδα που ανήκει το αντίστοιχο κέντρο.Πιο συγκεκριμένα για κάθε κέντρο $\mu_k \in R^d$ υπολογίζεται η απόσταση από κάθε σημείο $x \in R^d$ και η τοποθέτησή του στη ομάδα c γίνεται σύμφωνα με :

$$\forall k \neq c: \ \|x - \mu_c\|^2 < \|x - \mu_k\|^2 \tag{1}$$



(α′) K-μέσων για K=2 (β′) K-μέσων για K=4

Σχήμα 1: Εφαρμογή του αλγορίθμου K-μέσων σε τυχαίο σύνολο δεδομένων με K = 2 και K = 4 .

Στα διαγράμματα του σχήματος 1 αναπαριστάται η εφαρμογή του αλγορίθμου K-μέσων σε τυχαίο σύνολο δεδομένων 100 σημείων και αναπαράσταση των αντίστοιχων ομάδων και κέντρων στις 2 διαστάσεις για διαφορετικές τιμές του K , όπου K αριθμός ομάδων .Τα διαφορετικά χρώματα στα γραφήματα χρησιμοποιούνται για να απεικονήσουν τις διαφορετικές ομάδες όπου κάθε σημείο τοποθετήθηκε σύμφωνα με τον αλγόριθμο.Τονίζεται οτι ο χρωματισμός δεν σχετίζεται με τις ομάδες αλλά η εφαρμογή του χρώματος είναι αυθαίρετη.Αυτές οι ετικέτες συμπλέγματος δεν χρησιμοποιήθηκαν στην ομαδοποίηση αλλά είναι τα ίδια τα αποτελέσματα της ομαδοποίησης.

Τα βήματα που ακολουθεί ο αλγόριθμος είναι:

1. Δήλωση του αριθμού K όπου συμβολίζει τον αριθμό των επιθυμητών ομάδων που θέλουμε να διαχωρίσουμε το σύνολό μας , καθώς και αρχικοποίηση με τυχαία κέντρα.

2. Υπολογισμός τετραγωνικής ευκλείδιας απόστασης μεταξύ κάθε σημείου και κάθε κέντρου.Η τοποθέτηση σε ομάδα γίνεται σύμφωνα με τη μικρότερη απόσταση , το σημείο τοποθετείται στην ομάδα που απέχει μικρότερη απόσταση απο το κέντρο του.Ακολουθεί ανανέωση των κέντρων μέσω του υπολογισμού πλέον των νέων μέσων κάθε ομάδας.Αυτό επιτυγχάνεται με το άθροισμα των τιμών των σημείων που ανήκουν σε αυτή διαιρώντας το αποτέλεσμα με το πλήθος των σημείων που την απαρτίζουν.

3. Το δεύτερο βήμα επαναλαμβάνεται μέχρι να οδηγηθούμε σε σταθεροποίηση στις αναθέσεις.

Οπότε αν μια παρατήριση i βρίσκεται στην ομάδα K τότε ισχύει ότι $i \in C^k$.Σκοπός της ομαδοποίησης είναι η διαφορετικότητα/παραλλαγή εντός ομάδας να είναι όσο μικρότερη γίνεται.Το μέτρο αυτό ορίζεται ως $W(C^k)$ και αφορά το πόσο οι παρατηρήσεις εντός συμπλέγματος διαφέρουν μεταξύ τους.

$$\underset{C_1,\ldots,C_K}{minimize} \left\{ \sum_{k=1}^{K} W(C_k) \right\}. \tag{2}$$

Αυτό σημαίνει ότι κάθε παρατήρηση θα χωριστεί σε τόσες ομάδες όσο το K ώστε η εντός των ομάδων διαφοροποίηση για όλα τα K να είναι όσο μικρότερη γίνεται. Ο υπολογισμός γίνεται με βάση την τετραγωνική ευκλείδια απόσταση , πιο συγκεκριμένα:

$$W(C_k) = \frac{1}{|C_k|} \sum_{i,i' \in C_k} \sum_{j=1}^{p} (x_{ij} - x_{i'j})^2 \tag{3}$$

Όπως γίνεται φανερό από τον παραπάνω τύπο η διαφοροποίηση εντός των συστάδων για την κ-οστή ομάδα είναι το άθροισμα όλων των τετραγωνικών ευκλείδιων αποστάσεων μεταξύ των σημείων που την απαρτίζουν,το αποτέλεσμα διαιρείται με το πλήθος των σημείων που βρίσκονται στην ομάδα.Ο αλγόριθμος τη βέλτιστη ομαδοποίηση τη βρίσκει ύστερα απο πολλές επαναλήψεις μέχρι να καταλήξει σε μια σταθερή και βέλτιστη ομαδοποίηση διότι στο πρώτο βήμα οι αρχικοποιήσεις είναι τυχαίες . Σχετικά με την απόφαση για την τιμή του K που αφορά το πλήθος των ομάδων δεν είναι τόσο απλή και τυχαία.

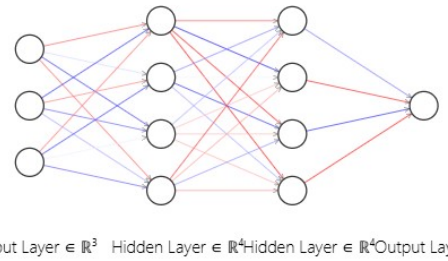
4.2 Δομή απλού νευρωνικού δικτύου.

Τα νευρωνικά δίκτυα αποτελούν είδος αλγόριθμου μηχανικής μάθησης όπου προσπαθούν να επεξεργαστούν δεδομένα λαμβάνοντας έμπνευση απο τον τρόπο που είναι δομημένος ο ανθρώπινος εγκέφαλος.Ένα νευρωνικό δίκτυο αποτελείται από επίπεδα , κάθε επίπεδο αποτελείται από κόμβους-τεχνητούς νευρώνες οι οποίοι συνδέονται με τους κόμβους του επόμενου και του προηγούμενου επιπέδου.Μέσω αυτών γίνεται επεξεργασία των δεδομένων εισόδου και παράγεται μια έξοδος , αυτό επιτυγχάνεται μέσω εκπαίδευσης σε μεγάλα σύνολα δεδομένων και μοντελοποιούν περίπλοκες σχέσεις μεταξύ δεδομένων εισόδου-εξόδου κάνοντας γενικεύσεις και λαμβάνοντας αποφάσεις.Ένα βασικό νευρωνικό δίκτυο αποτελείται από τρία είδη επιπέδων:

1. Επίπεδο εισόδου , όπου εισέρχονται τα διανύσματα των χαρακτηριστικών των δεδομένων εισόδου.

2. Κρυφό επίπεδο , εκεί εκτελούνται διάφοροι μετασχηματισμοί και υπολογισμοί και βρίσκεται μεταξύ εισόδου και εξόδου.

3. Επίπεδο εξόδου ,το επίπεδο αυτό είναι το τελικό επίπεδο όπου το αποτέλεσμά του εξαρτάται από τις ανάγκες του αντίστοιχου προβλήματος.



Σχήμα 2: Αναπαράσταση της δομής ενός απλού νευρωνικού δικτύου .

Στο σχήμα 2 αναπαριστάται η δομή ενός απλού νευρωνικού δικτύου . αποτελούμενο απο ένα επίπεδο εισόδου τριών νευρώνων,δύο κρυφά επίπεδα από τέσσερις νευρώνες το

καθένα πλήρες συνδεδεμένα μεταξύ τους ,δηλαδή κάθε νευρώνας από ένα επίπεδο συνδέεται με όλους τους νευρώνες του επόμενου.Τέλος υπάρχει ένα επίπεδο εξόδου το οποίο αποτελείται από ένα νευρώνα.

Ένα απλό νευρωνικό δίκτυο εμπρόσθιας τροφοδότισης δέχεται ως είσοδο ένα διάνυσμα p μεταβλητών , $X = (X_1, X_2, ..., X_p)$ και σχηματίζεται μια μη γραμμική συνάρτηση $f(x)$ για την πρόβλεψη του Y.

$$f(X) = \beta_0 + \sum_{k=1}^{K} \beta_k h_k(X) = \beta_0 + \sum_{k=1}^{K} \beta_k g\left(w_{k0} + \sum_{j=1}^{p} w_{kj} X_j\right). \tag{4}$$

Ο παραπάνω τύπος αποτυπώνει τη πράξη που εφαρμόζεται στο διάνυσμα εισόδου $f(x)$, πιο συγκεκριμένα με τον όρο b_0 αναφερόμαστε στη μεροληψία του νευρώνα ο οποίος συμβάλει στη μετατόπιση της συνάρτησης ενεργοποίησης ώστε να μαθαίνει καλύτερα το μοντέλο,σε αυτόν προστίθενται το γινόμενο των μεροληπτικών όρων κάθε νευρώνα του κρυφού επιπέδου για K νευρώνες με τη συνάρτηση ενεργοποίησης η οποία παίρνει ως όρισμα το άθροισμα για όλους τους νευρώνες του γινομένου των βαρών και κάθε χαρακτηριστικό της εισόδου. Ως συνάρτηση ενεργοποίησης προτιμάται η $ReLu$ η οποία ορίζεται ως εξής :

$$g_z = (z)_+ = \begin{cases} 0, & if z \leq 0 \\ z, & otherwise \end{cases} \tag{5}$$

Η συνάρτηση ενεργοποίησης αυτή αποθηκεύεται και εκτελείται πιο αποτελεσματικά απο τη σιγμοηδή.Σε αυτήν η είσοδος είναι γραμμική και το αποτέλεσμα για z μικρότερο ΄ή ίσο του μηδενός δίνει τιμή 0 ενώ δίνει z για κάθε άλλη περίπτωση.

Εκπαιδεύοντας ένα νευρωνικό δίκτυο εκτιμάται μια παράμετρος την οποία θέλουμε να ελαχιστοποιηθεί , συνήθως χρησιμοποιείται η απώλεια του τετραγωνικού σφάλματος που θέλουμε να ελαχιστοποιήσουμε:

$$\sum_{i=1}^{n} (y_i - f(x_i))^2 \tag{6}$$

Ουσιαστικά αυτό είναι το άθροισμα της τετραγωνικής διαφοράς της πραγματικής τιμής από αυτή που εκτιμάται.

Ακολουθεί παράδειγμα εφαρμογής του τύπου υπολογισμού της εξόδου y ενός απλού νευρωνικού δικτύου όπως αυτό το οποίο αναπαρίσταται στο σχήμα 3 για τυχαίο σημείο εισόδου $Q = [1, 1]$, ένα κρυφό επίπεδο τριών νευρώνων και επίπεδο εξόδου ενός νευρώνα.Τονίζεται ότι οι τιμές των βαρών , των μεροληπτικών όρων είναι τυχαίες επιλογές και χρησιμοποιούνται κυρίως για την εφαρμογή του τύπου:

$$w_{10} = 0, \quad w_{11} = 1, \quad w_{12} = 1, \quad w_{20} = 0, \quad w_{21} = -1, \quad w_{22} = -1$$

$$b_0 = 0, \quad b_1 = 1, \quad b_2 = -1$$

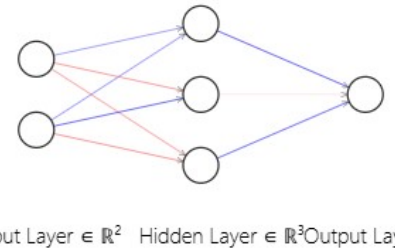
$$f(X) = 0 + [b_1 \cdot g\left(w_{10} + (w_{11} X_1) + (w_{12} X_2)\right)] + [b_2 \cdot g\left(w_{20} + (w_{21} X_1) + (w_{22} X_2)\right)]$$

$$f(X) = 0 + 1 \cdot g\left[(0 + X_1) + (1 \cdot X_2)\right] + (-1) \cdot g\left[(0 + (-X_1) + (-X_2)\right]]$$

$$= 0 + g(X_1 + X_2) - g(-X_1 - X_2)$$
$$= 0 + g(2) - g(-2)$$

Χρησιμοποιώντας ως συνάρτηση ενεργοποίησης τη $ReLu$ στη περίπτωση του $g(2)$ αφού $2 > 0$ η τιμή που επιστρέφεται είναι 2 ενώ στη περίπτωση που $g(-2)$ το $2 < 0$ επιστρέφεται η τιμή 0. Άρα το τελικό αποτέλεσμα είναι : $0 + 2 - 0 = 2$. Άρα ισχύει:

$$f\Big([1, 1]\Big) = 2$$



Σχήμα 3: Αρχικτεκτονική του νευρωνικού δικτύου που χρησιμοποιήθηκε για να λύσει το παραπάνω παράδειγμα.

4.3 Δείκτες και μετρικές αξιολόγισης που θα χρησιμοποιηθούν.

Adjusted Rand Index (ARI).

Είναι μια μετρική η οποία χρησιμοποιείται πολύ συχνά για να υπολογίσει την ομοιότητα ανάμεσα στις αναθέσεις δύο ομαδοποιήσεων. Είναι βελτίωση του δείκτη ((RI)) , λαμβάνοντας τιμές στο εύρος από -1 εως 1 όπου υψηλότερες τιμές υποδεικνύουν μεγαλύτερη ομοιότητα.Μπορεί να χρησιμοποιηθεί σε διάφορους αλγορίθμους ομαδοποίησης όπως:

1. Ομαδοποίηση βασισμένη σε αντιπροσώπους: K-μέσων , μοντέλα Gaussian Mix.
2. Ιεραρχική ομαδοποίηση : Συσσωρευτική.
3. Ομαδοποίηση βασισμένη στη πυκνότητα : DBSCAN.
4. Αλγόριθμοι ανίχνευσης κοινότητας : μέθοδος Louvain για ομαδοποίηση γραφημάτων.

Ο δείκτης Rand (RI) αποτελεί μέτρο συμφωνίας ομαδοποίησης με βάση τη καταμέτρηση του αριθμού των ζευγών σημείων δεδομένων που συγκεντρώνονται σταθερά σε δύο διαφορετικές ομάδες . Ο δείκτης ARI συμπεριλαμβάνει σε αυτό και την τυχαιότητα στην ομαδοποίηση εξασφαλίζοντας μια πιο ουσιαστική αξιολόγηση. Ο υπολογισμός γίνεται ως εξής:

$$RI = \frac{a + b}{a + b + c + d} \tag{7}$$

Όπου αν έχουμε δύο ομάδες C1 , C2 , ο συμβολισμός a αφορά τον αριθμό των ζευγών δειγμάτων που ανήκουν στην ίδια ομάδα τόσο στη C1 όσο και στη C2 , ο συμβολισμός b , αναφέρεται στον αριθμό των ζευγών δειγμάτων σε διαφορετικές ομάδες και στο C1 και C2 , η τιμή c αφορά τον αριθμό των ζευγών δειγμάτων στην ίδια ομάδα στο C1 αλλά σε διαφορετικές ομάδες στο C2 και η τιμή d αφορά τον αριθμό των ζευγών των δειγμάτων σε διαφορετικές ομάδες στο C1 αλλά στην ίδια στο C2.

Η αναμενόμενη τιμή για τον δείκτη RI υπολογίζεται ως εξής :

$$E = \frac{\sum \binom{n_i}{2} \cdot \sum \binom{m_j}{2}}{\binom{N}{2}} \tag{8}$$

Όπου n_i, m_j ο αριθμός των σημείων σε κάθε ομάδα.Πλέον η τιμή ARI υπολογίζεται :

$$ARI = \frac{RI - E}{1 - E} \tag{9}$$

Το εύρος τιμών που λαμβάνει είναι από -1 έως 1. Τιμές κοντά στο 1 σημαίνει ότι υπάρχει τέλεια συμφωνία ανάμεσα στις δύο ομαδοποιήσεις, το 0 αντιστοιχεί σε τυχαία ομαδοποίηση ενώ οι αρνητικές τιμές σηματοδοτούν χειρότερη ομαδοποίηση και απο τη τυχαιότητα.

4.4 Νευρονοποίηση αλγορίθμου K-μέσων.

Για να μπορέσουμε να εφαρμόσουμε τη τεχνική αυτή στον αλγόριθμο ομαδοποίησης K-μέσων αρχικά θα πρέπει να αποτυπώσουμε τη συνάρτηση απόφασης του αλγορίθμου $g_c(x)$ με τη χρήση του νευρωνικού δικτύου.Το συγκεκριμένο νευρωνικό δίκτυο που θα αφομοιώνει αυτή την απόφαση θα αποτελείται απο το επίπεδο εισόδου όπου εισέρχεται το διάνυσμα χαρακτηριστικών κάθε σημείου , ένα κρυφό επίπεδο το οποίο αποτελείται από τόσους νευρώνες όσο είναι ο αριθμός K δηλαδή οι ομάδες που επιθυμούμαι να διαχωρίσουμε τα δεδομένα και ένα επίπεδο εξόδου στο οποίο λαμβάνεται απλά η απόφαση για την ομάδα , δεν θεωρείται νευρώνας διότι δεν εκτελεί κάποια πράξη απλά επιλέγει την μικρότερη απόσταση.Επομένως στο τελευταίο επίπεδο δεν διαθέτει λειτουργίες νευρώνα άρα ούτε μεροληπτικό όρο bias ούτε βάρη weights . Το δίκτυο τονίζεται ότι δεν εκπαιδεύεται άρα τα βάρη και οι μεροληπτικοί όροι παραμένουν σταθερά μετά τον υπολογισμό τους. Επίσης το δίκτυο αποτυπώνει μια συγκεκριμένη ομαδοποίηση για συγκεκριμένα σημεία κάθε φορά και συγκεκριμένα βάρη και μεροληπτικούς όρους.Αυτό οφείλεται στο γεγονός ότι ο υπολογισμός αυτών των τιμών εξαρτάται άμεσα από τα κέντρα που έχουν προκύψει από την εφαρμογή του αλγορίθμου προγενέστερα στα δεδομένα.Άρα πρώτα απαιτείται να εφαρμόσουμε τον αλγόριθμο K-μέσων στα δεδομένα και έπειτα με τα κέντρα που θα έχουν προκύψει με το πέρας του να χτίσουμε το νευρωνικό δίκτυο. Η αρχιτεκτονική του είναι η ακόλουθη:

$$h_k = w'_k \cdot x + b_k \qquad \text{(Επίπεδο 1)} \tag{10}$$

$$f_c = \min\{h_k\} \quad \text{για } k \neq c \qquad \text{(Επίπεδο 2)} \tag{11}$$

Με h_k αναφερόμαστε στην τιμή κάθε νευρώνα του 1ου κρυφού επιπέδου και εκτελείται γραμμική πράξη με την αντίστοιχη είσοδο και με f_c αναφερόμαστε στην επιλογή της δεύτερης μικρότερης τιμής όπου αντιστοιχεί στο όριο απόφασης.

Για τον υπολογισμό των βαρών :

$$w_k = 2 \cdot (\mu_c - \mu_k) \tag{12}$$

Με μ_c το κέντρο της ομάδας στόχο και μ_k το κέντρο της ομάδα ανταγωνιστή.
Για τον μεροληπτικό όρο:

$$b_k = \|\mu_k\|^2 - \|\mu_c\|^2 \tag{13}$$

Με μ_c και μ_k να έχουν την ίδια ερμηνεία με παραπάνω. Ακολουθεί παράδειγμα με εφαρμογή του αλγορίθμου K-μέσων και έπειτα της νευρονοποίησης του για τυχαία σημεία: $X_1 = [0,0], X_2 = [0,1], X_3 = [1,0], X_4 = [1,1]$τα σημεία τέσσερα και είναι στις δύο διαστάσεις.Θα εφαρμόσουμε τον αλγόριθμο K-μέσων για δύο ομάδες.Ξεκινώντας τη διαδικασία αρχικά έχουμε τυχαία κέντρα για τις δύο ομάδες $m_1 = [0,0]$ και $m_2 = [1,1]$.Έπειτα γίνεται υπολογισμός της τετραγωνικής ευκλείδιας απόστασης κάθε σημείου από κάθε κέντρο. Επομένως με χρήση του τύπου (1) έχουμε:
Για το σημείο $X_1 = (0,0)$:

$$d(X_1, m_1) = \sqrt{(0-0)^2 + (0-0)^2} = 0$$

$$d(X_1, m_2) = \sqrt{(1-0)^2 + (1-0)^2} = \sqrt{2}$$

Για το σημείο $X_2 = (0,1)$:

$$d(X_2, m_1) = \sqrt{(0-0)^2 + (1-0)^2} = 1$$

$$d(X_2, m_2) = \sqrt{(1-0)^2 + (1-1)^2} = 1$$

Για το σημείο $X_3 = (1,0)$:

$$d(X_3, m_1) = \sqrt{(1-0)^2 + (0-0)^2} = 1$$

$$d(X_3, m_2) = \sqrt{(1-1)^2 + (1-0)^2} = 1$$

Για το σημείο $X_4 = (1,1)$:

$$d(X_4, m_1) = \sqrt{(1-0)^2 + (1-0)^2} = \sqrt{2}$$

$$d(X_4, m_2) = \sqrt{(1-1)^2 + (1-1)^2} = 0$$

Τα σημεία X_1 και X_2 αντιστοιχίζονται στο m_1, ενώ τα σημεία X_3 και X_4 αντιστοιχίζονται στο m_2.Στη συνέχεια ακολουθεί ο υπολογισμός των νέων κέντρων και προκύπτουν τα εξής νέα κέντρα:
Για το νέο κέντρο m_1:

$$m_1' = \left(\frac{0+0}{2}, \frac{0+1}{2}\right) = (0, 0.5)$$

Για το νέο κέντρο m_2:

$$m_2' = \left(\frac{1+1}{2}, \frac{0+1}{2}\right) = (1, 0.5)$$

Με την επανάληψη του αλγορίθμου παρατηρείται σταθεροποίηση στις αναθέσεις των κέντρων οπότε τα κέντρα για τις δύο ομάδες είναι αυτά που προέκυψαν. Για τον υπολογισμό των βαρών και των μεροληπτικών όρων θα γίνει χρήση των τύπων (12) και (13).Εφόσον έχουμε δύο ομάδες θα έχουμε δύο νευρώνες στο πρώτο κρυφό επίπεδο και δύο νευρώνες εισόδου αφού έχουμε δύο χαρακτηριστικά.Άρα θα υπολογίσουμε δύο βάρη αφού K=2 και δύο μεροληπτικούς όρους.

Για τα βάρη w_1 και w_2, έχουμε:

$$w_1 = 2 \cdot (m_2 - m_1) = 2 \cdot \left(\begin{bmatrix} 1 \\ 0.5 \end{bmatrix} - \begin{bmatrix} 0 \\ 0.5 \end{bmatrix} \right) = 2 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

και

$$w_2 = 2 \cdot (m_1 - m_2) = 2 \cdot \left(\begin{bmatrix} 0 \\ 0.5 \end{bmatrix} - \begin{bmatrix} 1 \\ 0.5 \end{bmatrix} \right) = 2 \cdot \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

Για τους μεροληπτικούς όρους b_1 και b_2 , έχουμε:
Υπολογισμός b_1:

$$||m_1||^2 = 0^2 + 0.5^2 = 0 + 0.25 = 0.25$$

$$||m_2||^2 = 1^2 + 0.5^2 = 1 + 0.25 = 1.25$$

Άρα:

$$b_1 = ||m_1||^2 - ||m_2||^2 = 0.25 - 1.25 = -1$$

Υπολογισμός b_2:

$$||m_1||^2 = 0.25$$

$$||m_2||^2 = 1.25$$

Άρα:

$$b_2 = ||m_2||^2 - ||m_1||^2 = 1.25 - 0.25 = 1$$

Συνοπτικά:

$$b_1 = -1 \quad \text{και} \quad b_2 = 1$$

Για την νευρονοποίηση και τον υπολογισμό κάθε τιμής του νευρώνα h_k χρησιμοποιούμε τον τύπο (10) :
Υπολογισμός του h_1:

Για το $X_1 = [0, 0]$:

$$h_1 = [2, 0] \cdot \begin{bmatrix} 0 \\ 0 \end{bmatrix} + (-1) = 0 + 0 - 1 = -1$$

Για το $X_2 = [0, 1]$:

$$h_1 = [2, 0] \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} + (-1) = 0 + 0 - 1 = -1$$

Για το $X_3 = [1, 0]$:

$$h_1 = [2, 0] \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + (-1) = 2 + 0 - 1 = 1$$

Για το $X_4 = [1, 1]$:

$$h_1 = [2, 0] \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} + (-1) = 2 + 0 - 1 = 1$$

Υπολογισμός του h_2:

Για το $X_1 = [0, 0]$:

$$h_2 = [-2, 0] \cdot \begin{bmatrix} 0 \\ 0 \end{bmatrix} + 1 = 0 + 0 + 1 = 1$$

Για το $X_2 = [0, 1]$:

$$h_2 = [-2, 0] \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 1 = 0 + 0 + 1 = 1$$

Για το $X_3 = [1, 0]$:

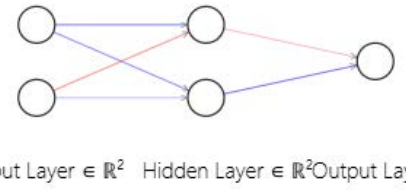
$$h_2 = [-2, 0] \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 = -2 + 0 + 1 = -1$$

Για το $X_4 = [1, 1]$:

$$h_2 = [-2, 0] \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1 = -2 + 0 + 1 = -1$$

Σύμφωνα λοιπόν με το επίπεδο εξόδου όπου δεν είναι νευρώνας με τη κλασική έννοια αλλά γίνεται επιλογή της μικρότερης τιμής - απόστασης απο κάθε σημείο - νευρώνα προκύπτει ότι από τον τύπο (10) για το επίπεδο εξόδου τίθεται ο περιορισμός ότι η επιλογή θα γίνεται για $k \neq c$ άρα εφόσον για c έχουμε τη μικρότερη έξοδο-απόσταση και εμείς θέλουμε για $k \neq c$να επιλέξουμε το f_c θα επιλέξουμε τη δεύτερη μικρότερη απόσταση όπου αυτή είναι η ανταγωνιστική ομάδα από αυτή που ανατίθεται το σημείο.Επειδή στο συγκεκριμένο παράδειγμα έχουμε μόνο δύο ομάδες αναγκαστικά η τιμή του f_c προκύπτει απο τη δεύτερη επιλογή άρα : Για το σημείο X_1 η μικρότερη απόσταση είναι η -1 του νευρώνα h_1 άρα τοποθετείται σύμφωνα με τη παραπάνω λογική στην ομάδα 1,η δεύτερη μικρότερη τιμή που τώρα λόγω δύο ομάδων είναι η τιμή 1 του h_2 είναι η τιμή του f_c για το σημείο X_2 επίσης η μικρότερη απόσταση είναι η -1 του νευρώνα h_1 και σύμφωνα με τα παραπάνω θα τοποθετηθεί στην ομάδα 1 και θα έχει ως τιμή f_c τη τιμή 1 του δεύτερου νευρώνα, ενώ για τα άλλα δύο σημεία X_3, X_4 που η μικρότερη απόσταση είναι επίσης η -1 βρίσκεται στον νευρώνα h_2 άρα θα τοποθετηθούν στην ομάδα 2 και με τιμή f_c την

επιλογή της δεύτερης μικρότερης τιμής που είναι η 1 του νευρώνα h_1 . Διαπιστώνεται ότι η ομαδοποίηση που προκύπτει απο το ισοδύναμο νευρωνικό του K-μέσων για τα συγκεκριμένα τυχαία σημεία και τα αντίστοιχα βάρη και μεροληπτικοί όροι που προέκυψαν από τα κέντρα που υπολογίστηκαν απο τον αλγόριθμο στο πρώτο στάδιο οδηγούν σε ισοδύναμη ομαδοποίηση , όπως είδαμε δεν υπάρχει εκπαίδευση των βαρών , το ισοδύναμο νευρωνικό δίκτυο φαίνεται στο σχήμα 4.



Σχήμα 4: Στο σχήμα φαίνεται το ισοδύναμο νευρωνικό δίκτυο που υλοποιεί τον αλγόριθμο K-μέσων για K=2, για τέσσερα σημεία 2 διαστάσεων-χαρακτηριστικών το καθένα.

4.5 Διαδικασία LRP.

Εφόσον πλέον έχει δημιουργηθεί το νευρωνικό δίκτυο το οποίο αντικατοπτρίζει την ομδοποίηση K-μέσων σε συγκεκριμένο σύνολο δεδομένων και για συγκεκριμένο πλήθος ομάδων , τώρα δίνεται η δυνατότητα μέσω της τεχνικής διάδοσης της συνάφειας ανάλογα με το επίπεδο ή αλλιώς LRP να αποδοθεί συνάφεια σε καθένα απο τα χαρακτηριστικά εισόδου των δεδομένων προκειμένου να διαπιστωθεί ποιό χαρακτηριστικό έχει συμβάλει στην τοποθέτηση ενός σημείου σε μία συγκεκριμένη ομάδα.Η τεχνική αυτή αξιοποιεί τη δομή του νευρωνικού δικτύου κάνοντας ένα πέρασμα προς τα πίσω , δηλαδή ξεκινώντας από το τελευταίο επίπεδο αποδίδουμε τη συνάφεια πηγαίνοντας προς τα πίσω ανά επίπεδο μέχρι να φτάσουμε στα χαρακτηριστικά εισόδου.Ουσιαστικά γίνεται αναδιανομή των ποσοτήτων συμβολής σε κάθε νευρώνα κάθε επιπέδου.Αυτοί οι κανόνες που θα παρουσιαστούν παρακάτω έχουν σχεδιαστεί σκόπιμα για το έργο της εξήγησης. Ξεκινώντας από το τελευταίο επίπεδο όπου είναι η έξοδος του νευρωνικού δικτύου f_c για κάθε σημείο ,διανέμεται η τιμή αυτή στο αμέσως προηγούμενο επίπεδο το οποίο είναι το κρυφό επίιπεδο hk που αποτελείται απο τόσους νευρώνες όσες και οι ομάδες δηλαδή αριθμό K. Η στρατηγική που ακολουθείται είναι η min-take-most (MTM) όπου μικρότερες είσοδοι σε αυτή τη συνάρτηση λαμβάνουν μεγαλύτερο μερίδιο ποσότητας προς αναδιανομή.Πιο συγκεκριμένα ο τύπος που εφαρμόζεται στο επίπεδο αυτό είναι:

$$R_k = \frac{\exp(-\beta h_k)}{\sum_{k \neq c} \exp(-\beta h_k)} f_c \tag{14}$$

Όπου f_c είναι η δεύτερη μικτότερη τιμή-απόσταση από την ανταγωνιστική ομάδα που λαμβάνει ένα σημείο , h_k είναι η τιμή του νευρώνα για την k-οστή ομάδα, δηλαδή η απόσταση του σημείου από τον νευρώνα h_k και β είναι μία θετική υπερπαράμετρος ακαμψίας η οποία παρεμβάλεται μεταξύ ομοιόμορφης στρατηγικής ανακατανομής (β=0) και μίας στρατηγικής min-take-all $b \to \infty$.Σε σύγκριση με αυτές τις δύο ακραίες περιπτώσεις η προσέγγιση αυτή επιτρέπει την πλαισίωση της εξήγησης δηλαδή τη μη ανακατανομή σε ανταγωνιστές που είναι πολύ μακριά και επομένως άσχετοι και ταυτόχρονα εξασφαλίζει τη συνέχεια της

εξήγησης καθώς μεταβαίνουμε από έναν πλησιέστερο ανταγωνιστή ομάδας σε άλλο. Η τιμή της υπολογίζεται ως εξής:

$$\beta = E[f_c]^{-1} \tag{15}$$

Όπου η παράμετρος ακαμψίας είναι αντιστρόφος ανάλογη της προσδοκίας των σκόρ σε ολόκληρο το σύνολο δεδομένων. Οι ενδιάμεσες βαθμολογίες συνάφειας που προκύπτουν από τους παραπάνω τύπους τώρα θα διανεμηθούν στο αμέσως προηγούμενο επίπεδο που είναι το επίπεδο εισόδου και τα χαρακτηριστικά που θέλουμε να μελετήσουμε ,πιο συγκεκριμένα ο τύπος που θα εφαρμοστεί για τον υπολογισμό της συνάφειας σε κάθε χαρακτηριστικό είναι :

$$R_i = \sum_{k \neq c} \frac{(x_i - m_{ik}) \cdot w_{ik}}{\sum_i (x_i - m_{ik}) \cdot w_{ik}} R_k \tag{16}$$

Ο τύπος αυτός λειτουργεί ως εξής : Για κάθε σημείο πολλών διαστάσεων , x_i είναι η διάσταση - χαρακτηριστικό i , απο αυτό αφαιρείται το $midpoint_k$ το οποίο είναι η ενδιάμεση απόσταση μεταξύ του κέντρου της ομάδας που έχει τοποθετηθεί το σημείο και του κάθε ανταγωνιστή k και υπολογίζεται από τον τύπο $m_k = (m_c + m_k)/2$, η διαφορά αυτή που προκύπτει πολλαπλασιάζεται με τη τιμή του βάρους του χαρακτηριστικού αυτού στον αντίστοιχο νευρώνα η πράξη αυτή γίνεται για όλα τα $k \neq c$, η ποσότητα αυτή διαιρείται με το άθροισμα της ποσότητας αυτής για όλα τα χαρακτηριστικά του σημείου και τέλος πολλαπλασιάζεται με τη συνάφεια του νευρώνα k.Όσον αφορά τη τιμή των ενδιάμεσων κέντρων λειτουργεί ως ουδέτερο σημείο ή αλλιώς σημείο αναφοράς καθώς μέσω αυτού βλέπουμε προς ποιά ομάδα και κατεύθυνση σπρώχνεται το σημείο δηλαδή ως προς την ομάδα που τοποθετήθηκε ή τον ανταγωνιστή.Επίσης , δημιουργεί στιγμιότυπα ομαδοποιήσεων για δύο ομάδες αφού κάθε φορά γίνεται σύγκριση μεταξύ ομάδας στόχου και ανταγωνστή προσεγγίζοντας ακόμα καλύτερα τη μαθηματική θεμελίωση Shapley. Για να γίνει πιο κατανοητή η εφαρμογή των τύπων ακολουθεί ο υπολογισμός του LRP για το παραπάνω παράδειγμα:

Από τις προηγούμενες διαδικασίες καταλήξαμε ότι: X_1 : $f_c = 1$ και ομάδα c=1, X_2 : $f_c = 1$ και ομάδα c=1, X_3 : $f_c = 1$ και ομάδα c=2, X_4 : $f_c = 1$ και ομάδα c=2,

Με χρήση του τύπου (14) στη συγκεκριμένη περίπτωση λόγω δύο ομάδων οδηγούμαστε ότι $h_k = f_c$, επειδή ως έξοδο του νευρωνικού παίρνουμε την δεύτερη μικρότερη απόσταση αναγκαστηκά εδώ προκύπτει παντού η τιμή 1.

Υπολογισμός συνάφειας στο κρυφό επίπεδο:

Για το X_1 : $R_2 = 1$
Για το X_2: $R_2 = 1$
Για το X_3: $R_1 = 1$
Για το X_4: $R_1 = 1$

Υπολογισμός midpoint:

Τα κέντρα που είχαν προκύψει με το πέρας του αλγορίθμου είναι για την ομάδα 1 $m_1 = [0, 0.5]$ και για την ομάδα 2 $m_2 = [1, 0.5]$, στη συγκεκριμένη περίπτωση που έχουμε μόνο δύο ομάδες άρα δύο νευρώνες στο κρυφό επίπεδο θα υπολογιστεί μόνο ένα ενδιάμεσο κέντρο:

$$m_k = \frac{([0, 0.5] + [1, 0.5])}{2} = [0.5, 0.5]$$

Υπολογισμός συνάφειας στα χαρακτηριστικά εισόσου:

Για την απόδοση της συνάφειας στα χαρακτηριστικά εισόδου θα γίνει χρήση του τύπου (16).Προηγουμένως τα βάρη που υπολογίστηκαν είναι $w_1 = [2, 0]$ και $w_2 = [-2, 0]$. Έχουμε δύο χαρακτηριστικά για καθένα απο τα τέσσερα σημεία που εισέρχονται στο νευρωνικό δίκτυο , άρα :

Για το $X_1 = [0, 0]$ με c=1 και k=2:

$$R_1 = \frac{(x_1 - m_{12}) * w_{12}}{(x_1 - m_{12}) * w_{12} + (x_2 - m_{22}) * w_{22}} * R_2 = \frac{(0 - 0.5) * (-2)}{(0 - 0.5) * (-2) + (0 - 0.5) * 0} * 1 = 1$$

$$R_2 = \frac{(x_2 - m_{22}) * w_{22}}{(x_1 - m_{12}) * w_{12} + (x_2 - m_{22}) * w_{22}} * R_2 = \frac{(0 - 0.5) * 0}{(0 - 0.5) * 0 + (0 - 0.5) * (-2)} * 1 = 0$$

Άρα για το σημείο X_1 παρατηρούμε ότι το χαρακτηριστικό που συμβάλει στην τοποθέτηση του στην ομάδα 1 είναι το πρώτο χαρακτηριστικό λόγω μεγαλύτερης τιμής-συνάφειας.

Για το $X_2 = [0, 1]$ με ομάδα c=1 και k = 2 :

$$R_1 = \frac{(x_1 - m_{12}) * w_{12}}{(x_1 - m_{12}) * w_{12} + (x_2 - m_{22}) * w_{22}} * R_2 = \frac{(0 - 0.5) * (-2)}{(0 - 0.5) * (-2) + (0 - 0.5) * 0} * 1 = 1$$

$$R_2 = \frac{(x_2 - m_{22}) * w_{22}}{(x_1 - m_{12}) * w_{12} + (x_2 - m_{22}) * w_{22}} * R_2 = \frac{(1 - 0.5) * 0}{(1 - 0.5) * 0 + (0 - 0.5) * (-2)} * 1 = 0$$

Αντίστοιχα και εδώ παρατηρούμε ότι στη τοποθέτηση του στην ομάδα 1 συνέβαλε το πρώτο χαρακτηριστικό με τιμή 1.

Για το $X_3 = [1, 0]$ με ομάδα c=2 και k = 1 :

$$R_1 = \frac{(x_1 - m_{11}) * w_{11}}{(x_1 - m_{11}) * w_{11} + (x_2 - m_{21}) * w_{21}} * R_1 = \frac{(1 - 0.5) * 2}{(0 - 0.5) * 0 + (1 - 0.5) * 2} * 1 = 1$$

$$R_2 = \frac{(x_2 - m_{21}) * w_{21}}{(x_1 - m_{11}) * w_{11} + (x_2 - m_{21}) * w_{21}} * R_1 = \frac{(0 - 0.5) * 0}{(1 - 0.5) * 2 + (0 - 0.5) * 0} * 1 = 0$$

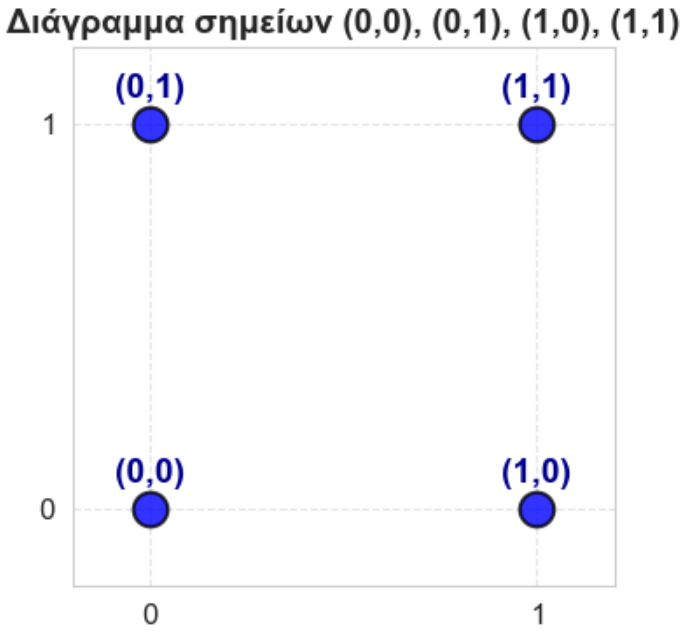
Πάλι παρατηρούμε ότι μεγαλύτερη συμβολή έχει το πρώτο χαρακτηριστικό.

Για το $X_4 = [1, 1]$ με ομάδα c=2 και k = 1

$$R_1 = \frac{(x_1 - m_{11}) * w_{11}}{(x_1 - m_{11}) * w_{11} + (x_2 - m_{21}) * w_{21}} * R_1 = \frac{(1 - 0.5) * 2}{(1 - 0.5) * 0 + (1 - 0.5) * 2} * 1 = 1$$

$$R_2 = \frac{(x_2 - m_{21}) * w_{21}}{(x_1 - m_{11}) * w_{11} + (x_2 - m_{21}) * w_{21}} * R_2 = \frac{(1 - 0.5) * 0}{(1 - 0.5) * 2 + (1 - 0.5) * 0} * 1 = 0$$

Τέλος και στο τελευταίο σημείο διαπιστώνεται ότι το χαρακτηριστικό που έχει τη μεγαλύτερη συνάφεια και επομένως συνέβαλε περισσότερο στη τοποθέτηση του στην ομάδα 2 είναι το πρώτο χαρακτηριστικό.Το αποτέλεσμα αυτό είναι απόλυτα λογικό διότι στα διανύσματα των σημείων που δηλώνουν συντεταγμένες η πρώτη συνιστώσα αφορά τη θέση του σημείου στο άξονα x και με βάση αυτό το χαρακτηριστικό προέκυψαν οι δύο ομάδες όπως γίνεται και ξεκάθαρο στο διάγραμμα του σχήματος 5.Τονίζεται ότι σύμφωνα με το επιστημονικό άρθρο στο οποίο βασίζεται η παρούσα εργασία , η θεωρία των Shapley values χρησιμοποιείται για να δικαιολογήσει σε θεωρητικό επίπεδο τη μέθοδο εξήγησης που χρησιμοποιεί η NEON ιδίως στις περιπτώσεις όπου K=2 όπου υπάρχει πλήρης ευθυγράμμιση με τη συγκεκριμένη μαθηματική θεμελίωση.Η θεωρία Shapley βασίζεται στη θεωρία παιγνίων δίνοντας απάντηση στο ποιά είναι η δίκαιη συνεισφορά κάθε παίκτη (χαρακτηριστικό) , στην τελική αξία (έξοδος).Παρέχει ένα αξιωματικό τρόπο κατανομής ευθύνης.Άρα η εξήγηση δεν είναι απλώς εμπειρική αλλά έχει πλήρη μαθηματική θεμελίωση με βάση τη θεωρία Shapley value.Αξιοσημείωτο είναι ότι η κατανομή της NEON ικανοποιεί την ικανότητα διατήρισης συνεισφοράς , είναι συνεχής και ανεξάρτητη από μετατοπίσεις στο χώρο εισόδου.Επίσης, αξίζει να τονίσουμε πως η διαδικασία απόδοσης συνάφειας υλοποιείται ως προς τους ανταγωνιστές και οχι αποκλειστικά ως προς την ομάδα που τοποθετήθηκε το σημείο , ουσιαστικά η συνάφεια ενός χαρακτηριστικού υπολογίζεται με βάση κατά πόσο συμβάλλει στο να ξεχωρίσει από τους ανταγωνιστές.



Σχήμα 5: Αναπαράσταση των τεσσάρων σημείων στο επίπεδο .Εδώ γίνεται ξεκάθαρη η θέση τους τόσο στο άξονα x όσο και στον άξονα y .

4.6 Νευρονοποίηση αλγορίθμου K-μέσων σε βαθιά νευρωνικά δίκτυα.

Η εφαρμογή του αλγορίθμου σε ένα βαθύ νευρωνικό δίκτυο γίνεται με τη χρήση ενός χάρτη χαρακτηριστικών που δίνεται ρητά ως ακολουθία αντιστοιχίσεων των επιπέδων $\Psi_(x) = \Psi_L \circ \cdots \circ \Psi_1(x)$ και ο χάρτης χαρακτηριστικών εκπαιδεύεται μέσω backpropagation για να παράγει την επιθυμητή δομή συμπλέγματος. Αν και έχουν προταθεί διάφορες διατυπώσεις βαθιών K-μέσων , στη περίπτωση αυτή ο αλγόριθμος της ομαδοποίησης θα δώσει λύση αξιοποιώντας τις αποστάσεις στον χώρο χαρακτηριστικών. Άρα χρησιμοποιώντας το ίδιο μοντέλο εκχώρησης σε ομάδες όπως με τον απλό K-μέσων αλλά αυτή τη φορά στον χώρο χαρακτηριστικών, έτσι αποφασίζουμε άν ένα σημείο ανήκει σε μια ομάδα c αν :

$$\forall k \neq c, \quad \|\Psi(x) - \mu_c\|^2 < \|\Psi(x) - \mu_k\|^2 \tag{17}$$

4.6.1 Αρχιτεκτονική δικτύου.

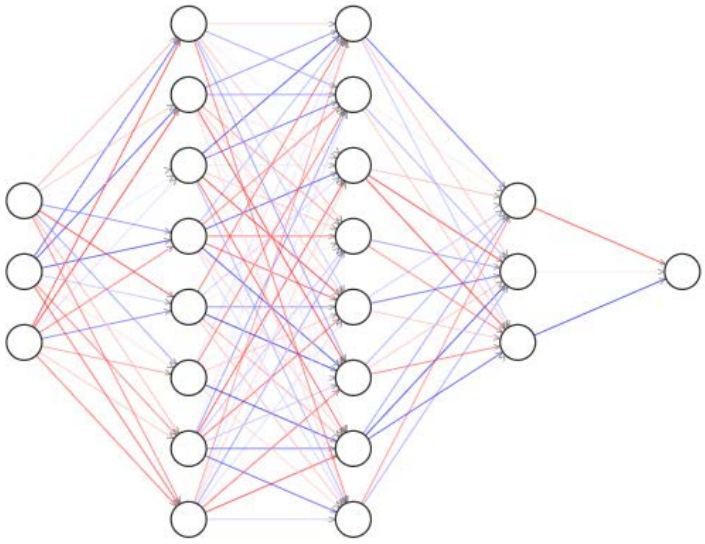
Έτσι η αρχιτεκτονική του δικτύου είναι:

$$a = \Psi_L \circ \cdots \circ \Psi_1(x) \quad (\text{Επίπεδο 1...L}) \tag{18}$$

$$h_k = w'_k \cdot a + b_k \quad (\text{Επίπεδο L} + 1) \tag{19}$$

$$f_c = \min\{h_k\} \quad \text{για } k \neq c \quad (\text{Επίπεδο L} + 2) \tag{20}$$

Όπως φαίνεται στη παραπάνω αρχιτεκτονική τα επίπεδα 1 έως L αντιστοιχούν σε οποιοδήποτε είδος νευρωνικού δικτύου είτε αφορά ένα δίκτυο εμπρόσθιας τροφοδότησης είτε κάποιο συνελικτικό δίκτυο.Αρχικά τα δεδομένα εισέρχονται στο νευρωνικό δίκτυο αντιστοιχίζεται σε κάποιον χώρο χαρακτηριστικών μέσω μη γραμμικών πράξεων και στο αποτέλεσμα αυτού εφαρμόζεται η τεχνική που είδαμε αρχικά για τον απλό K-μέσων αλλά πλέον στον χώρο χαρακτηριστικών και όχι στα αρχικά δεδομένα.Άρα τα δύο επίπεδα που αντιστοιχούν στην αναπαράσταση του αλγορίθμου K-μέσων τοποθετούνται στο τέλος του νευρωνικού προηγούμενου δικτύου , μια τέτοια αρχιτεκτονική φαίνεται στο νευρωνικό δίκτυο του σχήματος 6. Τονίζεται ότι ο υπολογισμός των βαρών και των μεροληπτικών όρων είναι ίδιος με αυτούς που αναφέρονται στους τύπους (12) και (13). Επίσης για τον υπολογισμό της συνάφειας στα δύο ανώτερα στρώματα χρησιμοποιείται η τεχνική LRP οι μαθηματικοί τύποι (14) και (16) , όπως παρατηρούμε από το επίπεδο 1 έως L +1 έχουμε τη μορφή τυπικών νευρωνικών δικτύων οπότε για τη διάδοση της συνάφειας μέσω αυτών των στρωμάτων και προς τα χαρακτηριστικά εισόδου χρησιμοποιούνται κανόνες διάδοσης που έχουν σχεδιαστεί για κάθε είδος νευρωνικού δικτύου όπως τα συνελικτικά και τα LSTM .



Σχήμα 6: Στο συγκεκριμένο σχήμα αναπαριστάται ένα τυχαίο βαθύ νευρωνικό δίκτυο σύμφωνα με την αρχιτεκτονική που αναφέρθηκε παραπάνω.

Στο νευρωνικό δίκτυο του σχήματος 6 οι νευρώνες εισόδου είναι τρείς και ακολουθούν τρία κρυφά επίπεδα , τα δύο πρώτα αποτελούνται από οχτώ νευρώνες , τα δεδομένα διαπερνόντας από αυτά τα κρυφά επίπεδα έχουν υποστεί μετασχηματισμούς σε έναν χώρο χαρακτηριστικών και στη συνέχεια εισέρχονται στους τρείς νευρώνες του τρίτου κρυφού επιπέδου που αντιστοιχεί στην ομαδοποίηση K-μέσων με K=3 όπου K αριθμός ομάδων .Τέλος το επίπεδο εξόδου αποτελείται από έναν νευρώνα που όπως έχει αναφερθεί κάνει την επιλογή της μικρότερης απόστασης για την τοποθέτηση του σημείου σε ομάδα.Η επιλογή του αριθμού των νευρώνων σε όλα τα επίπεδα και στον αριθμό των ομάδων είναι τυχαία με σκοπό τη χρήση τους στο συγκεκριμένο παράδειγμα για την αποτύπωση ενός βαθιού νευρωνικού δικτύου.

5 Εφαρμογές και μεθοδολογία.

Θα ακολουθήσουν παραδείγματα εφαρμογών της συγκεκριμένης τεχνικής , αρχικά θα γίνει υλοποίηση του αλγορίθμου στο σύνολο δεδομένων του iris με χρήση του απλού νευρωνικού δικτύου που αναπαριστά τον αλγόριθμο K-μέσων , στη συνέχεια ακολουθεί δεύτερο παράδειγμα στο ίδιο σύνολο δεδομένων αλλά με τη δημιουργία και χρήση βαθιού νευρωνικού δικτύου μετασχηματίζοντας τα δεδομένα εισόδου σε χώρο χαρακτηριστικών και τέλος ακολουθεί τρίτο παράδειγμα στο σύνολο δεδομένων BBC news με σκοπό την ομαδοποίηση των κειμένων και την απόδοση συνάφειας στις λέξεις που τα απαρτίζουν , εδώ γίνεται χρήση της απλής αρχιτεκτονκής του αλγορίθμου K-μέσων. Και στις τρείς εφαρμογές ύστερα από τον υπολογισμό της συνάφειας κάθε χαρακτηριστικού για την οπτικοποίηση του αποτελέσματος προτείνεται η χρήση ραβδογράμματος που κάθε στήλη θα αναπαριστά ένα χαρακτηριστικό.Στα δύο πρώτα παραδείγματα η τελική συνεισφορά κάθε χαρακτηριστικο-

ύ και κατέπέκταση η οπτικοποίηση του στο ραβδόγραμμα γίνεται μέσω του υπολογισμού της μέσης τιμής μεταξύ των συναφειών κάθε χαρακτηριστικού , ενώ στη τρίτη εφαρμογή γίνεται καταμέτρηση της συχνότητας των λέξεων με υψηλότερη συνάφεια και προβολή των δέκα λέξεων με μεγαλύτερη συχνότητα εμφάνισης και μεγαλύτερη συνάφεια.

5.1 Εφαρμογή του απλού νευρωνικού δικτύου του K-μέσων στο σύνολο δεδομένων Iris.

Το σύνολο δεδομένων Iris περιλαμβάνει τρία διαφορετικά είδη λουλουδιών από 50 δείγματα το καθένα άρα συνολικά 150 δείγματα καθώς και κάποια επιπλέον χαρακτηριστικά για τα λουλούδια , ένα είδος λουλουδιών διαχωρίζεται γραμμικά από τα άλλα δύο αλλά τα άλλα δύο δεν διαχωρίζονται γραμμικά μεταξύ τους.Οι στήλες-χαρακτηριστικά του συγκεκριμένου συνόλου δεδομένων είναι :

1. id.
2. Sepal Length cm.
3. Sepal Width cm.
4. Petal Length cm.
5. Petal Width cm.
6. Species.

Ακολουθεί στιγμιότυπο από τις πέντε πρώτες εγγραφές ύστερα από τη φόρτωση του συνόλου δεδομένων.

Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
1	5.1	3.5	1.4	0.2	Iris-setosa
2	4.9	3.0	1.4	0.2	Iris-setosa
3	4.7	3.2	1.3	0.2	Iris-setosa
4	4.6	3.1	1.5	0.2	Iris-setosa
5	5.0	3.6	1.4	0.2	Iris-setosa

Πίνακας 1: Δείγμα από τις πέντε πρώτες εγγραφές του συνόλου δεδομένων Iris με τη μορφή dataframe.

Προεπεξεργασία δεδομένων:

Αρχικά γίνεται αφαίρεση της στήλης Id διότι είναι απλά μία αρίθμιση των δειγμάτων χωρίς να συμβάλει με ουσιαστικό τρόπο στην υλοποίηση της διαδικασίας και της ομαδοποίησης.Επίσης θα γίνει αφαίρεση και της στήλης Species διότι είναι μια κατηγορική μεταβλητή και για την εφαρμογή του K-μέσων χρειάζονται αριθμητικές μεταβλητές για τους μαθηματικούς υπολογισμούς.Στο πλαίσιο της προεπεξεργασίας είναι και η εφαρμογή του μετασχηματισμού MinMaxScaler.

Μετασχηματισμός MinMaxScaler:

Είναι μια δημοφιλής τεχνική κανονικοποίησης δεδομένων με σκοπό τον μετασχηματισμό των χαρακτηριστικών εντός ενός συγκεκριμένου εύρους τιμών το [0,1] στη συγκεκριμένη περίπτωση.Ο μετασχηματισμός αυτός εξασφαλίζει ότι όλα τα χαρακτηριστικά συμβάλουν το ίδιο στο μοντέλο κλιμακώνοντας τα στο ίδιο εύρος , στην οπτικοποίηση τους και βελτιώνει τη γενίκευση του μοντέλου . Ο μαθηματικός τύπος που εφαρμόζεται για κάθε τιμή είναι :

$$Xscaled = \frac{X - Xmin}{Xmax - Xmin} \tag{21}$$

Ουσιαστικά για κάθε τιμή X που θέλουμε να κλιμακωθεί αφαιρείται η μικρότερη τιμή του συνόλου δεδομένων και η διαφορά αυτή διαιρείται με τη διαφορά που προκύπτει απο την αφαίρεση της μικρότερης τιμής του συνόλου δεδομένων από τη μεγαλύτερη τιμή του συνόλου δεδομένων.

Εφαρμογή του αλγορίθμου συσταδοποίησης K-μέσων.

Στο σημείο αυτό ορίζεται ο επιθυμητός αριθμός ομάδων δηλαδή $K = 3$ γιατί σύμφωνα με το σύνολο δεδομένων τρείς είναι οι διαφορετικές κατηγορίες φυτών.Με αυτό το βήμα έχουν δημιουργηθεί τα τελικά κέντρα των τριών ομάδων καθώς και η ετικέτα κάθε δεδομένου.Επίσης οι αρχικές και πραγματικές ετικέτες Spacies για κάθε σημείο έχουν αποθηκευτεί ξεχωριστά ως πίνακας-στήλη στη μεταβλητή Xtrue ,έτσι μπορούμε να αξιολογήσουμε την ποιότητα της ομαδοποίησης με τη χρήση του δείκτη ARI που αναλύθηκε προηγουμένως. Η τιμή του είναι $ari = 0.7163421126838476$.

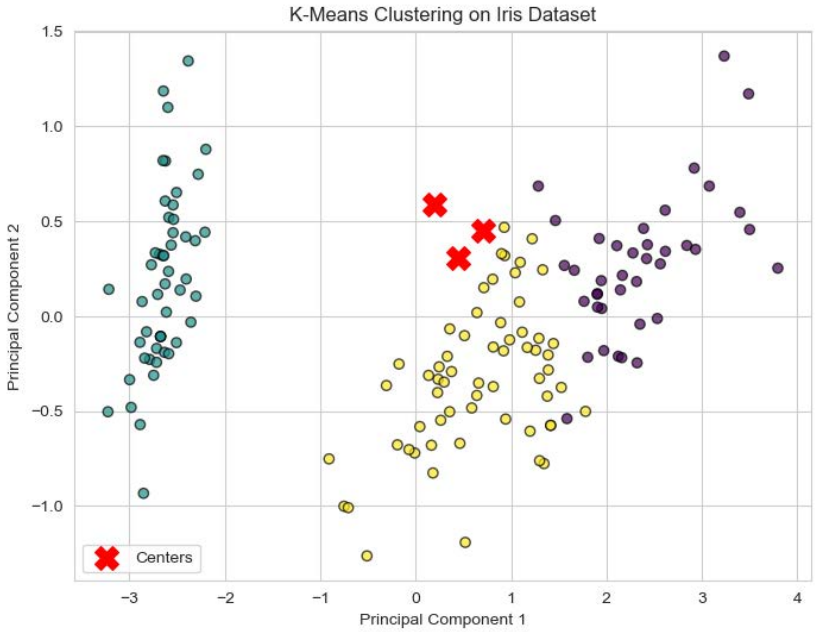
Εφαρμογή PCA για μείωση της διάστασης.

Η άμεση οπτικοποίηση του αποτέλεματος της ομαδοποίησης των δεδομένων όπως φαίνεται στο σχήμα 7 λόγω τεσσάρων διαστάσεων-χαρακτηριστικών δεν είναι δυνατή επομένως θα πρέπει να γίνει μείωση της διάστασης σε δύο διαστάσεις , αυτό μπορεί να επιτευχθεί με χρήση του αλγορίθμου PCA ή ανάλυση κύριων συνιστωσών διατηρώντας παράλληλα όσο το δυνατόν περισσότερες από τις αρχικές πληροφορίες .Πιο συγκεκριμένα ο αλγόριθμος στοχεύει να βρεί μια αναπαράσταση χαμηλών διαστάσεων διατηρώντας παράλληλα μεγάλη διακύμανση μεταξύ των δεδομένων διότι με τη διατήριση της διακύμανσης διατηρούμε και σημαντικές πληροφορίες.Αναλυτικά:

1. Κεντράρισμα δεδομένων.
2. Υπολογισμός πίνακα συνδιακύμανσης.
3. Αποσύνθεση ιδιοτιμών.
4. Επιλογή κύριων συνιστωσών.
5. Προβολή των δεδομένων στις επιλεγμένες συνιστώσες.

Τονίζεται ότι οι ιδιοτιμές ποσοτικοποιούν τη διακύμανση που αναγράφεται σε κάθε κύρια συνιστώσα και τα ιδιοδιανύσματα υποδεικνύουν τις κατευθύνσεις της μέγιστης διακύμανσης στα δεδομένα.Έτσι για την επιλογή των κύριων συνιστωσών στη συγκεκριμένη περίπτωση θέλουμε 2 συνιστώσες , από τις ταξινομημένες ιδιοτιμές σε φθίνουσα σειρά

επιλέγονται οι δυο πρώτες με τη μεγαλύτερη διακύμανση.Ακολυθεί η οπτικοποίηση της ομαδοποίησης K-μέσων στο σύνολο δεδομένων έπειτα από την εφαρμογή ανάλυσης κύριων συνιστωσών:



Σχήμα 7: Στο σχήμα αναπαριστάται η ομαδοποίηση K-μέσων στο σύνολο δεδομένων Iris .

Στο διάγραμμα που απεικονίζεται στο σχήμα 7 φαίνεται το αποτέλεσμα της ομαδοποίησης K-μέσων ύστερα απο την εφαρμογή ανάλυσης κύριων συνιστωσών στα αρχικά δεδομένα για την μείωση της διάστασης με σκοπό την οπτικοποίηση της διαδικασίας, επίσης φαίνονται τα κέντρα που υπολογίστηκαν καθώς και με διαφορετικούς χρωματισμούς φαίνονται οι τρείς διαφορετικές ομάδες που προκύπτουν.

5.1.1 Νευρονοποίηση για το συγκεκριμένο παράδειγμα

Για να ξεκινήσει η συγκεκριμένη διαδικασία ήταν απαραίτητο πρώτα να εφαρμοστεί ο αλγόριθμος K-μέσων στα κανονικοποιημένα δεδομένα διότι για τους υπολογισμούς των βαρών και των μεροληπτικών όρων απαιτούνται τα κέντρα που προκύπτουν στο τέλος της ομαδοποίησης.Επειδή όπως είδαμε τα χαρακτηριστικά είναι τέσσερα και οι ομάδες τρείς , τα κέντρα που προκύπτουν είναι διάστασης (3, 4).Τα κέντρα που προέκυψαν είναι :

$$\begin{bmatrix} 0.7073 & 0.4509 & 0.7970 & 0.8248 \\ 0.1961 & 0.5908 & 0.0786 & 0.0600 \\ 0.4413 & 0.3074 & 0.5757 & 0.5492 \end{bmatrix}$$

Με βάση αυτές τις τιμές που προέκυψαν και με την υλοποίηση των τύπων (12) και (13) προκύπτει ο πίνακας με τις τιμές των βαρών και τις τιμές των μεροληπτικών όρων αντίστοιχα.

Βάρη:

$$\begin{bmatrix} 1.0223 & -0.2800 & 1.4368 & 1.5296 \\ 0.5320 & 0.2870 & 0.4427 & 0.5512 \\ -0.4903 & 0.5669 & -0.9941 & -0.9784 \end{bmatrix}$$

Παρατηρούμε ότι η διάστασή τους είναι (3,4) γεγονός λογικό αφού κάθε δεδομένο που εισέρχεται στο νευρωνικό αποτελείται από τέσσερα χαρακτηριστικά καθώς και σύμφωνα με το τύπο του άρθρου για τον υπολογισμό των βαρών λαμβάνουμε μόνο τη διαφορά ανάμεσα στην ομάδα στόχο και τον ανταγωνιστή της άρα τρία βάρη όσοι και οι νευρώνες .

Μεροληπτικοί όροι:

$$\begin{bmatrix} -1.6217 & -1.0968 & 0.5249 \end{bmatrix}$$

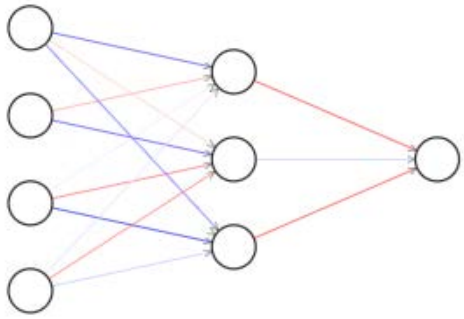
Η διάσταση του είναι (1,3) γιατί τρείς νευρώνες στο κρυφό επίπεδο όπως αναφέρθηκε προηγουμένως .

Αρχιτεκτονική νευρωνικού:

Εφόσον έχουν υπολογιστεί πλέον τα βάρη και οι μεροληπτικοί όροι δημιουργείται το νευρωνικό που θα αναπαραστήσει την συγκεκριμένη ομαδοποίηση .Έτσι θα έχουμε τέσσερις νευρώνες εισόδου , τρείς νευρώνες στο κρυφό επίπεδο όσο το K=3 και επειδή όπως έχει αναφερθεί το επίπεδο εξόδου αποτελείται από ένα νευρώνα ο οποίος δεν είναι νευρώνας με τη κλασική έννοια δηλαδή δεν έχει βάρη ούτε μεροληπτικό όρο απλά επιλέγει τη μικρότερη απόσταση ,δεν περιλαμβάνεται στην αρχιτεκτονική του δικτύου.Πιο συγκεκριμένα για κάθε σημείο θα προκύπτουν τρείς αποστάσεις οι οποίες θα αποθηκευτούν σε ένα πίνακα διάστασης (150,3) και σε αυτόν θα γίνει ανά στήλες η επιλογή της μικρότερης τιμής και με βάση αυτή θα τοποθετούνται τα σημεία σε ομάδες.Το μοντέλο που αναπαριστάται στο σχήμα 8 επιβεβαιώνεται η αρχιτεκτονική του μέσω της παρακάτω σχέσης:

$$KmeansModel(\ (linear) : Linear(in\,features = 4, out\,features = 3, bias = True)$$

Το μοντέλο που δημιουργήθηκε δεν εκπαιδεύεται και τα βάρη παραμένουν σταθερά ,άρα γίνεται μόνο ένα πέρασμα των δεδομένων από το δίκτυο.



Σχήμα 8: Η αρχιτεκτονική του δικτύου για την ομαδοποίηση K-μέσων για K=3 στο σύνολο δεδομένων Iris με τέσσερα χαρακτηριστικά και νευρώνες εισόδου , τρείς νευρώνες στο κρυφό επίπεδο και έναν νευρώνα εξόδου.

Παρατίθενται στιγμιότυπο από τον πίνακα των αποστάσεων που αναφέρθηκε με διάσταση (150,3) , ο πίνακας αφορά τις αποστάσεις που έχει καθένα από τα 150 σημεία που εισέρχονται στο νευρωνικό δίκτυο από τα τρία κέντρα που υπολογίστηκαν αρχικά σύμφωνα με τον μαθηματικό τύπο (10).Στο στιγμιότυπο αναπαρίσταται οι αποστάσεις για τα τις πέντε πρώτες εγγραφές:

$$\begin{bmatrix} -1.4084 & -0.7463 & 0.6621 \\ -1.4068 & -0.8356 & 0.5712 \\ -1.5113 & -0.8487 & 0.6626 \\ -1.4793 & -0.8605 & 0.6189 \\ -1.4484 & -0.7491 & 0.6993 \end{bmatrix}$$

Σύμφωνα με τη λογική που έχουμε περιγράψει το πρώτο σημείο θα τοποθετηθεί στην ομάδα όπου απέχει μικρότερη απόσταση από το κέντρο της .Παρατηρούμε ότι στη τρίτη στήλη της υπάρχει η μικρότερη απόσταση άρα θα τοποθετηθεί σε αυτή την ομάδα ,το ίδιο συμβαίνει και για τις υπόλοιπες τέσσερις εγγραφές άρα συμπεραίνουμε ότι τα πέντε πρώτα σημεία ανήκουν στην ίδια ομάδα.Επίσης σύμφωνα πάλι με τον τύπο (10) η τιμή του fc είναι η δεύτερη μικρότερη απόσταση ,όπου πάλι παρατηρούμε ότι στο συγκεκριμένο στιγμιότυπο οι τιμές αυτές βρίσκονται στη δεύτερη στήλη.Εφόσον γίνει επιλογή της ομάδας που ανήκει κάθε σημείο στη συνέχεια γίνεται επιλογή του fc για τα ίδια σημεία και οι τιμές αυτές αποθηκεύονται σε μία λίστα διάστασης (150,1) διότι έχουμε 150 εγγρφές και γίνεται επιλογή μίας τιμής για καθένα.

Έτσι έχουμε:

$$\begin{bmatrix} -0.7463 \\ -0.8356 \\ -0.8487 \\ -0.8605 \\ -0.7491 \end{bmatrix}$$

Το διάνυσμα αυτό είναι απόσπασμα της λίστας με τις τιμές fc των πέντε πρώτων εγγραφών.Με το απόσπασμα αυτό επιβεβαιώνεται η διαδικασία που έχει αναλυθεί και ο μαθηματικός τύπος (11).Πιο συγκεκριμένα συγκρίνοντας τις τιμές της παραπάνω λίστας με τον πίνακα αποστάσεων από τα τρία κέντρα παρατηρούμε ότι σε καθένα από τα πέντε δείγματα η τιμή fc είναι η δεύτερη μικρότερη τιμή αφού η πρώτη αφορά τη τοποθέτηση του σημείου στην ομάδα αυτή με τη μικρότερη απόσταση ,για το συγκεκριμένο απόσπασμα των πέντε δειγμάτων βρίσκεται στη τρίτη στήλη γιαυτό τον λόγο ανήκουν στην ίδια ομάδα και επίσης τυχαίνει η δεύτερη μικρότερη τιμή να βρίσκεται στη δεύτερη στήλη του πίνακα αποστάσεων γεγονός το οποίο επιβεβαιώνεται από τις τιμές της παραπάνω λίστας.Έχοντας ολοκληρώσει τη διαδικασία της ομαδοποίησης μέσω του νευρωνικού δικτύου η τιμή του δείκτη ari είναι $ari = 0.559629277257056 9$.Παρατηρούμαι ότι η απόδοση της ομαδοποίησης έχει μειωθεί αρκετά σε σχέση με την απόδοση του απλού αλγορίθμου K-μέσων στο ίδιο σύνολο δεδομένων,το γεγονός αυτό όμως σύμφωνα με τη πρόταση του άρθρου δεν αποτελεί πρόβλημα αφού η διαδικασία αυτή όπως αναφέρθηκε και αρχικά δεν έχει ως στόχο την αύξηση της απόδοσης και τη βελτίωση της ομαδοποίησης αλλά αποτελεί έναν τρόπο αναπαράστασης της απλής ομαδοποίησης K-μέσων με κύριο σκοπό την απόδοση συνάφειας και σημαντικότητας στα χαρακτηριστικά εισόδου.

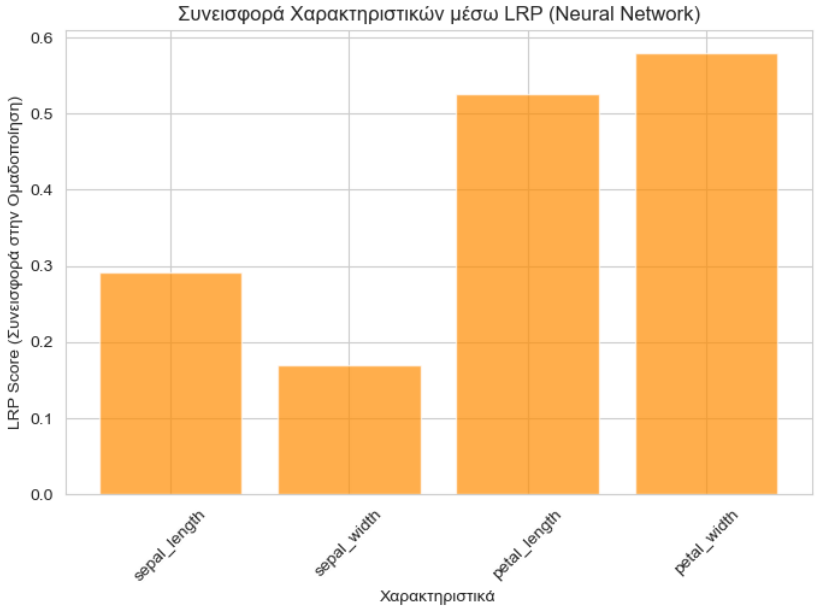
5.1.2 Απονομή συνάφειας στα χαρακτηριστικά εισόδου , διαδικασία LRP.

Εφόσον έγινε η επιλογή των τιμών fc για κάθε δείγμα ξεκινάει η διαδικασία της απόδοσης συνάφειας στα χαρακτηριστικά εισόδου (διαδικασία LRP) δηλαδή με την μέθοδο αυτή θα διαπιστωθεί ποιό από τα τέσσερα χαρακτηριστικά κάθε δείγματος συνέβαλε περισσότερο για την τοποθέτηση του σε μία ομάδα. Για την υλοποίηση των μαθηματικών τύπων (14),(15),(16) έχουν δημιουργηθεί κατάλληλες συναρτήσεις αξιοποιώντας προκειμένου να εκτελεστούν οι μαθηματικές πράξεις με ακρίβεια και να ακολουθούνται αυστηρά οι μαθηματικοί τύποι που προτείνει ο συγκεκριμένος αλγόριθμος.Έτσι προκύπτει ένας πίνακας διάστασης (150,4) με τις συνάφειες των τεσσάρων χαρακτηριστικών για καθένα από τα 150 δείγματα .

$$\begin{bmatrix} -2.06085641e-01 & -2.85553302e-01 & -8.73270375e-01 & -8.73871033e-01 \\ -3.18308130e-01 & -3.18039001e-02 & -1.07643808e+00 & -1.08024668e+00 \\ -3.62939385e-01 & -1.46424463e-01 & -1.03776976e+00 & -9.99101792e-01 \\ -4.21363510e-01 & -9.30063898e-02 & -1.00934278e+00 & -1.05771069e+00 \\ -2.21385925e-01 & -3.34980181e-01 & -8.45590950e-01 & -8.45288570e-01 \end{bmatrix}$$

Ο παραπάνω πίνακας είναι ένα απόσπασμα από τον συνολικό πίνακα συναφειών των χαρακτηριστικών εισόδου για τα πέντε πρώτα δείγματα .Σε ολόκληρο τον πίνακα προκύπτουν εξίσου τόσο θετικές όσο και αρνητικές τιμές LRP , η ύπαρξη των θετικών τιμών δείχνει ποιό χαρακτηριστικό έχει τη μεγαλύτερη συμβολή στη πρόβλεψη του μοντέλου ενώ η ύπαρξη αρνητικών τιμών δείχνει ποιό χαρακτηριστικό λειτουργεί αποτρεπτικά στη πρόβλεψη του μοντέλου. Στο συγκεκριμένο απόσπασμα τυχαίνει οι τιμές συνάφειας για όλα τα χαρακτηριστικά και για τα πέντε δείγματα να είναι αρνητικές,για παράδειγμα στο πρώτο δείγμα αν λάβουμε υπόψιν τις απόλυτες τιμές των τιμών παρατηρούμε ότι το χαρακτηριστικό της τέταρτης στήλης έχει τη μεγαλύτερη τιμή άρα μεγαλύτερη συμβολή .Το γεγονός όμως ότι η τιμή είναι αρνητική σημαίνει ότι το χαρακτηριστικό αυτό αντιτίθεται

στη πρόβλεψη του μοντέλου.Για την αναπαράσταση της σημαντικότητας των χαρακτηριστικών θα χρησιμοποιηθεί ραβδόγραμμα για τα τέσσερα χαρακτηριστικά , πριν τη δημιουργία του ραβδογράμματος θα εφαρμοστεί στις τιμές συνάφειας που υπολογίστηκαν η απόλυτη τιμή και στη συνέχεια θα υπολογιστεί η μέση τιμή στον ίδιο πίνακα σε κάθε στήλη.Έτσι το αποτέλεσμα που προκύπτει είναι το εξής:



Σχήμα 9: Ραβδόγραμμα με τη μέση τιμή των τιμών LRP για κάθε χαρακτηριστικό .

Στο διάγραμμα του σχήματος 9 αναπαριστάται η συνεισφορά κάθε χαρακτηριστικού στην διαδικασία της ομαδοποίησης αφού πρώτα έχει εφαρμοστεί η απόλυτη τιμή και στη συνέχεια η μέση τιμή σε κάθε στήλη του πίνακα συναφειών.Επίσης γίνεται ξεκάθαρο το γεγονός ότι τα δύο χαρακτηριστικά με τη μεγαλύτερη συμβολή στην διαδικασία ομαδοποίησης Κ-μέσων των φυτών του αρχικού συνόλου δεδομένων σε τρείς ομάδες είναι το τρίτο και τέταρτο χαρακτηριστικό δηλαδή το μήκος και το πλάτος πετάλου.Το αποτέλεσμα αυτό φαίνεται λογικό , δηλαδή ότι το πλάτος και το ύψος του πετάλου συμβάλει περισσότερο στη διάκριση μεταξύ των διαφορετικών κατηγοριών των φυτών σε σχέση με το μήκος και το πλάτος του σεπάλου.

5.2 Εφαρμογή του αλγορίθμου Κ-μέσων στο σύνολο δεδομένων Iris με αναπαράσταση μέσω βαθιού νευρωνικού δικτύου.

Στο συγκεκριμένο παράδειγμα θα πραγματοποιηθεί η υλοποίηση του αλγορίθμου Κ-μέσων στο ίδιο σύνολο δεδομένων με τη διαφορά ότι τα αρχικά δεδομένα θα μετασχηματιστούν μέσω του περάσματος τους από ένα βαθύ νευρωνικό δίκτυο εμπρόσθιας τροφοδότισης.Πιο αναλυτικά τα αρχικά δεδομένα θα περάσουν από ένα νευρωνικό δίκτυο που θα οριστεί στη συνέχεια και στα αποτελέσματα που θα εξαχθούν από το πρώτο νευρωνικό δίκτυο θα εφαρμοστεί ο απλός αλγόριθμος Κ-μέσων για Κ=3.Με τα τρία διαφορετικά κέντρα που θα προκύψουν στο τέλος της διαδικασίας θα υπολογιστούν τα βάρη και οι μεροληπτικοί όροι,όπου πάλι θα παραμείνουν σταθερά καθώς ο νευρονοποιημένος αλγόριθμος Κ-μέσων

δεν εκπαιδεύεται.Αφού έχει δημιουργηθεί πλέον το νευρωνικό δίκτυο του αλγορίθμου με τα βάρη και τους μεροληπτικούς όρους που προέκυψαν από την εφαρμογή της ομαδοποίησης στα μετασχηματισμένα δεδομένα ,ενώνονται τα δύο νευρωνικά δίκτυα σχηματίζοντας ένα βαθύ νευρωνικό δίκτυο που δέχεται ως είσοδο τα αρχικά μη μετασχηματισμένα δεδομένα.

Προεπεξεργασία δεδομένων:

Στο στάδιο αυτό πραγματοποιείται η ίδια διαδικασία με το προηγούμενο παράδειγμα , δηλαδή γίνεται φόρτωση του αρχείου με τα δεδομένα του συνόλου δεδομένων Iris , αφαιρούνται οι στήλες Id και Species και εφαρμόζουμε την κανονικοποίηση Min-Max scaler όπως ακριβώς και στο προηγούμενο παράδειγμα.

5.2.1 Δημιουργία πρώτου νευρωνικού δικτύου.

Στο σημείο αυτό τα αρχικά κανονικοποιημένα δεδομένα θα περάσουν από το νευρωνικό δίκτυο που θα οριστεί και στα αποτελέσματα της εξόδου του θα εφαρμοστεί ο αλγόριθμος K-μέσων.Το συγκεκριμένο νευρωνικό δίκτυο που θα οριστεί θα αποτελείται από τέσσερεις νευρώνες στο επίπεδο είσοδου αφού η διάσταση κάθε δείγματος είναι τέσσερα όπου τέσσερα τα χαρακτηριστικά,στη συνέχεια ορίζεται ένα κρυφό επίπεδο δεκαέξι νευρώνων ώστε τα αρχικά χαρακτηριστικά να μετασχηματιστούν και τέλος ένα επίπεδο εξόδου τεσσάρων νευρώνων . Η επιλογή του αριθμού των νευρώνων στο επίπεδο εξόδου δεν είναι τυχαία διότι η έξοδος αυτή θα αποτελεί είσοδο στο δεύτερο νευρωνικό που θα αντιστοιχεί στην ομάδοποίηση και θα ενωθεί με το πρώτο.Όσον αφορά το πρώτο νευρωνικό δίκτυο που θα αναλυθεί και αποτελεί ένα είδος κωδικοποιητή θα πραγματοποιηθεί εκπαίδευση και για επιβεβαίωση της αρχιτεκτονικής που περιγράφηκε παρατίθεται το παρακάτω απόσπασμα :

$$Deep\begin{pmatrix} layers : Sequential \\ (0) : Linear(in\_features = 4, out\_features = 16, bias = True) \\ (1) : ReLU() \\ (2) : BatchNorm1d(16, \epsilon = 10^{-5}, momentum = 0.1, \\ \quad affine = True, track\_running\_stats = True) \\ (3) : Linear(in\_features = 16, out\_features = 4, bias = True) \\ (4) : ReLU() \end{pmatrix}$$

Στην παραπάνω αρχιτεκτονική παρατηρείται ότι το νευρωνικό δίκτυο είναι γραμμικό , πλήρες συνδεδεμένο.Γίνεται ξεκάθαρο πως η είσοδος είναι διάστασης τεσσάρων νευρώνων ακολουθεί ένα κρυφό επίπεδο δεκαέξι νευρώνων όπου μετασχηματίζει την είσοδο αυξάνοντας τη διαστατικότητα του χώρου αναπαράστασης και το οποίο συνδέεται με ένα επίπεδο εξόδου τεσσάρων νευρώνων.Η συνάρτηση ενεργοποίησης που χρησιμοποιείται στο νευρωνικό δίκτυο είναι η Relu και η κανονικοποίηση Batch Normalization για τα οποία ισχύει:

Συνάρτηση ενεργοποίησης ReLu:

Ως συνάρτηση ενεργοποίησης χρησιμοποιείται η ReLu προσθέτοντας μη γραμμικότητα στις πράξεις βοηθώντας το μοντέλο να μάθει πιο πολύπλοκα μοτίβα , η συνάρτηση αυτή εκτελεί την εξής λειτουργία επιστρέφει μηδέν για τις αρνητικές τιμές που δέχεται και την ίδια την τιμή αν η τιμή που δέχεται είναι θετική:

$$ReLU(x) = \max(0, x) \tag{22}$$

$$ReLU'(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases} \tag{23}$$

Κανονικοποίηση Batch Normalization:

Η κανονικοποίηση αυτή ομαλοποιεί τις τιμές του προηγούμενου επιπέδου βελτιώνοντας την σταθερότητα της εκπαίδευσης και επιταχύνοντας τη σύγκλιση.Η λειτουργία της για κάθε δείγμα είναι:

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \tag{24}$$

$$y_i = \gamma \hat{x}_i + \beta \tag{25}$$

Όπου :

x_i : Είσοδος στο επίπεδο BatchNorm για το δείγμα i
μ_B : Μέσος όρος της δέσμης (batch mean)
σ_B^2 : Διακύμανση της δέσμης (batch variance)
ϵ : Μικρή σταθερά για αποφυγή διαίρεσης με το μηδέν
$\hat{x}_i$: Κανονικοποιημένο χαρακτηριστικό (normalized feature)
γ, β : Μαθηματικές παράμετροι εκπαίδευσης (learnable parameters)

Εκπαίδευση του νευρωνικού δικτύου:

Για την εκπαίδευση του συγκεκριμένου νευρωνικού δικτύου θα χρειαστούμε 180 εποχές όπου το μοντέλο προσπαθεί να αναπαράγει τα ίδια δεδομένα με τα αρχικά , χρησιμοποιείται ο βελτιστοποιητής Adam για την εκπαίδευση και ως συνάρτηση απώλειας η Mean Squared Error (MSE loss) η οποία μετρά τη διαφορά μεταξύ εισόδου και εξόδου του δικτύου.Σκοπός είναι αυτή η συνάρτηση απώλειας να είναι όσο το δυνατόν μικρότερη γίνεται . Το δίκτυο εξάγει χαρακτηριστικά απο τα δεδομένα εισόδου χωρίς να προχωρά σε αλλαγές στα βάρη (backpropagation), ουσιαστικά το μοντέλο ανακατασκευάζει τα αρχικά δεδομένα εισόδου.

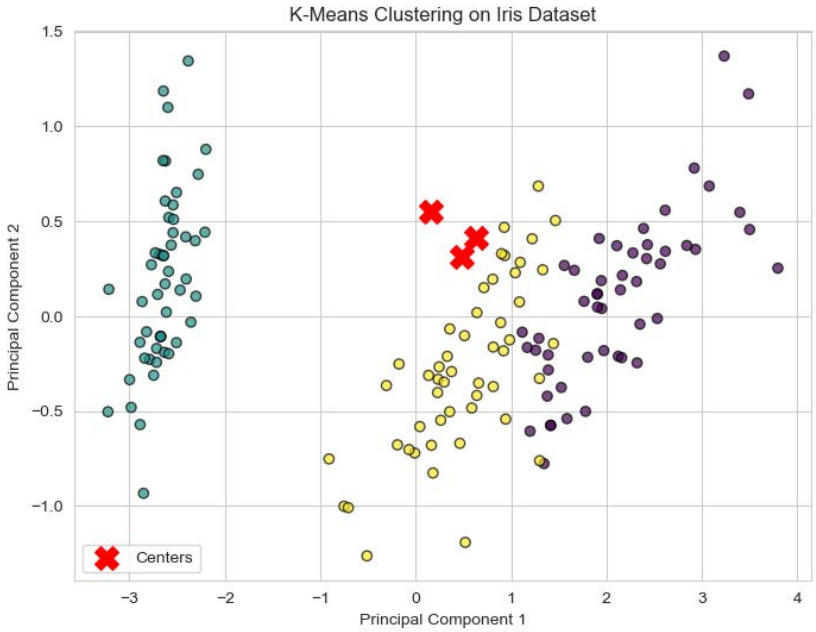
Συνάρτηση απώλειας Mean Squared Error:

Η συνάρτηση απώλειας MSE υπολογίζεται ως ο μέσος όρος των τετραγωνικών διαφορών των πραγματικών τιμών y_i και των προβλεπόμενων τιμών $\hat{y}_i$.Πιο συγκεκριμένα αθροίζεται η τετραγωνική διαφορά ανάμεσα στην πραγματική τιμή και την προβλεπόμενη τιμή για όλα τα δείγματα και στη συνέχεια το άθροισμα αυτό διαιρείται με τον αριθμό των δειγμάτων n. Ο μαθηματικός τύπος που εκτελεί τη συγκεκριμένη διαδικασία είναι:

$$\mathcal{L}_{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{26}$$

Εφαρμογή του αλγορίθμου K-μέσων στα μετασχηματισμένα δεδομένα εισόδου:

Στο σημείο αυτό θα εφαρμοστεί ο αλγόριθμος συσταδοποίησης K-μέσων για K=3 στα δεδομένα που προέκυψαν από την έξοδο του νευρωνικού δικτύου που περιγράφηκε προηγουμένως με σκοπό να προκύψουν τα αντίστοιχα κεντρα όπως φαίνεται και στο σχήμα 10 , για τον υπολογισμό των βαρών και τον μεροληπτικών όρων στη συνέχεια.Η εφαρμογή της ομαδοποίησης στα μετασχηματισμένα δεδομένα δίνει τιμή στον δείκτη $ari = 0.7274709390355676$.Παρατίθεται διάγραμμα αναπαράστασης της ομαδοποίησης στα δεδομένα αφού έχει εφαρμοστεί πρώτα ο αλγόριθμος μείωσης της διάστασης PCA όπως ακριβώς έγινε και στη προηγούμενη εφαρμογή :



Σχήμα 10: Διάγραμμα που αντιστοιχεί στην ομαδοποίηση K-μέσων για K=3.

Στο διάγραμμα του σχήματος 10 απεικονίζεται η ομαδοποίηση K-μέσων για K=3 στα μετασχηματισμένα δεδομένα που προέκυψαν μέσω του περάσματος τους από το νευρωνικό δίκτυο.Έχει εφαρμοστεί ο αλγόριθμος ανάλυσης κύριων συνιστωσών (PCA) για μείωση της αρχικής διάστασης από τέσσερα σε δύο διαστάσεις διατηρώντας μεγάλο μέρος της κύριας πληροφορίας.

5.2.2 Νευρονοποίηση του αλγορίθμου K-μέσων.

Στο σημείο αυτό γίνεται η νευρονοποίηση του αλγορίθμου K-μέσων με K=3 στα μετασχηματισμένα πλέον δεδομένα που προέκυψαν ως αποτέλεσμα στην έξοδο του πρώτου νευρωνικού δικτύου που μόλις περιγράφηκε. Όσον αφορά τον υπολογισμό των κέντρων των ομάδων , των βαρών και των μεροληπτικών όρων ακολουθείται ακριβώς η ίδια διαδικασία που περιγράφηκε και αναλύθηκε στην πρώτη εφαρμογή που γίνεται απευθείας νευρονοποίηση και εφαρμογή του αλγορίθμου στα μη μετασχηματισμένα δεδομένα.Ουσιαστικά στη περίπτωση αυτή η μόνη διαφορά είναι οτι τα αρχικά δεδομένα έχουν μετασχηματιστεί μέσω ενός αρχικού νευρωνικού δικτύου - κωδικοποιητή και ύστερα σε αυτά εφαρμόζεται η ίδια

διαδικασία που αναλύθηκε. Επίσης ίδια παραμένει και η διαδικασία ανάθεσης κάθε εγγραφής σε ομάδα όπως και οι υπολογισμοί των τιμών που χρειάζονται για τη διαδικασία LRP δηλαδή η τιμή fc , επιλέγοντας ουσιαστικά την αμέσως επόμενη μικρότερη τιμή .

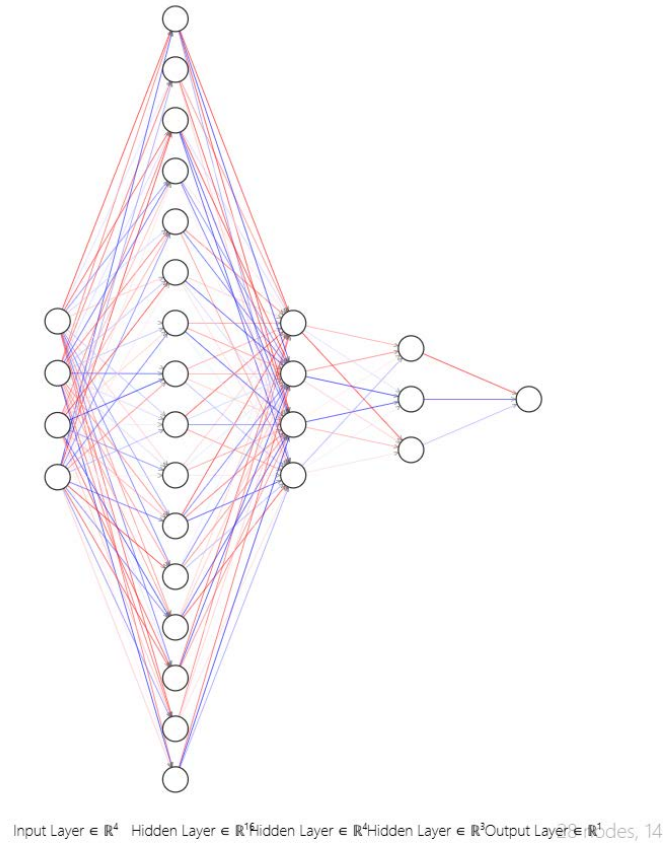
5.2.3 Αρχιτεκτονική και ένωση των δύο δικτύων.

Εφόσον έχουν υπολογιστεί τα βάρη και οι μεροληπτικοί όροι καθώς και τα κέντρα των τριών ομάδων , για την αναπαράσταση της συσταδοποίησης θα γίνει χρήση του ίδιου νευρωνικού δικτύου που έγινε στην προηγούμενη εφαρμογή , χρησιμοποιείται η ίδια αρχιτεκτονική με τη διαφορά πλέον ότι υπολογίστηκαν διαφορετικά βάρη και διαφορετικοί μεροληπτικοί όροι.Άρα για την ομαδοποίηση θα χρησιμοποιηθεί το ίδιο νευρωνικό δίκτυο με ένα κρυφό επίπεδο τριών νευρώνων όσες και οι ομάδες και επίπεδο εισόδου τεσσάρων νευρώνων ώστε να μπορέσει να γίνει η ένωση μεταξύ των δύο δικτύων,επίσης όπως και στη προηγούμενη περίπτωση τα βάρη και οι μεροληπτικοί όροι παραμένουν σταθεροί δεν εκπαιδεύονται .Στο τέλος πάλι θα γίνει επιλογή της μικρότερης απόστασης ώστε να γίνει τοποθέτηση του σημείου στην αντίστοιχη ομάδα. Για επιβεβαίωση της αρχιτεκτονικής που έχει περιγραφεί ακολουθεί η παρακάτω αναπαράσταση της ένωσης των δύο νευρωνικών δικτύων :

$$
DeepKmeans\left(\begin{array}{l} deep : Deep\left(\begin{array}{c} layers : Sequential \\ (0) : Linear(in\_features = 4, \\ \quad out\_features = 16, bias = True) \\ (1) : ReLU() \\ (2) : BatchNorm1d(16, \epsilon = 10^{-5}, momentum = 0.1, \\ \quad affine = True, track\_running\_stats = True) \\ (3) : Linear(in\_features = 16, \\ \quad out\_features = 4, bias = True) \\ (4) : ReLU() \end{array}\right) \\ kmeans\_layer : KmeansModel1\left(\begin{array}{c} linear : Linear(in\_features = 4, \\ \quad out\_features = 3, bias = True) \end{array}\right) \end{array}\right)
$$

Στην παραπάνω αρχιτεκτονική του τελικού βαθιού νευρωνικού δικτύου , η οποία αναπαριστάται και στο σχήμα 11 φαίνεται αναλυτικά η ένωση μεταξύ των δύο νευρωνικών που αρχικά δημιουργήθηκαν ξεχωριστά.Πιο αναλυτικά , τα αρχικά δεδομένα ύστερα από την προεπεξεργασία τους περνάνε από το νευρωνικό δίκτυο Deep το οποίο λειτουργεί ως κωδικοποιήτης δηλαδή μετασχηματίζει τα άρχικά δεδομένα σε έναν χώρο χαρακτηριστικών υψηλότερης διάστασης αφού από τέσσερα που είναι τα χαρακτηριστικά και τέσσερις οι νευρώνες εισόδου μετασχηματίζονται μέσω του κρυφού επιπέδου το οποίο αποτελείται από δεκαέξι νευρώνες σε χώρο υψηλότερης διάστασης .Στη συνέχεια το επίπεδο εξόδου αποτελείται από τέσσερις νευρώνες διότι το νευρωνικό δίκτυο ομαδοποίησης που δημιουργήθηκε για το σύνολο δεδομένων Iris διαθέτει τέσσερα χαρακτηριστικά άρα τέσσερις νευρώνες εισόδου.Όταν τα δεδομένα μετασχηματιστούν από το νευρωνικό δίκτυο Deep τα αποτελέσματα του μετασχηματισμού αποθηκεύονται σε έναν πίνακα διάστασης (150,4) και πλέον θα εφαρμοστεί σε αυτά τα δεδομένα ο απλός αλγόριθμος ομαδοποίησης K-μέσων για K=3.Έτσι προκύπτουν τρία κέντρα με τα οποία υπολογίζουμε τα βάρη και τους μεροληπτικούς όρους που θα χρησιμοποιηθούν στο δεύτερο νευρωνικό δίκτυο που αντιστοιχεί στην

αναπαράσταση της ομαδοποίησης το οποίο έχει την ίδια αρχιτεκτονική με το νευρωνικό δίκτυο που χρησιμοποιήθηκε στην πρώτη εφαρμογή αλλά με τη διαφορά ότι πλεόν κάνει την αναπαράσταση σε μετασχηματισμένα δεδομένα.Για τον λόγο αυτόν γίνεται η ένωση μεταξύ επιπέδου εξόδου του Deep και επιπέδου εισόδου του KmeansModel1.Πλέον από το βαθύ ,τελικό νευρωνικό δίκτυο διέρχονται τα αρχικά δεδομένα μη μετασχηματισμένα και διαδοχικά εκτελούνται όλες οι παραπάνω διαδικασίες ώστε να προκύψει το τελικό αποτέλεσμα της ομαδοποίησης και η απόδοση της ομαδοποίησης είναι $ari = 0.56208670095518$.



Σχήμα 11: Αναπαράσταση της αρχιτεκτονικής του τελικού βαθιού νευρωνικού δίκτυου .

Στο σχήμα 11 αναπαριστάται η αρχιτεκτονική του τελικού βαθιού νευρωνικού δίκτυου , με τέσσερις νευρώνες εισόδου , τρία κρυφά επίπεδα που αποτελούνται από δεκαέξι νευρώνες ,τέσσερις νευρώνες και τρείς αντίστοιχα και μία έξοδο.Το νευρωνικό δίκτυο είναι γραμμικό και πλήρες συνδεδεμένο.

5.2.4 Απονομή συνάφειας στα δεδομένα εισόδου ,διαδικασία LRP

Στη περίπτωση του βαθιού νευρωνικού δικτύου όπου το ένα μέρος του αποτελεί το δίκτυο της συσταδοποίησης και το άλλο μέρος ένα τυπικό νευρωνικό δίκτυο εμπρόσθιας τροφοδότισης η συνάφεια διαδίδεται μέχρι ένα σημείο με τον κανόνα LRP που ορίζεται στο άρθρο για την διαδικασία της νευρονοποίησης και από εκείνο το σημείο και μετά χρησιμοποιούνται οι κλασικοί κανόνες των νευρωνικών δικτύων για να επιστρέψουμε πίσω στα δεδομένα εισόδου.Σύμφωνα με την παραπάνω αρχιτεκτονική που περιγράφηκε ξεκινώντας από το επίπεδο εξόδου και πηγαίνοντας διαδοχικά στα προηγούμενα επίπεδα χρησιμοποιε-

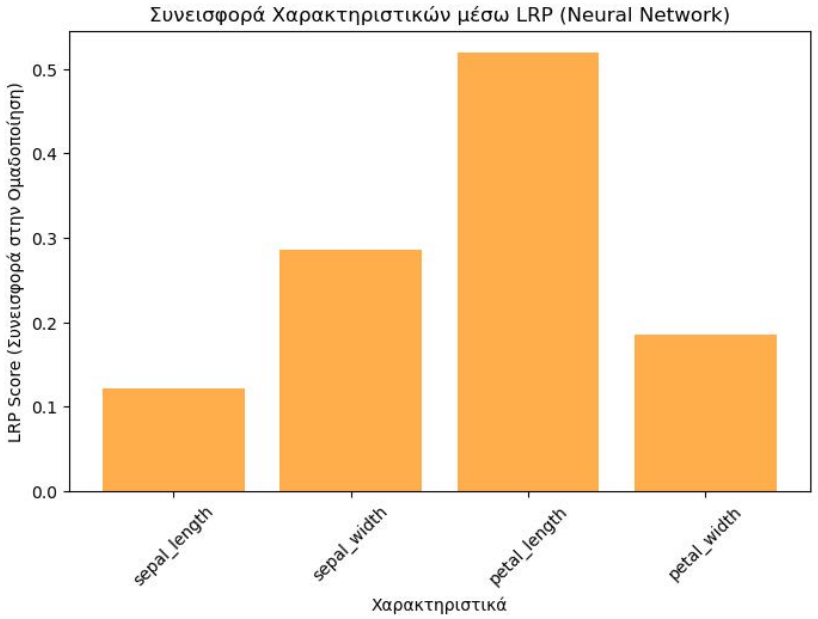
ίται η διαδικασία LRP μέχρι το τρίτο επίπεδο από το τέλος δηλαδή το επίπεδο εισόδου του δεύτερου νευρώνικού δικτύου που αποτελεί ταυτόχρονα επίπεδο εξόδου για το πρώτο νευρωνικό δίκτυο.Στο τμήμα αυτό ακολουθούνται πιστά οι μαθηματικοί τύποι (14),(15),(16) για τον υπολογισμό της συνάφειας σε καθένα από τους τέσσερις νευρώνες του τρίτου επιπέδου από το τέλος μέσω της διαδικασίας LRP.Από εκείνο το σημείο και πίσω μέχρι το επίπεδο εισόδου η συνάφεια που έχει υπολογιστεί θα περάσει από τους δεκαέξι νευρώνες και ύστερα στα δεδομένα εισόδου με βάση τους κανόνες διάδοσης για ένα τυπικό βαθύ νευρωνικό δίκτυο.Για την λειτουργία αυτή έχει υλοποιηθεί αντίστοιχη συνάρτηση όπου πάλι ξεκινώντας από το επίπεδο εξόδου (έχει γίνει υπολογισμός LRP έως εκεί) γίνεται η εξής λειτουργία:

1. Στη περίπτωση που το επίπεδο που συναντάμε είναι γραμμικό ,πλήρες συνδεδεμένο γίνεται πολλαπλασιασμός της συνάφειας με τα βάρη του επιπέδου λαμβάνοντας υπόψιν τη συνάφεια του προηγούμενου επιπέδου και τα βάρη του τρέχοντος.Κατά τον υπολογισμό αυτό γίνεται και μία κανονικοποίηση ώστε να αποφευχθούν οι μηδενικές τιμές προσθέτοντας μια πολύ μικρή τιμή ε για σταθερότητα.

2. Για τα επίπεδα ReLu η συνάφεια παραμένει μόνο για θετικές ενεργοποιήσεις της συνάρτησης ενώ οι αρνητικές μηδενίζοντας διατηρώντας έτσι τη λογική ότι μόνο οι θετικές τιμές στην συνάρτηση ενεργοποίησης συμβάλουν στη λειτουργία του δικτύου.

3. Για οποιοδήποτε άλλο επίπεδο όπως το batchnorm η τιμή της συνάφειας που έχει υπολογιστεί μέχρι εκείνο το σημείο δεν αλλάζει παραμένει αμετάβλητη.

Με την ακριβώς ίδια λογική που περιγράφηκε και στη πρώτη εφαρμογή η μεγαλύτερη τιμή αντιστοιχεί στο χαρακτηριστικό με τη μεγαλύτερη συνάφεια στην ομαδοποίηση έτσι λαμβάνοντας την απόλυτη τιμή του πίνακα συνάφειας και ύστερα τη μέση τιμή για κάθε χαρακτηριστικό έχουμε:



Σχήμα 12: Αναπαράσταση της συνάφειας των χαρακτηριστικών εισόδου .

Στο διάγραμμα του σχήματος 12 απεικονίζεται η συνάφεια των χαρακτηριστικών εισόδου . Στο συγκεκριμένο παράδειγμα όπου τα χαρακτηριστικά έχουν περάσει από ένα βαθύ νευρωνικό δίκτυο τριών κρυφών επιπέδων και έχουν μετασχηματιστεί σε έναν χώρο υψηλότερης διάστασης η διαδικασία υπολογισμού απόδοσης συνάφειας είναι τροποποιημένη διότι χρησιμοποιούνται δύο νευρωνικά δίκτυα ένα που αναπαριστά την ομαδοποίηση K-μέσων και ένα το οποίο χρησιμοποιείται ως κωδικοποιητής για τον μετασχηματισμό των δεδομένων.Έτσι παρατηρούμε ότι το χαρακτηριστικό που είχε καθοριστικό ρόλο στην συσταδοποίηση και στην απάντηση του γιατί ένα σημείο ανήκει σε μία συγκεκριμένη ομάδα είναι το petal length, ύψος πετάλου.Το αμέσως επόμενο χαρακτηριστικό με μεγαλύτερη συμβολή είναι sepal width πλατος σεπάλου.

6 Ερμηνεία συσταδοποίησης κειμένων και εξαγωγή θεμάτων.

6.1 Χρήση του απλού νευρωνικού δικτύου K-μέσων στο σύνολο δεδομένων BBC News dataset για ανάλυση κειμένου.

Στην εφαρμογή που θα ακολουθήσει θα χρησιμοποιηθεί το σύνολο δεδομένων BBC News Dataset με σκοπό τη χρήση της παραπάνω τεχνικής στα κείμενα του συνόλου δεδομένων προκειμένου να διαπιστωθεί ποιές λέξεις σε κάθε κείμενο κάθε ομάδας συνέβαλαν στη διάκριση των κείμενων και τη τοποθέτησή τους σε διαφορετικές ομάδες . Το BBC News Dataset περιλαμβάνει ειδησεογραφικά άρθρα απο τον ιστότοπο του BBc ,δημιουργήθηκε για να χρησιμοποιηθεί κυρίως σε εργασίες επεξεργασίας φυσικής γλώσσας NLP όπως ταξινόμηση κειμένων ,ανάλυση συναισθήματος ή εξαγωγή θεμάτων. Το σύνολο δεδομένων περιλαμβάνει 2.225 άρθρα από το BBc και περιλαμβάνει πέντε διαφορετικές κατηγορίες ειδήσεων :

1. Business
2. Entertainment
3. Politics
4. Tech
5. Sport

Κάθε άρθρο αποτελείται από τις εξής τέσσερις στήλες :

1. Το όνομα του αρχείου (filename).
2. Τον τίτλο (title).
3. Το περιεχόμενο του άρθρου (content).
4. Την κατηγορία (category).

Ακολουθεί απόσπασμα από τη μορφή του συνόλου δεδομένων για τις πέντε πρώτες εγγραφές:

Α/Α	**Κατ.**	**Όνομα Αρ.**	**Τίτλος**	**Περιεχόμενο**
1	business	001.txt	Ad sales boost Time Warner profit	Quarterly profits at US media giant TimeWarner...
2	business	002.txt	Dollar gains on Greenspan speech	The dollar has hit its highest level against ...
3	business	003.txt	Yukos unit buyer faces loan claim	The owners of embattled Russian oil giant Yuk...
4	business	004.txt	High fuel prices hit BA's profits	British Airways has blamed high fuel prices f...
5	business	005.txt	Pernod takeover talk lifts Domecq	Shares in UK drinks and food firm Allied Dome...

Πίνακας 2: Απόσπασμα από το BBC News Dataset (πρώτες 5 εγγραφές)

Στη συγκεκριμένη εφαρμογή εφόσον γνωρίζουμε ήδη τις πέντε κατηγορίες θα εφαρμόσουμε τον αλγόριθμο K-μέσων για K=5 ώστε τα κείμενα να τοποθετηθούν σε πέντε διαφορετικές ομάδες σύμφωνα με τις κατηγορίες που αναφέρθηκαν.Στη συνέχεια αφού προκύψουν τα κέντρα των πέντε διαφορετικών ομάδων θα χρησιμοποιήσουμε την ίδια μεθοδολογία με τα προηγούμενα παραδείγματα για τη νευρονοποίηση του αλγορίθμου καθώς και τον υπολογισμό των μεροληπτικών όρων και των βαρών.Αφού αναπαραστήσουμε την ομαδοποίηση με τη χρήση του αντίστοιχου νευρωνικού δικτύου θα χρησιμοποιήσουμε τις ίδιες διαδικασίες απόδοσης συνάφειας LRP ώστε να αποδόσουμε τιμή συμβολής σε κάθε λέξη κάθε κειμένου κάθε ομάδας ώστε να εξηγηθεί ποιές λέξεις ήταν σημαντικές ώστε το συγκεκριμένο κείμενο να τοποθετηθεί στη συγκεκριμένη ομάδα.Η διαφορά σε σχέση με τις προηγούμενες εφαρμογές είναι ότι θα πρέπει επειδή έχουμε κείμενα πλέον να προηγηθεί μία ειδική προεπεξεργασία ώστε να φέρουμε τα κείμενα σε μορφή διανυσμάτων ώστε να μπορούν να χρησιμοποιηθούν στο νευρωνικό δίκτυο.Για τις συγκεκριμένες διαδικασίες έχουν υλοποιηθεί συναρτήσεις οι οποίες εκτελούν τις διαδικασίες που απαιτούνται για την εκτέλεση της εφαρμογής .Τέλος θα γίνει σύγκριση των αποτελεσμάτων της διαδικασίας LRP με τα αποτελέσματα απλής καταμέτρησης των πιο συχνών λέξεων που εμφανίζονται σε κάθε ομάδα καθώς και με τα αποτελέσματα των λέξεων που προκύπτουν απο τη μέθοδο LDA σε κάθε ομάδα.

Προεπεξεργασία δεδομένων:

Για να μπορέσει να εφαρμοστεί ο αλγόριθμος που έχει υλοποιηθεί θα πρέπει τα δεδομένα να μην έχουν τη μορφή κειμένου όπως είναι η αρχική τους μορφή ,έτσι θα πρέπει να αντιστοιχηθεί κάθε κείμενο σε ένα διάνυσμα , για τον λόγο αυτόν θα χρειαστούμε μόνο τη στήλη content που περιέχει τα περιέχομενα των άρθρων αλλά και τη στήλη category ως true labels για να αξιλογήσουμε τη ποιότητα της ομαδοποίησης.Αρχικά αφού έχουμε κρατήσει μόνο τη στήλη content από το σύνολο δεδομένων μετατρέποντας το σε λίστα και αποθηκεύοντας την στη μεταβλητή (text) θα πρέπει να δημιουργηθούν tokens.Δηλαδή

κάθε άρθρο που είναι ένα σύνολο λέξεων και φράσεων θα διαχωριστεί σε λέξεις (tokens) έτσι θα αποτελεί κάθε άρθρο μία λίστα λέξεων.Πριν προχωρήσουμε σε αυτόν τον διαχωρισμό θα φορτώσουμε το αρχείο stopwords της βιβλιοθήκης nltk ώστε να αφαιρεθούν από τα κείμενα γενικές λέξεις . Για τον σχηματισμό των tokens έχει υλοποιηθεί συνάρτηση η οποία παίρνει διαδοχικά κάθε εγγραφή-άρθρο αρχικά μετατρέπει όλα τα κεφαλαία γράμματα σε μικρά , στη συνέχεια αφαιρεί χαρακτήρες όπως εισαγωγικά και σύμβολα και οποιονδήποτε χαρακτήρα δεν ανήκει στο αγγλικό αλφάβητο και στους αριθμούς 0-9 , έπειτα ΄σπάει ΄ τις φράσεις σε επιμέρους λέξεις διαχωρίζοντας τις λέξεις με κόμμα και πρίν επιστρέψει τα tokens κάνει έλεγχο αν η λέξη βρίσκεται στο αρχείο των stopwords , αν δεν βρίσκεται στο αρχείο προστίθεται στη λίστα tokens.Στο τέλος προκύπτει μια λίστα από λίστες όπου κάθε εσωτερική λίστα αντιστοιχεί σε κάθε άρθρο και στο περιεχόμενο της βρίσκονται οι λέξεις (tokens) διαχωρισμένες με κόμμα. Για την αντιστοίχιση κάθε άρθρου-εγγραφής σε διάνυσμα θα χρησιμοποιηθεί το προεκπαιδευμένο μοντέλο Word2Vec της βιβλιοθήκης gensim.

6.1.1 Μοντέλο Word2Vec

Το Word2Vec είναι μια αποτελεσματική τεχνική ενσωμάτωσης λέξεων.Πιο συγκεκριμένα το μοντέλο αντιστοιχεί κάθε λέξη που υπάρχει στο λεξιλόγιο σε ένα διάνυσμα , οι λέξεις με παρόμοια σημασία ή όμοιες λέξεις θα έχουν ΄κοντινές΄ διανυσματικές αναπαραστάσεις δηλαδή όμοια διανύσματα.Το μοντέλο έχει ως σκοπό να εξάγει σημασιολογικές σχέσεις μεταξύ των λέξεων.Το Word2Vec είναι ένα νευρωνικό δίκτυο δυο επιπέδων εκπαιδευμένο, όπου στην είσοδο του νευρωνικού δικτύου εισέρχονται τα έγγραφα/κείμενα , στη συγκεκριμένη περίπτωση εισέρχονται τα all tokenized δηλαδή η λίστα από λίστες που προέκυψε από τα tokens κάθε κειμένου .Με την είσοδο στο δίκτυο γίνεται κωδικοποίηση 1-hot για κάθε λέξη που υπάρχει στο κείμενο. Στο κρυφό επίπεδο υπάρχουν τόσοι νευρώνες όσο το μήκος του διανύσματος που έχουμε επιλέξει για την αναπαράσταση κάθε λέξης. Το επίπεδο εξόδου αποτελείται απο τόσους νευρώνες όσο και το επίπεδο εισόδου και ουσιαστικά περιέχει πιθανότητες για κάθε λέξη εισόδου να είναι η λέξη που αναμένουμε για συγκεκριμένη είσοδο.Με το τέλος της εκπαίδευσης έχουμε τη λέξη ενσωμάτωσης η οποία είναι τα βάρη του κρυφού επιπέδου .Στη συγκεκριμένη εργασία θα γίνει χρήση προεκπαιδευμένου μοντέλου Google News που περιλαμβάνει ενσωματώσεις λέξεων για καλύτερη ποιότητα embedding και σταθερότητα.Καταλήγουμε σε διανυσματικές αναπαραστάσεις μεγέθους 300 διαστάσεων για κάθε λέξη κάθε κειμένου.Επόμενο βήμα είναι να καταλήξουμε σε ένα μόνο διάνυσμα 300 διαστάσεων που θα αντιστοιχεί σε κάθε κείμενο , αυτό θα το πετύχουμε υπολογίζοντας τη μέση τιμή των διανυσμάτων των λέξεων για κάθε κείμενο και έτσι προκύπτει η τελική αναπαράσταση όπου κάθε εγγραφή/κείμενο αντιστοιχίζεται σε ένα διάνυσμα διάστασης 300.Άρα πλέον έχουμε έναν πίνακα με τα embeddings των λέξεων μεγέθους 2225 γραμμών και 300 στηλών.

Ακολουθεί στιγμιότυπο από τον πίνακα με τα embeddings των κειμένων:

Γραμμή	**0**	**1**	**2**	$\cdots$	**299**	**300**
0	-0.341540	-0.133578	0.356442	$\cdots$	0.121867	-0.081342
1	-0.275671	0.071958	0.193290	$\cdots$	0.072121	0.073692
2	-0.435356	0.017695	0.109920	$\cdots$	0.096529	0.290121
3	-0.193838	-0.062683	0.334585	$\cdots$	0.031269	-0.053603
4	-0.262244	0.003551	0.218501	$\cdots$	0.093813	0.088719
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
2225	-0.137269	0.028891	0.008384	$\cdots$	0.124355	0.090482

Πίνακας 3: Πίνακας με τις διανυσματικές αναπαραστάσεις των κειμένων.

Εφαρμογή αλγορίθμου K-μέσων στις διανυσματικές αναπαραστάσεις των κειμένων:

Εφόσον οι κατηγορίες που διαθέτουμε στο σύνολο δεδομένων είναι πέντε όπως αναφέραμε και προηγουμένως, θα εφαρμόσουμε τον αλγόριθμο ομαδοποίησης K-μέσων για K=5 με βάση τον τύπο (1) όπου K ο αριθμός των ομάδων.Με την ολοκλήρωση της διαδικασίας αυτής θα έχουμε τα πέντε κέντρα για κάθε μια από τις πέντε ομάδες , ο δείκτης ARI μεταξύ των ομάδων που προκύπτουν από την ομαδοποίηση K-μέσων και των πραγματικών ομάδων (true labels) ισούται με $ari = 0.7438666231645302$. Τα κέντρα που προκύπτουν από την ομαδοποίηση K-μέσων είναι πέντε όσες και οι ομάδες και είναι διανύσματα διάστασης 300 όσο είναι και η διάσταση της εισόδου.

6.1.2 Νευρονοποίηση της ομαδοποίησης K-μέσων για το σύνολο δεδομένων BBC News.

Στο σημείο αυτό αφού έχουν υπολογιστεί τα κέντρα των πέντε ομάδων μπορούμε να προχωρήσουμε στον υπολογισμό των βαρών και των μεροληπτικών όρων σύμφωνα με τους τύπους (12) και (13) .Τα βάρη είναι διάστασης (5,300) γιατί όπως αναφέραμε προηγουμένως 5 είναι οι ομάδες άρα και οι νευρώνες του κρυφού επιπέδου και 300 είναι η διάσταση των δεδομένων εισόδου. Επίσης οι μεροληπτικοί όροι είναι διάστασης (5,1) όσοι δηλαδή και οι νευρώνες.Το νευρωνικό δίκτυο που θα χρησιμοποιηθεί για να αναπαραστήσει τη συγκεκριμένη ομαδοποίηση θα έχει ένα κρυφό επίπεδο με αριθμό νευρώνων όσες και οι ομάδες δηλαδή 5 νευρώνες , το επίπεδο εισόδου θα αποτελείται από 300 νευρώνες όσο είναι δηλαδή η διάσταση των δεδομένων εισόδου και το επίπεδο εξόδου όπου όπως έχουμε αναφέρει και στις προηγούμενες εφαρμογές δεν είναι τυπικός νευρώνας αλλά κάνει απλά την επιλογή της μικρότερης απόστασης προκειμένου να τοποθετηθεί κάθε embedding σε μια από τις πέντε ομάδες , επομένως θα γίνει χρήση των μαθηματικών τύπων (10) και (11).

Ακολουθεί η αρχιτεκτονική του νευρωνικού δικτύου ως επιβεβαίωση της παραπάνω διαδικασίας που θα υλοποιήσει τη συγκεκριμένη ομαδοποίηση :

$$KmeansModel(\ (linear) : Linear(in features = 300, out features = 5, bias = True)$$

Κάθε embedding το οποίο είναι η διανυσματική αναπαράσταση κάθε κειμένου με διάσταση 300 θα εισέρχεται στο παραπάνω νευρωνικό δίκτυο και με βάση το μαθηματικό

τύπο (10) προκύπτει μια τιμή/απόσταση σε καθένα από τους πέντε νευρώνες στη συνέχεια γίνεται χρήση του τύπου (11) για να επιλεχθεί η μικρότερη από αυτές και να τοποθετηθεί το embedding στη συγκεκριμένη ομάδα.Έτσι ο πίνακας με τις αποστάσεις που προκύπτει είναι διάστασης (2225,5).Ακολουθεί στιγμιότυπο με τις αποστάσεις που προκύπτουν για τις πέντε πρώτες εγγραφές:

$$\begin{bmatrix} -0.3164 & -0.8021 & 0.1883 & 0.1112 & -0.4857 \\ 0.2652 & -0.3755 & 0.4802 & 0.6182 & -0.6407 \\ 0.4654 & -0.4126 & 0.5002 & 0.7022 & -0.8781 \\ 0.0055 & -0.8072 & 0.1536 & 0.2334 & -0.8126 \\ -0.0153 & -0.8899 & 0.1347 & 0.0486 & -0.8747 \end{bmatrix}$$

Όπως προκύπτει από τον πίνακα για το embedding του πρώτου κειμένου η μικρότερη τιμή βρίσκεται στον δεύτερο νευρώνα ,αντίστοιχα για το δεύτερο ,τρίτο και τέταρτο κείμενο η μικρότερη τιμή βρίσκεται στον πέμπτο νευρώνα ενώ για το πέμπτο κείμενο η μικρότερη τιμή βρίσκεται στον δεύτερο νευρώνα .Επομένως το συμπέρασμα που προκύπτει είναι ότι το πρώτο και πέμπτο κείμενο ανήκουν στην ίδια ομάδα καθώς και το δεύτερο , τρίτο ,τέτρατο κείμενο ανήκουν στην ίδια ομάδα αλλά διαφορετικές μεταξύ τους.Η ομαδοποίηση μέσω του νευρωνικού δικτύου έχει δείκτη $ari = 0.6482842720008718$,η απόδοση είναι χαμηλότερη σε σχέση με την απόδοση έπειτα από την εφαρμογή του απλού K-μέσων αλλά το γεγονός αυτό είναι κάτι το οποίο περιμέναμε σύμφωνα με το άρθρο αφού όπως έχουμε αναφέρει και στις προηγούμενες εφαρμογές η τεχνική αυτή δε χρησιμοποιείται για να βελτιστοποιήσει την ομάδοποίηση αλλά για να δώσει εξήγηση με βάση τα χαρακτηριστικά εισόδου στο πώς λειτούργησε ο αλγόριθμος και στο ποιό χαρακτηριστικό συνέβαλε στην τοποθέτηση μιας εγγραφής σε μια ομάδα.Για να προχωρήσουμε με τον υπολογισμό της συνάφειας πρέπει να υπολογίσουμε τις τιμές f_c που όπως προκύπτει από τον μαθηματικό τύπο (11) είναι η δεύτερη μικρότερη τιμή .Έτσι για τα πέντε πρώτα κείμενα/εγγραφές οι τιμές f_c είναι :

$$\begin{bmatrix} -0.48567687 \\ -0.37549898 \\ -0.4126304 \\ -0.80719407 \\ -0.87465542 \end{bmatrix}$$

6.1.3 Απονομή συνάφειας στις λέξεις που συνέβαλαν στην ομαδοποίηση των κειμένων.

Στο σημείο αυτό γίνεται υπολογισμός της συμβολής κάθε χαρακτηριστικού από κάθε εγγραφή στην ομαδοποίηση , για την διαδικασία αυτή θα χρησιμοποιηθούν οι μαθηματικοί τύποι (14),(15),(16) και οι αντίστοιχες συναρτήσεις που υλοποιήθηκαν μέσω της γλώσσας Python . Η διαδικασία είναι ακριβώς η ίδια με αυτή που ακολουθήσαμε στις δύο προηγούμενες εφαρμογές με τη διαφορά ότι πλέον ώς είσοδο στο νευρωνικό δίκτυο έχουμε το διάνυσμα που αντιστοιχεί σε κάθε κείμενο .Τα διανύσματα ενσωμάτωσης δημιουργήθηκαν με τη τεχνική Word2Vec που περιγράψαμε σε προηγούμενη ενότητα.Τα διανύσματα είναι διάστασης 300 οπότε κατά τη διαδικασία LRP που πραγματοποιήσαμε αποδίδεται συνάφεια/τιμή σε κάθε μια από τις τριακόσιες διαστάσεις , επομένως με την ακριβώς αντίστροφη διαδικασία από αυτή που χρησιμοποιήσαμε για να υπολογίσουμε το διάνυσμα μέσης τιμής των διανυσμάτων κάθε λέξης του κειμένου , θα αποδώσουμε την συνάφεια

σε κάθε διάνυσμα λέξεων και κατέπέκταση στις αρχικές λέξεις. Για να επιτύχουμε αυτή τη διαδικασία έχει υλοποιηθεί αντίστοιχη συνάρτηση η οποία έχει ως στόχο να μεταφέρει τη συνάφεια (16) του κάθε embedding στα αντίστοιχα tokens απο τα οποία απαρτίζεται.Η συνάρτηση κάνει την εξής λειτουργία:

1. Λαμβάνει τα embeddings των tokens του κάθε εγγράφου.
2. Λαμβάνει τον πίνακα με τα διανύσματα LRP κάθε εγγράφου/embedding.
3. Υπολογίζει το άθροισμα των embeddings ανά διάσταση.Για κάθε διάσταση κανονικοποιεί κάθε embedding ως ποσοστό του συνολικού embedding της διάστασης και πολλαπλασιάζει το ποσοστό αυτό με τη συνάφεια της αντίστοιχης διάστασης.
4. Η διαδικασία αυτή επαναλαμβάνεται για όλες τις διαστάσεις και στο τέλος προκύπτει πίνακας με τις συνάφειες των tokens.

Με την ολοκλήρωση της παραπάνω διαδικασίας θα έχουμε για κάθε κείμενο τις λέξεις και τις αντίστοιχες συνάφειες τους. Ως παράδειγμα παρατίθεται οι λέξεις και δίπλα η συνάφεια που απονέμεται σε αυτές μέσω της διαδικασίας LRP για το κείμενο 1:

Κείμενο 1

Λέξη: quarterly, **Συνάφεια:** -0.0087
Λέξη: profits, **Συνάφεια:** -0.0811
Λέξη: us, **Συνάφεια:** -0.0199
Λέξη: media, **Συνάφεια:** -0.0578
Λέξη: giant, **Συνάφεια:** -0.0092
Λέξη: timewarner, **Συνάφεια:** 0.0074
Λέξη: jumped, **Συνάφεια:** -0.0685
Λέξη: 76, **Συνάφεια:** -0.0314
Λέξη: 1, **Συνάφεια:** 0.0133
Λέξη: 13bn, **Συνάφεια:** -0.0457
Λέξη: 600μ, **Συνάφεια:** 0.0114
Λέξη: three, **Συνάφεια:** 0.0104
Λέξη: months, **Συνάφεια:** 0.0588
Λέξη: december, **Συνάφεια:** -0.0552
Λέξη: 639μ, **Συνάφεια:** -0.0104
Λέξη: year, **Συνάφεια:** 0.0020
Λέξη: earlier, **Συνάφεια:** -0.0485
Λέξη: firm, **Συνάφεια:** -0.109
Λέξη: one, **Συνάφεια:** -0.0911
Λέξη: biggest, **Συνάφεια:** 0.0086
Λέξη: investors, **Συνάφεια:** 0.0276
Λέξη: google, **Συνάφεια:** 0.0361
Λέξη: benefited, **Συνάφεια:** -0.0516
Λέξη: sales, **Συνάφεια:** 0.0797
...
Λέξη: us, **Συνάφεια:** 0.0000
Λέξη: version, **Συνάφεια:** 0.0003
Λέξη: world, **Συνάφεια:** 0.0000

Λέξη: warcraft, **Συνάφεια:** -0.0001

Για να μπορέσουμε να εξάγουμε σημαντικά συμπεράσματα μέσα από αυτή τη διαδικασία απαιτείται μια επιπλέον ενέργεια ,πιο συγκεκριμένα σε κάθε κείμενο κάθε ομάδας ξεχωριστά θα απομωνωθούν οι λέξεις με τις συνάφειες τους και στη συνέχεια έχοντας αυτές τις λέξεις θα γίνει καταμέτρηση της συχνότητας εμφάνισης τους σε κάθε κείμενο κάθε ομάδας.Έπειτα θα γίνει υπολογιμσός του μέσου όρου των συναφειών κάθε λέξης ανάλογα με τη συχνότητα εμφάνισης τους στα κείμενα.Επομένως λέξεις με μεγαλύτερη συχνότητα εμφάνισης θα έχουν μια καλύτερη τιμή μέσου όρου.Τέλος επιλέγουμε τις λέξεις με τον μεγλύτερο μέσο όρο συναφειών και έτσι θα εξάγουμε συμπέρασμα για το ποιές λέξεις ήταν καθοριστικής σημασίας για τη τοποθέτηση ενός κειμένου σε μία συγκεκριμένη ομάδα.Για την επίτευξη αυτής της διαδικασίας έχει υλοποιηθεί κατάλληλη συνάρτηση προκειμένου να επιστραφούν οι 10 λέξεις με μεγαλύτερο μέσο όρο συνάφειας που συναντάμε σε κάθε ομάδα.Για να κατανοήσουμε πόσο καλά έχει λειτουργήσει ο αλγόριθμος και να αξιολογήσουμε τα αποτελέσματα που μας επιστρέφει θα συγκρίνουμε τις λέξεις που προέκυψαν απο τη διαδικάσια LRP με χρήση νευρωνικού δικτύου , με τη διαδικασία LDA για κάθε ομάδα και με την απλή καταμέτρηση συχνότερων λέξεων σε κάθε ομάδα. Στη περίπτωση της απλής καταμέτρησης της συχνότητας έχει οριστεί συνάρτηση με την οποία για κάθε ομάδα σε κάθε κείμενο που περιλαμβάνεται σε αυτή θα καταμετρούνται οι λέξεις , από αυτές θα κρατήσουμε τις δέκα λέξεις με τη μεγαλύτερη συχνότητα.Μέσω αυτής της διαδικάσιας θα μπορέσουμε να συγκρίνουμε τις δύο μεθόδους και να αξιολογήσουμε αν συσχετίζονται σε κάθε ομάδα οι λέξεις που εμφανίζονται περισσότερες φορές με τις λέξεις που παρουσιάζουν μεγαλύτερη συνάφεια που απονομήθηκε στις λέξεις μέσω της διαδικασίας LRP.Παράλληλα για επιπλέον συγκρίσεις θα εφαρμόσουμε ανεξάρτητα από τις παραπάνω διαδικασίες , μια διαδικασία topic modeling με χρήση του αλγορίθμου Latent Dirichlet allocation (LDA).

6.1.4 Μοντελοποίηση θέματος με χρήση του αλγορίθμου LDA.

Η μοντελοποίηση θέματος είναι μια διαδικασία επεξεργασίας φυσικής γλώσσας (NLP) χωρίς επίβλεψη . Αυτή η τεχνική αποτελεί μία μέθοδο εξόρυξης κειμένου σε μεγάλα σύνολα κειμένων για να εντοπίσει ένα σύνολο όρων που προέρχονται από τα έγγραφα αυτά και αντιπροσωπεύουν ένα σύνολο θεμάτων που εντοπίζονται στα κείμενα. Τα μοντέλα θεμάτων εντοπίζουν συγκεκριμένες ,κοινές λέξεις κλειδιά σε ένα σύνολο κειμένων και ομαδοποιούν τις λέξεις σε ένα σύνολο θεμάτων. Τα μοντέλα θεμάτων είναι μια μορφή ανάλυσης κειμένου που βασίζεται σε μοντέλα μηχανικής μάθησης και χρησιμοποιείται για να σχολιάσει θεματικά μεγάλα σώματα κειμένου.Ουσιαστικά τα μοντέλα θεμάτων δεν χρησιμοποιούν ετικέτες γιαυτό και αποτελεί μέθοδο μη εποπτευόμενης μάθησης .Πιο συγκεκριμένα αντιμετωπίζει κάθε έγγραφο σε μια συλλογή κειμένων ως μια ΄τσάντα λέξεων΄ αγνοώντας τη σειρά μέσα στο κείμενο αλλά εστιάζοντας στο πόσο συχνά εντοπίζονται μέσα σε κάθε κείμενο και πόσο συχνά συνυπάρχουν με άλλες λέξεις. Η πλειοψηφία των προσεγγίσεων της μοντελοποίησης θεμάτων ξεκινούν με τη δημιουργία μήτρας εγγράφου-όρου , η μήτρα αυτή περιλαμβάνει τα έγγραφα σε κάθε σειρά και τις μεμονωμένες λέξεις ως στήλες ή αντίστροφα.Κάθε τιμή στον πίνακα δηλώνει τη συχνότητα με την οποία κάθε μεμονωμένη λέξη εμφανίζεται σε κάθε έγγραφο.Στη συνέχεια ο πίνακας αυτός χρησιμοποιείται για τη δημιουργία ενός διανυσματικού χώρου όπου n λέξεις ισούται με n διαστάσεις. Η τιμή μίας δεδομένης σειράς δηλώνει τη θέση αυτού του εγγράφου στο διανυσματικό χώρο , έτσι το μοντέλο αντιμετωπίζει την εγγύτητα στον διανυσματικό χώρο ως έγγραφα που μοιράζονται παρόμοιο περιεχόμενο ή θέματα.Στη συνέχεια το μοντέλο ομαδοποιεί τις λέξεις

που συνήθως συνυπάρχουν σε σύνολα θεμάτων .Κάθε θέμα μοντελοποιείται ως κατανομή πιθανοτήτων σε ένα λεξιλόγιο λέξεων , έπειτα κάθε έγγραφο αναπαρίσταται ως προς αυτά τα θέματα.

6.1.5 Ο αλγόριθμος LDA.

Η λανθάνουσα κατανομή Dirichlet είναι ένα πιθανοτικός αλγόριθμος μοντελοποίησης θεμάτων.Αυτό σημαίνει ότι δημιουργεί θέματα ταξινομώντας λέξεις και έγγραφα μεταξύ θεμάτων σύμφωνα με κατανομές πιθανοτήτων. Ο αλγόριθμος χρησιμοποιεί τη μήτρα εγγράφου-όρου και δημιουργεί κατανομές θεμάτων (λίστες λέξεων με αντίστοιχες πιθανότητες) σύμφωνα με τη συχνότητα εμφανίσεων τους και τις συνεμφανίσεις . Με το τρόπο αυτό οδηγούμαστε στο συμπέρασμα ότι λέξεις που εμφανίζονται μαζί είναι πιθανότατα μέρος παρόμοιων θεμάτων. Εκχωρεί τα έγγραφα σε θέματα με βάση τις ομάδες λέξεων που εμφανίζονται σε ένα συγκεκριμένο έγγραφο. Κατά την ανάθεση θεμάτων ο αλγόριθμος χρησιμοποιεί τη μέθοδο δειγματοληψίας Gibbs. Πιο συγκεκριμένα ο τύπος είναι ο εξής:

Ο τύπος για την πιθανότητα ανάθεσης του θέματος z_i στη λέξη w_i είναι:

$$P(z_i = k \mid z_{-i}, w) \propto \frac{n_{d,k}^{-i} + \alpha}{\sum_{k'} n_{d,k'}^{-i} + K\alpha} \cdot \frac{n_{k,w}^{-i} + \beta}{\sum_{w'} n_{k,w'}^{-i} + V\beta} \tag{27}$$

όπου:

- z_i: το θέμα για τη λέξη i
- z_{-i}: οι αναθέσεις θεμάτων για όλες τις άλλες λέξεις εκτός της i
- w: η λέξη που εξετάζεται
- $n_{d,k}^{-i}$: πλήθος φορών που το θέμα k έχει ανατεθεί στο έγγραφο d, χωρίς τη λέξη i
- $n_{k,w}^{-i}$: πλήθος φορών που η λέξη w έχει ανατεθεί στο θέμα k, χωρίς τη λέξη i
- α, β: υπερπαράμετροι του Dirichlet
- K: συνολικός αριθμός θεμάτων
- V: μέγεθος λεξιλογίου

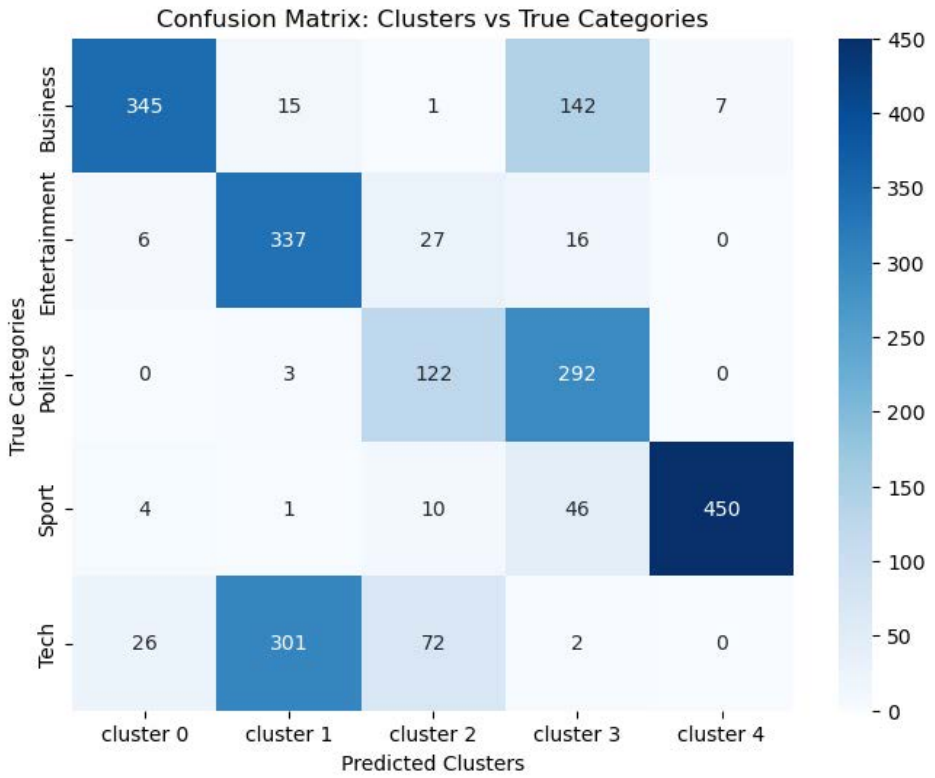
Η δειγματοληψία Gibbs είναι μια μέθοδος για την εκτίμηση των λανθανουσών μεταβλητών δηλαδή την ανάθεσης θεμάτων σε λέξεις στα έγγραφα. Είναι μια τεχνική Markov Chain Monte Carlo και χρησιμοποιείται επειδή η ακριβής εκτίμηση είναι υπολογιστικά ανεφάρμοστη. Στον LDA αλγόριθμο ουσιαστικά υπολογίζουμε τη πιθανότητα μια λέξη σε ένα έγγραφο να ανήκει σε ένα συγκεκριμένο θέμα.

6.1.6 Αξιολόγηση της ομαδοποίησης με χρήση πίνακα σύγχυσης (confusion matrix).

Ο πίνακας σύγχυσης είναι εργαλείο που χρησιμοποιείται για την αξιολόγηση της απόδοσης ενός αλγορίθμου ταξινόμησης. Ο πίνακας συγκρίνει τις πραγματικές κλάσεις δηλαδή τι είναι πραγματικά τα δεδομένα με τις προβλεπόμενες κλάσεις δηλαδή τι προβλέπει ο αλγόριθμος. Η αξιολόγηση του πίνακα σε περίπτωση δυαδικής ταξινόμησης γίνεται ως εξής:

- True Positive (TP): Το μοντέλο προέβλεψε θετικό, και ήταν όντως θετικό.
- False Positive (FP): Το μοντέλο προέβλεψε θετικό, αλλά ήταν αρνητικό (λέγεται και ¨τύπου I σφάλμα¨).
- False Negative (FN): Το μοντέλο προέβλεψε αρνητικό, αλλά ήταν θετικό (λέγεται και ¨τύπου II σφάλμα¨).
- True Negative (TN): Το μοντέλο προέβλεψε αρνητικό, και ήταν όντως αρνητικό.

Στη συγκεκριμένη εφαρμογή οι κατηγορίες/ομάδες είναι πέντε για τον λόγο αυτό ο πίνακας θα είναι ένας τετραγωνικός πίνακας 5x5 , άρα έχουμε :

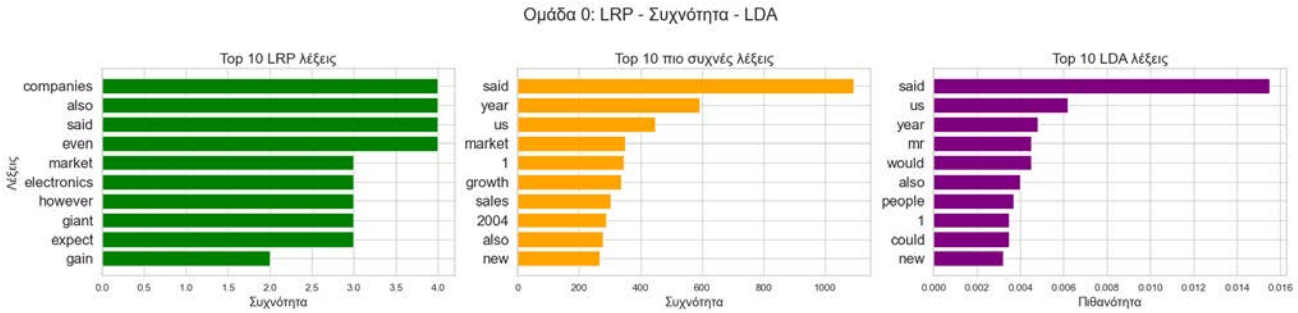


Σχήμα 13: Πίνακας σύγχυσης (confusion matrix) ο οποίος αναπαριστά στις γραμμές τις πραγματικές κλάσεις και στις στήλες τις προβλεπόμενες κλάσεις/ομάδες που πρέβλεψε ο αλγόριθμος K-μέσων για K=5 δηλαδή πέντε ομάδες.

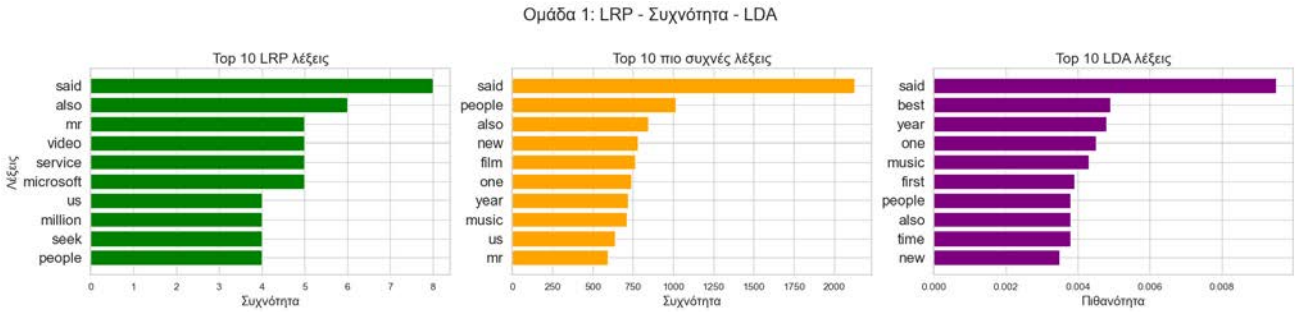
Από τον πίνακα του σχήματος 13 προκύπτει ότι στην πρώτη ομάδα (cluster 0 η πλειοψηφία των κειμένων είναι της κατηγορίας business.Η δεύτερη ομάδα (cluster 1) αποτελείται κυρίως από κείμενα της κατηγορίας entertainment και tech.Αντίστοιχα η τρίτη ομάδα (cluster 2) αποτελείται απο κείμενα κυρίως της κατηγορίας politics και λιγότερα tech .Η τέταρτη ομάδα (cluster 3) περιλαμβάνει κυρίως κείμενα της κατηγορίας politics και λιγότερα business ενώ η πέμπτη ομάδα (cluster 4) αποτελείται από κατηγορίες κειμένων sport . Τα αποτελέσματα της ομαδοποίησης προβάλλονται στο διάγραμμα του σχήματος 19.

6.1.7 Αποτελέσματα.

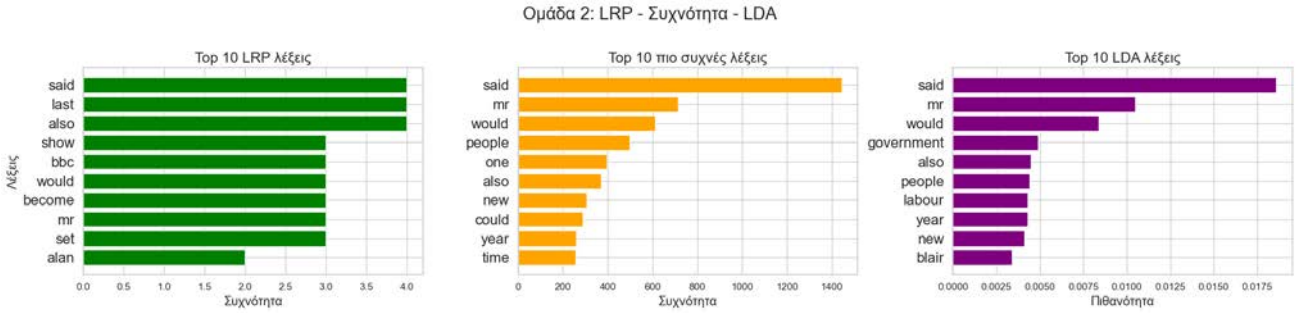
Για την οπτικοποίηση των παραπάνω τεχνικών θα γίνει χρήση ραβδογράμματος.Για κάθε ομάδα θα εμφανίζονται τρία διαγράμματα , το πρώτο θα αφορά τη διαδικασία LRP όπου έχουν επιλεχθεί οι δέκα λέξεις με τη μεγαλύτερη συνάφεια που έχουν μεγαλύτερο μέσο όρο , το δεύτερο διάγραμμα αφορά λέξεις με μεγαλύτερη συχνότητα εμφάνισης και το τρίτο διάγραμμα αφορά λέξεις που προκύπτουν απο την εφαρμογή του αλγορίθμου LDA.



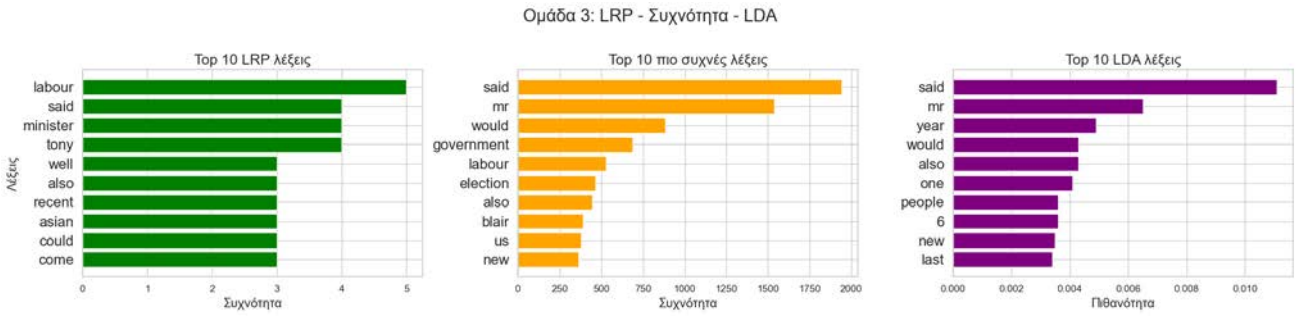
Σχήμα 14: Ραβδογράμματα ομάδας 0 που αφορουν κείμενα business.



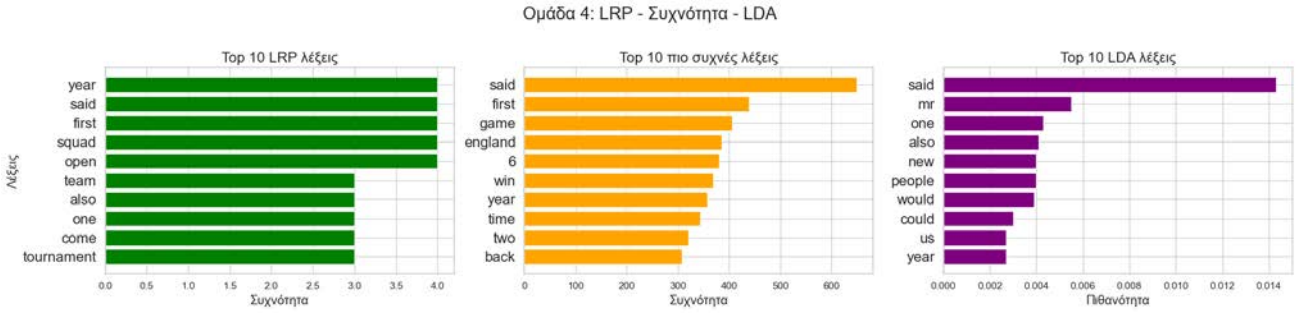
Σχήμα 15: Ραβδογράμματα ομάδας 1 που αφορούν κείμενα τόσο entertainment όσο και tech.



Σχήμα 16: Ραβδογράμματα ομάδας 2 που αφορούν κείμενα politics και tech.

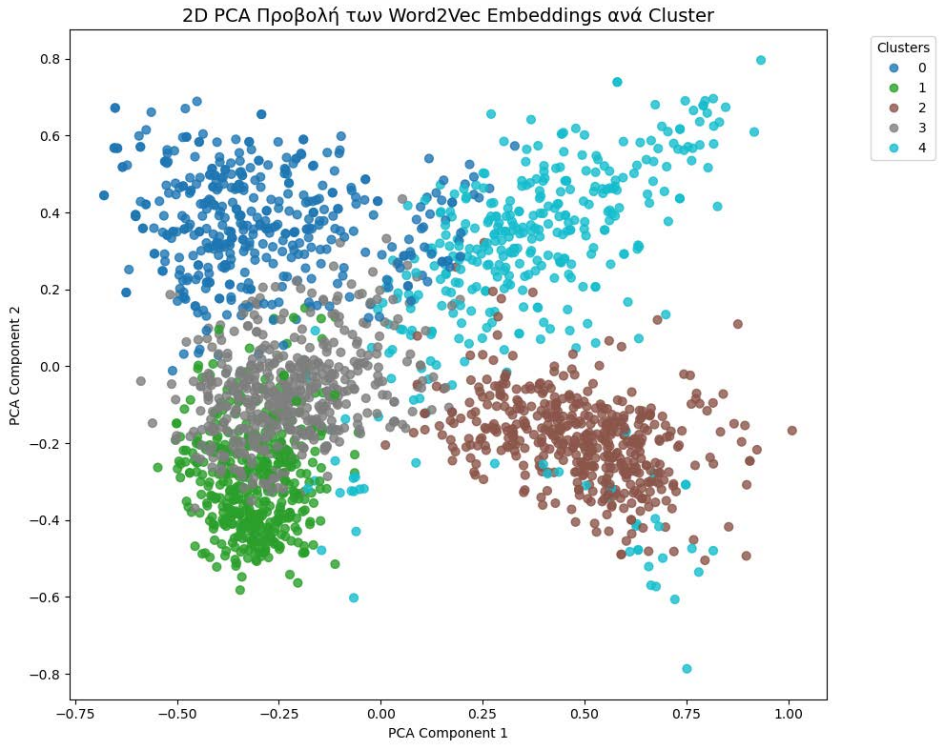


Σχήμα 17: Ραβδογράμματα ομάδας 3 που αφορούν κείμενα politics και business.



Σχήμα 18: Ραβδογράμματα ομάδας 4 που αφορούν κείμενα sport.

Αυτό το οποίο παρατηρείται στα σχετικά ραβδογράμματα 14 , 15 , 16 , 17 , 18 είναι ότι σε κάθε μια από τις πέντε ομάδες οι λέξεις που προκύπτουν από τον αλγόριθμο LRP , από την απλή καταμέτρηση συχνότητας λέξεων και από τον αλγόριθμο LDA είναι στη πλειοψηφία διαφορετικές σε κάθε περίπτωση. Στη πρώτη ομάδα η οποία αφορά κυρίως κείμενα της κατηγορίας business παρατηρούμε ότι με τη μέθοδο LRP προκύπτουν λέξεις οι οποίες σχετίζονται άμεσα με τη κατηγορία και δεν συναντάμε τις αντίστοιχες στη συχνότητα και στο LDAόπως companies,market,electronics,gain.Στην ομάδα 1 έχουμε κείμενα τόσο κατηγορίας entertainment όσο και tech όπου παρατηρούμε πάλι λέξεις στο LRP που περιγράφουν το θέμα και δεν συναντάμε αντίστοιχες στις άλλες δύο τεχνικές όπως video,service,microsoft.Στις ομάδες 2 και 3 που το μεγαλύτερος μέρος των ομάδων περιλαμβάνει κείμενα politics και λιγότερα tech για την 2 και business για την 3 αντίστοιχα παρατηρούμε πάλι λέξεις σχετικές με το θέμα στο LRP όπως bbc,labour,minister που δεν συναντάμε στις άλλες δύο μεθόδους.Τέλος στην ομάδα 4 έχουμε κείμενα κυρίως sport και προκύπτει ότι η μέθοδος LRP βρήκε λέξεις σχετικές με το θέμα άμεσα που επίσης δεν συναντάμε στις άλλες δύο μεθόδους όπως squad,team,tournament.



Σχήμα 19: Απεικόνιση των ομάδων που προέκυψαν από την συσταδοποίηση K-μέσων για K=5 με χρήση νευρωνικού δικτύου στο σύνολο δεδομένων BBC κειμένων.

Στο διάγραμμα του σχήματος 19 έχει εφαρμοστεί στα embeddings των κειμένων ο αλγόριθμος μείωσης διάστασης PCA ώστε να γίνει μείωση των αρχικών 300 διαστάσεων σε 2 ώστε να γίνει εφικτή η αναπαράσταση τους με χρήση διαγράμματος και στη συνέχεια ο χρωματισμός τους με βάση τις διαφορετικές ομάδες.Με μπλέ χρώμα αναφερόμαστε στα κείμενα της ομάδας tech , με πράσινο χρώμα σε κείμενα της ομάδας business, με καφέ χρώμα στην ομάδα sport , με γκρί χρώμα σε κείμενα της ομάδας entertainment και με γαλάζιο χρώμα σε κείμενα τόσο της ομάδας business όσο και της ομάδας politics.

Ανάλυση Θεματικών Περιοχών με LLM

Παρακάτω παρουσιάζεται η απόδοση γενικών θεματικών περιοχών για κάθε συστάδα, βασισμένη σε δύο σύνολα λέξεων: τις λέξεις LRP (με υψηλή συνάφεια μέσω explainable clustering) και τις συχνότερες λέξεις (FREQ). Για την εκτίμηση αξιοποιήθηκε ένα μεγάλο γλωσσικό μοντέλο (LLM).Με τη συγκεκριμένη διαδικασία ερμηνεύουμε τα αποτελέσματα των δυο μεθόδων με σκοπό να βρούμε μια γενική κατηγορία για κάθε ομάδα.Όπως προκύπτει απο το γλωσσικό μοντέλο οι λέξεις της μεθόδου LRP είναι πιο εννοιολογικά ξεκάθαρες και ειδικές σε σχέση με τις πιο συχνές λέξεις.Τέλος η διαδικασία αυτή , δηλαδή η νευρονοποίηση , η απόδοση συνάφειας στις αρχικές λέξεις και στη συνέχεια η ερμηνεία αυτών μέσω ενός γλωσσικού θα μπορούσε να θεωρηθεί μια ολοκληρωμένη μεθοδολογία εναλλακτικού τρόπου εξαγωγής θέματος για τα κείμενα.Το γλωσσικό μοντέλο έλαβε ως είσοδο το εξής ερώτημα:

Ερώτημα προς το LLM:

- Analyze the provided keywords.

- Generate a **generic topic label** that captures the broader, more general thematic area suggested by these words.
- After providing the generic topic labels for both LRP and FREQ sets for a cluster, write a brief "Observation" comparing the two generic topics.

Cluster 0
LRP Keywords: *companies, also, said, even, market, electronics, however, giant, expect, gain*
Generic Topic (LRP): Business & Industry News
FREQ Keywords: *said, year, us, market, 1, growth, sales, 2004, also, new*
Generic Topic (FREQ): Economic Reporting & Market Analysis
Observation: Both point to economic or business discussions. LRP hints more at corporate discourse within a sector, while FREQ is broader, covering general market reporting. .

Cluster 1
LRP Keywords: *said, also, mr, video, service, microsoft, us, million, seek, people*
Generic Topic (LRP): Technology & Corporate Announcements
FREQ Keywords: *said, people, also, new, film, one, year, music, us, mr*
Generic Topic (FREQ): Media & Entertainment Updates
Observation: LRP points more clearly towards technology and corporate strategy in digital media. FREQ is broader, covering general entertainment news (film, music)..

Cluster 2
LRP Keywords: *said, last, also, show, bbc, would, become, mr, set, alan*
Generic Topic (LRP): Media & Broadcasting News
FREQ Keywords: *said, mr, would, people, one, also, new, could, year, time*
Generic Topic (FREQ): Public Statements & Future Outlooks
Observation: LRP clearly orients towards media and broadcasting. FREQ is highly generic, indicating general announcements or discussions about future possibilities by an individual. LRP provides a more defined generic field.

Cluster 3
LRP Keywords: *labour, said, minister, tony, well, also, recent, asian, could, come*
Generic Topic (LRP): Political Discourse & Governance
FREQ Keywords: *said, mr, would, government, labour, election, also, blair, us, new*
Generic Topic (FREQ): Politics & Election News
Observation: Both are clearly political. LRP hints at specific policy discussions or international/community relations within politics, while FREQ is more centered on general government announcements and electoral politics.

Cluster 4
LRP Keywords: *year, said, first, squad, open, team, also, one, come, tournament*
Generic Topic (LRP): Sports & Competitions
FREQ Keywords: *said, first, game, england, 6, win, year, time, two, back*
Generic Topic (FREQ): Sports Results & Events
Observation: Both strongly indicate sports. LRP focuses on the organizational aspect

(teams, tournaments), while FREQ points more directly to game outcomes and specific events. The generic topic is very similar.

Βιβλιογραφία

References

[1] David Baehrens, Timon Schroeter, Stefan Harmeling, Motoaki Kawanabe, Katja Hansen, and Klaus-Robert Müller. How to explain individual classification decisions. *J. Mach. Learn. Res.*, 11:1803–1831, August 2010.

[2] Pei-Yin Chen and Jih-Jeng Huang. A hybrid autoencoder network for unsupervised image clustering. *Algorithms*, 12:122, 06 2019.

[3] Eduardo Laber, Lucas Murtinho, and Felipe Oliveira. Shallow decision trees for explainable k-means clustering, 2022.

[4] Scott Lundberg and Su-In Lee. A unified approach to interpreting model predictions, 2017.

[5] Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

[6] Martin Magdin, Juraj Benc, Štefan Koprda, Zoltán Balogh, and Daniel Tuček. Comparison of multilayer neural network models in terms of success of classifications based on emgucv, ml.net and tensorflow.net. *Applied Sciences*, 12(8), 2022.

[7] Tri Nguyen, Shahana Ibrahim, and Xiao Fu. Deep clustering with incomplete noisy pairwise annotations: a geometric regularization approach. In *Proceedings of the 40th International Conference on Machine Learning*, ICML'23. JMLR.org, 2023.

[8] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. "why should i trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 1135–1144, New York, NY, USA, 2016. Association for Computing Machinery.

[9] J.M. Zurada, A. Malinowski, and I. Cloete. Sensitivity analysis for minimization of input data dimension for feedforward neural network. In *Proceedings of IEEE International Symposium on Circuits and Systems - ISCAS '94*, volume 6, pages 447–450 vol.6, 1994.