ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

Συστήματα Σύστασης Μουσικής: Μια Συγκριτική Μελέτη των Κλασικών, Βαθιάς Μάθησης και Βασισμένων σε LLM Αλγορίθμων

Διπλωματική Εργασία

Απόστολος Χριστόφορος Ρουσσάς

Επιβλέπουσα: Ελένη Τουσίδου

Σεπτέμβριος 2025

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

Συστήματα Σύστασης Μουσικής: Μια Συγκριτική Μελέτη των Κλασικών, Βαθιάς Μάθησης και Βασισμένων σε LLM Αλγορίθμων

Διπλωματική Εργασία

Απόστολος Χριστόφορος Ρουσσάς

Επιβλέπουσα: Ελένη Τουσίδου

Σεπτέμβριος 2025

UNIVERSITY OF THESSALY

SCHOOL OF ENGINEERING

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Music Recommendation Systems: A comparative study of Classical, Deep Learning, and LLM-Based Algorithms

Diploma Thesis

Apostolos Christoforos Roussas

Supervisor: Eleni Tousidou

September 2025

Εγκρίνεται από την Επιτροπή Εξέτασης:

Επιβλέπουσα **Ελένη Τουσίδου**
Εργαστηριακό Διδακτικό Προσωπικό, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Πανεπιστήμιο Θεσσαλίας

Μέλος **Μιχαήλ Βασιλακόπουλος**
Καθηγητής, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Πανεπιστήμιο Θεσσαλίας

Μέλος **Γεώργιος Θάνος**
Εργαστηριακό Διδακτικό Προσωπικό, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Πανεπιστήμιο Θεσσαλίας

Ευχαριστίες

Θα ήθελα να ευχαριστήσω θερμά την Επιβλέπουσα Καθηγήτριά μου, κα. Ελένη Τουσίδου, για την αδιάκοπη καθοδήγηση, την εμπιστοσύνη και τον χρόνο που αφιέρωσε καθ' όλη τη διάρκεια της Διπλωματικής μου Εργασίας.

Ευχαριστώ επίσης τα μέλη της Τριμελούς Εξεταστικής Επιτροπής, κ. Μιχαήλ Βασιλακόπουλο, και κ. Γεώργιο Θάνο, για τον χρόνο που διέθεσαν να μελετήσουν την εργασία μου, καθώς και την υποστήριξη τους σε όλη την διάρκεια των σπουδών μου.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου, τους φίλους και συμφοιτητές μου για την πολύτιμη και καθοριστική υποστηριξή τους.

ΥΠΕΥΘΥΝΗ ΔΗΛΩΣΗ ΠΕΡΙ ΑΚΑΔΗΜΑΪΚΗΣ ΔΕΟΝΤΟΛΟΓΙΑΣ ΚΑΙ ΠΝΕΥΜΑΤΙΚΩΝ ΔΙΚΑΙΩΜΑΤΩΝ

«Με πλήρη επίγνωση των συνεπειών του νόμου περί πνευματικών δικαιωμάτων, δηλώνω ρητά ότι η παρούσα διπλωματική εργασία, καθώς και τα ηλεκτρονικά αρχεία και πηγαίοι κώδικες που αναπτύχθηκαν ή τροποποιήθηκαν στα πλαίσια αυτής της εργασίας, αποτελεί αποκλειστικά προϊόν προσωπικής μου εργασίας, δεν προσβάλλει κάθε μορφής δικαιώματα διανοητικής ιδιοκτησίας, προσωπικότητας και προσωπικών δεδομένων τρίτων, δεν περιέχει έργα/εισφορές τρίτων για τα οποία απαιτείται άδεια των δημιουργών/δικαιούχων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον και πληρούν τους κανόνες της επιστημονικής παράθεσης. Τα σημεία όπου έχω χρησιμοποιήσει ιδέες, κείμενο, αρχεία ή/και πηγές άλλων συγγραφέων, αναφέρονται ευδιάκριτα στο κείμενο με την κατάλληλη παραπομπή και η σχετική αναφορά περιλαμβάνεται στο τμήμα των βιβλιογραφικών αναφορών με πλήρη περιγραφή. Δηλώνω επίσης ότι τα αποτελέσματα της εργασίας δεν έχουν χρησιμοποιηθεί για την απόκτηση άλλου πτυχίου. Αναλαμβάνω πλήρως, ατομικά και προσωπικά, όλες τις νομικές και διοικητικές συνέπειες που δύναται να προκύψουν στην περίπτωση κατά την οποία αποδειχθεί, διαχρονικά, ότι η εργασία αυτή ή τμήμα της δεν μου ανήκει διότι είναι προϊόν λογοκλοπής».

Ο Δηλών

Απόστολος Χριστόφορος Ρουσσάς

Διπλωματική Εργασία

Συστήματα Σύστασης Μουσικής: Μια Συγκριτική Μελέτη των Κλασικών, Βαθιάς Μάθησης και Βασισμένων σε LLM Αλγορίθμων

Απόστολος Χριστόφορος Ρουσσάς

Περίληψη

Η παρούσα εργασία εξετάζει τη θεωρητική θεμελίωση και πρακτική υλοποίηση σύγχρονων Συστημάτων Σύστασης Μουσικής, εστιάζοντας στη συγκριτική μελέτη αλγοριθμικών προσεγγίσεων που αξιοποιούνται για την προσωποποιημένη πρόβλεψη μουσικών προτιμήσεων.

Σε θεωρητικό επίπεδο, πραγματοποιείται συστηματική ανασκόπηση των βασικών τεχνικών που έχουν διαμορφώσει το πεδίο: συνεργατικό φιλτράρισμα (collaborative filtering), φιλτράρισμα με βάση το περιεχόμενο (content-based filtering), τεχνικές παραγοντοποίησης πινάκων (matrix factorization), καθώς και σύγχρονες προσεγγίσεις βασισμένες σε βαθιά νευρωνικά δίκτυα, ενσωματώσεις (embeddings), πολυτροπικά μοντέλα και μεγάλα γλωσσικά μοντέλα (LLMs). Ιδιαίτερη έμφαση αποδίδεται στην αντιμετώπιση του προβλήματος ψυχρής εκκίνησης (cold-start), στη μοντελοποίηση της ακολουθιακής συμπεριφοράς των χρηστών, καθώς και στην ερμηνεία των αποτελεσμάτων.

Το πρακτικό σκέλος της εργασίας περιλαμβάνει την υλοποίηση εφαρμογής σύστασης μουσικού περιεχομένου, η οποία ενσωματώνει πολλαπλές αλγοριθμικές προσεγγίσεις και επιτρέπει τον πειραματισμό με βάση την απόδοση του κάθε αλγορίθμου. Η εφαρμογή αξιοποιεί δεδομένα χρηστών, πληροφορίες ακρόασης, ετικέτες και μεταδεδομένα τραγουδιών, συνδυάζοντας τόσο κλασικά μοντέλα όσο και σύγχρονες τεχνικές με ενσωματώσεις γλωσσικών χαρακτηριστικών. Η αποτίμηση των συστημάτων πραγματοποιείται με χρήση καθιερωμένων μετρικών, όπως Precision@k, Recall@k, NDCG, MRR, Hit Rate και Coverage, με στόχο την τεκμηριωμένη σύγκριση ως προς την ακρίβεια και την εξατομίκευση.

Λέξεις-κλειδιά:

Συστήματα Σύστασης Μουσικής, Συνεργατικό Φιλτράρισμα, Φιλτράρισμα με βάση το περιεχόμενο, Παραγοντοποίηση πινάκων, Διανύσματα, Αξιολόγηση Συστάσεων

Diploma Thesis

Music Recommendation Systems: A comparative study of Classical, Deep Learning, and LLM-Based Algorithms

Apostolos Christoforos Roussas

Abstract

This dissertation examines both the theoretical aspects and real-world applications of modern Music Recommendation Systems, with a particular focus on a comparative analysis of algorithmic methods for predicting individual music preferences.

On the theoretical side, systematic reviews of several core methods take place, like collaborative filtering, content-based filtering, and matrix factorization models, along with more recent developments around deep neural networks, embeddings, multimodal frameworks, and large language models (LLMs). Specific emphasis is given to addressing cold-start issues, modeling of user listening behavior as sequential, and providing interpretation for the results.

On the practical side, this thesis develops a music recommendation application that incorporates multiple algorithms and can be used to experiment according to their performance against each other. The application utilizes data from user interactions with music data, metadata around tracks, tags, and text-based embeddings, combining both classical models and modern techniques that incorporate language features. This study also uses typical evaluation metrics like Precision@k, Recall@k, NDCG, MRR, Hit Rate, and Coverage to evaluate each method in terms of preciseness and personalization.

Keywords:

Music Recommender Systems, Collaborative Filtering, Content-Based Filtering, Matrix Factorization, Embeddings, Evaluation Metrics

Πίνακας περιεχομένων

Ευχαριστίες **ix**

Περίληψη **xii**

Abstract **xiii**

Πίνακας περιεχομένων **xv**

Κατάλογος σχημάτων **xvii**

Κατάλογος πινάκων **xix**

Συντομογραφίες **xxi**

1 Εισαγωγή **1**

1.1 Αντικείμενο της διπλωματικής 1

1.1.1 Συνεισφορά 2

1.2 Οργάνωση του τόμου 3

2 Σχετικές Εργασίες **5**

3 Θεωρητικό Υπόβαθρο **9**

3.1 Συνεργατικό Φιλτράρισμα (Collaboritive Filtering) 9

3.1.1 Αλγόριθμος KNN με Βάση τον Χρήστη (UserKNN) 10

3.1.2 Αλγόριθμος KNN με Βάση το Αντικείμενο (ItemKNN) 11

3.1.3 Παραγοντοποίηση Πινάκων με SVD (Singular Value Decomposition) 12

3.2 Νευρωνικό Συνεργατικό Φιλτράρισμα (Neural Collaborative Filtering) . . . 13

3.3 BERT και Μεγάλα Γλωσσικά Μοντέλα στην Επεξεργασία Περιεχομένου Συστάσεων . . . 14
3.4 Μετρικές Αξιολόγησης . . . 15

4 Χαρακτηριστικά, Προεπεξεργασία και Διαχωρισμός Συνόλου Δεδομένων 19
4.1 Σύνολο Δεδομένων . . . 19
4.2 Προεπεξεργασία Δεδομένων . . . 20
4.2.1 Αρχείο artists.dat . . . 21
4.2.2 Αρχείο user_artists.dat . . . 22
4.2.3 Αρχείο tags.dat . . . 22
4.2.4 Αρχείο user_taggedartists.dat . . . 22
4.3 Διαχωρισμός Δεδομένων σε Σύνολα Εκπαίδευσης και Αξιολόγησης . . . 23

5 Μεθολογία και Αξιολόγηση Αλγορίθμων 25
5.1 Αρχιτεκτονική και Βασική Διεπαφή . . . 25
5.2 Ανάλυση Τεχνικών Βελτιστοποίησης Συστάσεων . . . 26
5.3 Αλγόριθμοι συνεργατικού φιλτραρίσματος . . . 28
5.3.1 Εύρεση βέλτιστου αριθμού γειτόνων (k) . . . 28
5.3.2 Αλγόριθμος KNN με βάση τον χρήστη (UserKNN) . . . 28
5.3.3 Αλγόριθμος KNN με βάση το Αντικείμενο (ItemKNN) . . . 31
5.3.4 Αλγόριθμος Παραγοντοποίησης Πινάκων (SVD) . . . 33
5.4 Αλγόριθμοι Βαθιάς Μάθησης: Neural Matrix Factorization (NeuMF) . . . 36
5.5 Αλγόριθμος με βάση το περιεχόμενο: Sentence-BERT με Last.fm Tags . . . 39
5.6 Συγκριτική επίδοση των αλγορίθμων . . . 41

6 Συμπεράσματα και Μελλοντικές Επεκτάσεις 47
6.1 Συμπεράσματα . . . 47
6.2 Μελλοντικές Επεκτάσεις . . . 48

Βιβλιογραφία 51

Κατάλογος σχημάτων

3.1 Διάγραμμα ροής NeuMF 14

4.1 Στατιστικά Δεδομένων προ επεξεραγασίας 20
4.2 Πυκνότητα Δεδομένων προ επεξεργασίας 20
4.3 Στατιστικά Δεδομένων μετά επεξεραγασίας 21
4.4 Πυκνότητα Δεδομένων μετά επεξεργασίας 21

5.1 Ροή Βασικής διεπαφής 26
5.2 Γραφική αναπαράσταση μετρικών UserKNN 30
5.3 Γραφική αναπαράσταση μετρικών ItemKNN 32
5.4 Γραφική αναπαράσταση μετρικών SVD 35
5.5 Γραφική αναπαράσταση μετρικών NeuMF 38
5.6 Γραφική αναπαράσταση μετρικών BERT 41
5.7 Heatmap μετρικών UserKNN, ItemKNN, SVD, NeuMF 44
5.8 Heatmap μετρικών BERT 45
5.9 Bar Chart μετρικών UserKNN, ItemKNN, SVD, NeuMF, BERT 45

Κατάλογος πινάκων

4.1 Συγκριτικά αποτελέσματα πριν και μετά το φιλτράρισμα 23

5.1 Αποτελέσματα αξιολόγησης για $k = 5$ 28
5.2 Αποτελέσματα αξιολόγησης για $k = 10$ 28
5.3 Αποτελέσματα αξιολόγησης για $k = 20$ 28
5.4 Αποτελέσματα αξιολόγησης UserKNN 30
5.5 Αποτελέσματα αξιολόγησης ItemKNN 32
5.6 Αποτελέσματα αξιολόγησης SVD 34
5.7 Αποτελέσματα αξιολόγησης NeuMF 37
5.8 Αποτελέσματα αξιολόγησης BERT 40
5.9 Συγκριτικά αποτελέσματα στα @5 42
5.10 Συγκριτικά αποτελέσματα στα @10 42
5.11 Αποτελέσματα αξιολόγησης στα @20 43
5.12 Αποτελέσματα αξιολόγησης στα @50 43
5.13 Αποτελέσματα αξιολόγησης με MRR και Coverage 44

Συντομογραφίες

CF	Collaborative Filtering (Συνεργατικό Φιλτράρισμα)
CB	Content-Based (Βάσει Περιεχομένου)
NCF	Neural Collaborative Filtering (Νευρωνικό Συνεργατικό Φιλτράρισμα)
KNN	k-Nearest Neighbors (k Πλησιέστεροι Γείτονες)
SVD	(Truncated) Singular Value Decomposition (Αποσύνθεση Μοναδιαίων Τιμών)
MF	Matrix Factorization (Παραγοντοποίηση Πίνακα)
MLP	Multi-Layer Perceptron (Πολυστρωματικό Perceptron)
NeuMF	Neural Matrix Factorization (Νευρωνική Παραγοντοποίηση Πίνακα)
BERT	Bidirectional Encoder Representations from Transformers
Sentence-BERT	Μοντέλο BERT για ενσωματώσεις προτάσεων
TF–IDF	Term Frequency–Inverse Document Frequency
MFCC	Mel-Frequency Cepstral Coefficients
STFT	Short-Time Fourier Transform
CQT	Constant-Q Transform
PCA	Principal Component Analysis
NDCG	Normalized Discounted Cumulative Gain
DCG	Discounted Cumulative Gain
IDCG	Ideal DCG
MRR	Mean Reciprocal Rank
RMSE	Root Mean Squared Error
Hit Rate@K	Ποσοστό Επιτυχιών στα Top–K
Coverage	Κάλυψη (μοναδικά αντικείμενα που προτάθηκαν)
Precision@K	Ακρίβεια στις Top–K προτάσεις
Recall@K	Ανάκληση στις Top–K προτάσεις
Cosine similarity	Συνημιτονική ομοιότητα

Pearson correlation	Συσχέτιση Pearson
Euclidean distance	Ευκλείδεια απόσταση
Cold-start	Πρόβλημα ψυχρού ξεκινήματος
Long-tail	«Μακρά ουρά» / ποινή δημοτικότητας
Implicit feedback	Σιωπηρή/έμμεση ανάδραση
Top–N	Κορυφαίες N συστάσεις
ReLU	Rectified Linear Unit
Dropout	Κανονικοποίηση με τυχαία απενεργοποίηση μονάδων
Adam	Adaptive Moment Estimation (βελτιστοποιητής)
Epoch	Επανάληψη εκπαίδευσης πάνω σε όλο το σύνολο
CUDA	Compute Unified Device Architecture
PyTorch	Βιβλιοθήκη βαθιάς μάθησης
HetRec2011	Σύνολο δεδομένων Last.fm (2011)

Κεφάλαιο 1

Εισαγωγή

Η ραγδαία αύξηση της ψηφιακής μουσικής και η μαζική διάθεσή της μέσω διαδικτυακών υπηρεσιών έχουν καταστήσει τα Συστήματα Σύστασης (Recommender Systems) αναπόσπαστο μέρος της εμπειρίας του τελικού χρήστη. Ειδικότερα, τα Συστήματα Σύστασης Μουσικής (Music Recommender Systems – MRS) στοχεύουν στην προσωποποιημένη πρόβλεψη μουσικών προτιμήσεων, επιδιώκοντας να προτείνουν στον χρήστη νέο μουσικό περιεχόμενο βασισμένο στο ιστορικό ακροάσεων, στις αλληλεπιδράσεις του με τα αντικείμενα ή σε μεταδεδομένα των κομματιών. Ο τομέας αυτός συνδυάζει τεχνικές από τη μηχανική μάθηση, την εξαγωγή γνώσης από δεδομένα και την αναπαράσταση πληροφορίας, ενώ αντιμετωπίζει προκλήσεις όπως το πρόβλημα ψυχρής εκκίνησης (cold-start), την πολυπλοκότητα της μουσικής αισθητικής, την ερμηνευσιμότητα των συστάσεων και την ανάγκη για συνδυασμό ετερογενών πηγών πληροφορίας (social tags, playlists, lyrics, κ.ά.).

1.1 Αντικείμενο της διπλωματικής

Η παρούσα διπλωματική εργασία επικεντρώνεται στη μελέτη, υλοποίηση και συγκριτική αξιολόγηση αλγορίθμων σύστασης μουσικού περιεχομένου. Σε θεωρητικό επίπεδο, πραγματοποιείται βιβλιογραφική επισκόπηση βασικών τεχνικών όπως η σύσταση βάση ομοιότητας αντικειμένων (item-based collaborative filtering), η σύσταση με βάση το περιεχόμενο (content-based filtering), τα υβριδικά μοντέλα, οι τεχνικές παραγοντοποίησης πινάκων (matrix factorization), καθώς και των πρόσφατων εξελίξεων με χρήση νευρωνικών δικτύων (neural networks), Sentence-BERT embeddings και πολυτροπικών μοντέλων. Παρουσιάζονται επίσης οι κύριες μετρικές αξιολόγησης (Precision@k, Recall@k, NDCG, κ.α.), προκειμένου

να υποστηριχθεί τεκμηριωμένη σύγκριση της αποτελεσματικότητας των προσεγγίσεων.

Στο πρακτικό μέρος, αναπτύσσεται μία εφαρμογή σύστασης μουσικής, η οποία επιτρέπει την εναλλαγή μεταξύ διαφορετικών μοντέλων και τεχνικών βασισμένη σε πραγματικά δεδομένα που παρέχονται από το Last.fm [1], μια πλατφόρμα μουσικής με δεδομένα για χρήστες και τις αλληλεπιδράσεις τους με καλλιτέχνες. Το σύστημα αξιοποιεί πληροφορίες χρηστών, μεταδεδομένα τραγουδιών και ετικέτες (tags), καθώς και γλωσσικές ενσωματώσεις περιγραφών με χρήση Sentence-BERT, ενσωματώνοντας τόσο παραδοσιακές μεθόδους όσο και σύγχρονες αρχιτεκτονικές βαθιάς μάθησης και προεκπαιδευμένων μεγάλων γλωσσικών μοντέλων. Μέσω αυτής της πειραματικής διαδικασίας, συγκρίνεται η απόδοση των αλγορίθμων ως προς την ακρίβεια, την κάλυψη, την ερμηνευσιμότητα και την εξατομίκευση των συστάσεων.

1.1.1 Συνεισφορά

Η συνεισφορά της διπλωματικής συνοψίζεται ως εξής:

1. Μελετήθηκαν σε βάθος οι κυριότερες κατηγορίες συστημάτων σύστασης μουσικής, με έμφαση στο συνεργατικό φιλτράρισμα, στο φιλτράρισμα με βάση το περιεχόμενο, στα νευρωνικά δίκτυα και στις υβριδικές προσεγγίσεις, συγκεντρώνοντας έτσι πλούσιο υλικό σε μία μόνο έρευνα.

2. Υλοποιήθηκαν πέντε διαφορετικές αλγοριθμικές προσεγγίσεις: (α) αλγόριθμος συνεργατικού φιλτραρίσματος με βάση τον χρήστη (UserKNN), (β) αλγόριθμος συνεργατικού φιλτραρίσματος με βάση το αντικείμενο (ItemKNN), (γ) αλγόριθμος συνεργατικού φιλτραρίσματος μέσω παραγοντοποίησης πινάκων (SVD) , (δ) υβριδικός αλγόριθμος νευρωνικών δικτύων και παραγοντοποίησης πινάκων (NeuMF) , (ε) αλγόριθμος σύστασης με βάση το περιεχόμενο χρησιμοποιώντας διανύσματα Sentence-BERT. Με αυτόν τον τρόπο η παρούσα έρευνα καλύπτει ένα ευρύ φάσμα αλγορίθμων.

3. Πραγματοποιήθηκε πειραματική αξιολόγηση των παραπάνω μεθόδων σε πραγματικό dataset (Last.fm/HetRec2011), με χρήση πολλαπών μετρικών (Recall, NDCG κ.ά.), αναλύθηκε η συγκριτική επίδοση τους και εντοπίστηκαν οι συνθήκες στις οποίες κάθε προσέγγιση υπερέχει. Έτσι, οι αξιολογήσεις είναι πιο αξιόπιστες και άμεσα συγκρίσιμες, αφού αναφέρονται στο ίδιο dataset, γεγονός που καθιστά αυτή τη διπλωματική σημείο αναφοράς για μελλοντικές επεκτάσεις.

4. Αναπτύχθηκε λειτουργική εφαρμογή που ενσωματώνει τους παραπάνω αλγορίθμους και επιτρέπει πειραματισμό με διαφορετικά σύνολα χαρακτηριστικών και παραμέτρων.

1.2 Οργάνωση του τόμου

Η εργασία οργανώνεται ως εξής: Στο Κεφάλαιο 2 παρουσιάζεται η σχετική βιβλιογραφία και προηγούμενες εργασίες στον τομέα των συστημάτων σύστασης μουσικής. Το Κεφάλαιο 3 αναλύει το θεωρητικό υπόβαθρο, καλύπτοντας τις βασικές αρχές των μεθοδολογιών που υιοθετούνται και τις μετρικές αξιολόγησης. Στο Κεφάλαιο 4 περιγράφεται το σύνολο δεδομένων, τα χαρακτηριστικά του, η δομή του, καθώς και η διαδικασία προεπεξεργασίας και παραγωγής συνόλων εκπαίδευσης και αξιολόγησης (train/test split). Το Κεφάλαιο 5 επικεντρώνεται στην αρχιτεκτονική του συστήματος, την υλοποίηση των αλγορίθμων και την παρουσίαση-ερμηνεία των αποτελεσμάτων τους. Τέλος, στο Κεφάλαιο 6 παρατίθενται τα συμπεράσματα, συνοψίζονται τα βασικά ευρήματα της εργασίας και προτείνονται κατευθύνσεις για μελλοντική έρευνα και βελτιστοποίηση.

Κεφάλαιο 2

Σχετικές Εργασίες

Η έρευνα στο πεδίο των μουσικών συστημάτων συστάσεων έχει οδηγήσει σε ποικίλες μεθοδολογικές προσεγγίσεις, καλύπτοντας από κλασικές τεχνικές συνεργατικού φιλτραρίσματος έως σύγχρονες υλοποιήσεις βασισμένες σε βαθιά μάθηση και πολυτροπικά δεδομένα. Το παρόν κεφάλαιο παρουσιάζει και σχολιάζει ενδεικτικές σχετικές εργασίες, οι οποίες είτε αποτέλεσαν άμεση έμπνευση για τη μεθοδολογία της παρούσας εργασίας είτε μελετήθηκαν ως αντιπροσωπευτικά παραδείγματα της υφιστάμενης βιβλιογραφίας.

Μία από τις πιο επιδραστικές και ευρέως υλοποιημένες προσεγγίσεις συστάσεων είναι η σύσταση με βάση την ομοιότητα των αντικειμένων (item-to-item collaborative filtering), το οποίο χρησιμοποιήθηκε για πρώτη φορά σε μεγάλη κλίμακα στο Amazon. Όπως περιγράφουν οι G.Linden et al. [2], η βασική ιδέα είναι η εύρεση παρόμοιων αντικειμένων βάσει των αλληλεπιδράσεων χρηστών με αυτά και η παραγωγή συστάσεων με βάση την ομοιότητα μεταξύ αντικειμένων. Η τεχνική αυτή έχει το πλεονέκτημα ότι είναι υπολογιστικά αποδοτική και επιτρέπει προκατασκευή των πινάκων ομοιότητας, όμως περιορίζεται σε σενάρια όπου η πυκνότητα αλληλεπιδράσεων είναι επαρκής.

Αν και η χρήση ομοιότητας αντικειμένων υπήρξε ιδιαίτερα αποδοτική, η ανάγκη για πιο αφηρημένη και γενικεύσιμη αναπαράσταση προτιμήσεων οδήγησε στην καθιέρωση τεχνικών παραγοντικής ανάλυσης. Η παραγοντική ανάλυση και ειδικότερα η Matrix Factorization έχει κυριαρχήσει στην ακαδημαϊκή κοινότητα για προβλήματα συστάσεων, ιδίως μετά τον διαγωνισμό Netflix Prize, όπου οι κορυφαίες λύσεις βελτιστοποίησης του αλγορίθμου συστάσεων του Netflix στηρίχθηκαν σε τεχνικές παραγοντοποίησης πινάκων. Σύμφωνα με τους Y.Koren et al. [3] ,η τεχνική αυτή επιδιώκει τη χαμηλοδιάστατη αναπαράσταση χρηστών και αντικειμένων σε κοινό χώρο, στον οποίο το εσωτερικό τους γινόμενο επιχειρεί

να αποτυπώσει τη προτιμησιακή τους συμπεριφορά. Παρά τα πλεονεκτήματα σε όρους γενίκευσης, η μέθοδος δεν ενσωματώνει άμεσα περιεχομενικά ή κοινωνικά χαρακτηριστικά, κάτι που αποτελεί περιορισμό για το πεδίο της μουσικής.

Ωστόσο, σε περιπτώσεις όπου απουσιάζουν ιστορικά δεδομένα, οι παραδοσιακές τεχνικές συχνά αποτυγχάνουν, γεγονός που οδήγησε στην ανάπτυξη πολυτροπικών μοντέλων για την αντιμετώπιση του προβλήματος ψυχρής εκκίνησης (cold-start). Σε μελέτη που πραγματοποιήθηκε από τους Y.Deldjoo et al. [4], προτείνεται ένα deep multimodal σύστημα συστάσεων που στοχεύει στην αντιμετώπιση του προβλήματος ψυχρής εκκίνησης (cold-start) για μουσικά αντικείμενα. Το σύστημα αυτό συνδυάζει φωνητικά χαρακτηριστικά, μεταδεδομένα και κοινωνικές ετικέτες μέσω συγχώνευσης νευρωνικών αναπαραστάσεων. Η στρατηγική του ενσωματώνει ποικίλες πηγές πληροφορίας που εμπλουτίζουν τις συστάσεις ακόμη και για τραγούδια χωρίς ιστορικά δεδομένα αλληλεπίδρασης. Έτσι και η παρούσα εργασία αξιοποιεί ομοειδή δεδομένα, όπως tag metadata, αντλώντας έμπνευση από τη συγκεκριμένη προσέγγιση.

Εκτός από την ενσωμάτωση ακουστικών χαρακτηριστικών και μεταδεδομένων, ιδιαίτερο βάρος έχει δοθεί και στις κοινωνικές σχέσεις, καθώς επηρεάζουν σημαντικά τη μουσική κατανάλωση και την αξιολόγηση των συστάσεων. Μία ενδιαφέρουσα οπτική προσφέρει το έργο των A.Bellogin et al. [5], στο οποίο εξετάζεται η διαφοροποίηση στις συστάσεις σε ένα πραγματικό κοινωνικό μουσικό δίκτυο (Last.fm), υποδεικνύοντας ότι η ετερογένεια στις προτιμήσεις απαιτεί δυναμικά και ευπροσάρμοστα μοντέλα. Η ανάλυση αυτή αποτέλεσε κίνητρο για τη μελέτη προφίλ χρηστών στο πλαίσιο αυτής της διατριβής, λαμβάνοντας υπόψη πιθανές διαφοροποιήσεις στον ρυθμό κατανάλωσης μουσικού περιεχομένου.

Η ανάλυση κοινωνικών προτύπων, αν και ενισχύει την εξατομίκευση, φέρνει στην επιφάνεια ζητήματα ισότιμης μεταχείρισης και ενδεχόμενης μεροληψίας. Η δίκαιη αναπαράσταση χρηστών και δημιουργών στα recommender συστήματα αναδεικνύεται ως κρίσιμο ερευνητικό πρόβλημα. Οι K.Dinnissen και C.Bauer [6], προσφέρουν μία συνοπτική ανασκόπηση της έννοιας της δικαιοσύνης στα μουσικά συστήματα συστάσεων, δίνοντας έμφαση στις πολυεπίπεδες επιπτώσεις στα ενδιαφερόμενα μέλη. Αν και η παρούσα εργασία δεν εστιάζει άμεσα στη δίκαιη κατανομή, λαμβάνει υπόψη τέτοιους προβληματισμούς κατά τον σχεδιασμό της αξιολόγησης.

Παράλληλα με τις ηθικές διαστάσεις, η βιβλιογραφία εξετάζει την ενίσχυση των συστάσεων μέσω περιεχομενικών χαρακτηριστικών, με ιδιαίτερη έμφαση στην αξιοποίηση ετικε-

τών και σημασιολογικών περιγραφών. Στο άρθρο του ο M.Schedl [7], περιγράφει τα κύρια μοντέλα που βασίζονται σε χαρακτηριστικά με βάση το περιεχόμενο (content-based filtering) καθώς και την εξέλιξή τους στο πεδίο της μουσικής. Έμφαση δίνεται στην επεξεργασία ετικετών (tags) και περιγραφών, όπως και στη χρήση μεταδεδομένων. Η διπλωματική αυτή ακολουθεί παρόμοια προσέγγιση, ενσωματώνοντας embeddings βασισμένα σε music tags για την παραγωγή συστάσεων.

Τέλος οι B.Amiri et al. [8], παρέχουν μία εκτενή επισκόπηση των τάσεων και των τεχνολογικών κατευθύνσεων στα μουσικά συστήματα συστάσεων, εντοπίζοντας τα ερευνητικά κενά και τα μελλοντικά ερωτήματα που παραμένουν ανοιχτά. Στο πλαίσιο αυτής της επισκόπησης επιβεβαιώνεται η ανάγκη συνδυασμού διαφορετικών τύπων πληροφορίας (κοινωνικής, ακουστικής, σημασιολογικής).

Κεφάλαιο 3

Θεωρητικό Υπόβαθρο

3.1 Συνεργατικό Φιλτράρισμα (Collaboritive Filtering)

Μια από τις βασικότερες μεθόδους ανάπτυξης αποδοτικών συστημάτων συστάσεων μουσικής είναι αυτή του συνεργατικού φιλτραρίσματος. Το συνεργατικό φιλτράρισμα (collaborative filtering - CF) αξιοποιεί μοτίβα χρήστη-αντικειμένου για να προβλέψει προτιμήσεις άλλων χρηστών. Ειδικότερα, τα συστήματα που χρησιμοποιούν αυτού του είδους αλγορίθμους στηρίζονται στη λογική ότι χρήστες με παρόμοιο ιστορικό αλληλεπιδράσεων (ακροάσεων στη συγκεκριμένη περίπτωση) τείνουν να έχουν παρόμοιες αλληλεπιδράσεις και στο μέλλον. Για παράδειγμα, αν σε δύο χρήστες αρέσουν οι ίδιοι καλλιτέχνες και ο ένας ακούσει κάποιον καινούργιο, είναι πιθανό να αρέσει και στον δεύτερο, οπότε το σύστημα θα του τον πρότεινε. Ο υπολογισμός της ομοιότητας μεταξύ δύο χρηστών (ιδίως για αλγορίθμους γειτνίασης) προκύπτει από τη χρήση ομοιότητας συνημιτόνου (cosine similarity), η οποία μετράει τη γωνιακή ομοιότητα ανάμεσα σε δύο διανύσματα αξιολογήσεων (με τιμές οι οποίες κυμαίνονται από 0 έως 1 για πλήρη ομοιότητα), χωρίς να λαμβάνεται υπόψη το μέγεθός τους. Συγκεκριμένα, για δύο διανύσματα αξιολογήσεων $\vec{r_u}$ και $\vec{r_v}$ δύο χρηστών (ή δύο αντικειμένων), η cosine similarity ορίζεται ως:

$$\text{sim}(u, v) = \frac{\vec{r_u} \cdot \vec{r_v}}{\|\vec{r_u}\| \cdot \|\vec{r_v}\|} \tag{3.1}$$

Οι αλγόριθμοι συνεργατικού φιλτραρίσματος χωρίζονται σε 2 κύριες υποκατηγορίες: τις μεθόδους βασισμένες σε γειτνίαση και τις μεθόδους βασισμένες σε τεχνικές παραγοντοποίησης πινάκων [2, 3]. Για τις ανάγκες αυτής της διπλωματικής υλοποιούνται τρεις αλγόριθμοι συνεργατικού φιλτραρίσματος δύο από την πρώτη υποκατηγορία (ο αλγόριθμος ”κ πλησιέ-

στεροι γείτονες (KNN) με βάση τον Χρήστη" και ο αλγόριθμος "κ πλησιέστεροι γείτονες (KNN) με βάση το αντικείμενο") και ένας από την δεύτερη (ο αλγόριθμος παραγοντοποίησης πινάκων με Singular Value Decomposition (SVD)).

3.1.1 Αλγόριθμος KNN με Βάση τον Χρήστη (UserKNN)

Ο αλγόριθμος γειτνίασης με βάση τον χρήστη (UserKNN) αποτελεί μία από τις πιο διαδεδομένες και κατανοητές μεθόδους συνεργατικού φιλτραρίσματος. Βασίζεται στην παραδοχή ότι χρήστες με παρόμοιο ιστορικό προτιμήσεων είναι πιθανό να έχουν και παρόμοια συμπεριφορά στο μέλλον. Για την εκτίμηση της ομοιότητας μεταξύ χρηστών, χρησιμοποιούνται μετρικές όπως η ομοιότητα συνημιτόνου (cosine similarity), ενώ όταν εφαρμόζεται κανονικοποίηση ή κεντράρισμα των αξιολογήσεων, συχνά προτιμάται η συσχέτιση Pearson (Pearson Correlation). Η διαδικασία ξεκινά με την αναπαράσταση κάθε χρήστη ως ένα διάνυσμα που περιλαμβάνει τις αλληλεπιδράσεις του με τα αντικείμενα (π.χ. ακροάσεις κομματιών). Κατόπιν, για κάθε χρήστη-στόχο, εντοπίζονται οι k πιο όμοιοι χρήστες, και οι αξιολογήσεις αυτών χρησιμοποιούνται για την εκτίμηση των προτιμήσεών του. Το τελικό σκορ ενός αντικειμένου υπολογίζεται μέσω ενός σταθμισμένου μέσου όρου των αξιολογήσεων των γειτόνων, με βάρη τις τιμές ομοιότητας. Ο παραπάνω υπολογισμός για έναν χρήστη-στόχο u, το αντικείμενο i που εξετάζεται, την αξιολόγηση ενός γειτονικού χρήστη v για το αντικείμενο i ($r_{v,i}$) και την ομοιότητα του με αυτόν $sim(u, v)$, γίνεται με τον τύπο:

$$\hat{r}_{u,i} = \frac{\sum\limits_{v \in N_k(u)} sim(u, v) \cdot r_{v,i}}{\sum\limits_{v \in N_k(u)} |sim(u, v)|} \tag{3.2}$$

Η μέθοδος UserKNN είναι ευρέως χρησιμοποιούμενη και καθιερωμένη στον χώρο της μουσικής σύστασης. Χαρακτηριστικό παράδειγμα αποτελεί η μελέτη των Bellogín et al. [5], στην οποία χρησιμοποιούνται δεδομένα από το Last.fm για να εξεταστεί πώς επηρεάζουν οι μουσικές ετικέτες, οι κοινωνικές συνδέσεις και τα μοτίβα προτιμήσεων την ποιότητα των συστάσεων. Αντίστοιχα, ο Majumdar [9] αξιοποιεί ιστορικό ακροάσεων και κοινωνικά δίκτυα χρηστών για την ενίσχυση της ακρίβειας και της εξατομίκευσης των προτάσεων. Παρότι εύχρηστος και ερμηνεύσιμος, ο UserKNN παρουσιάζει ορισμένες προκλήσεις, όπως η ευαισθησία σε αραιά δεδομένα (sparsity) και η αυξημένη υπολογιστική απαίτηση καθώς μεγαλώνει το πλήθος των χρηστών, αφού απαιτείται δυναμικός επαναϋπολογισμός των ομοιοτήτων. Ωστόσο, λόγω της απλότητάς του και της ευκολίας ενσωμάτωσης νέων τεχνικών, παραμένει

μέθοδος αναφοράς για τη σύγκριση πιο σύνθετων αλγορίθμων.

3.1.2 Αλγόριθμος KNN με Βάση το Αντικείμενο (ItemKNN)

Ο αλγόριθμος γειτνίασης με βάση το αντικείμενο (ItemKNN) διαφοροποιείται από τον UserKNN ως προς το σημείο υπολογισμού των ομοιοτήτων. Αντί να συγκρίνονται χρήστες μεταξύ τους, συγκρίνονται αντικείμενα (π.χ. τραγούδια ή καλλιτέχνες) μεταξύ τους, βάσει των χρηστών που τα έχουν αξιολογήσει ή ακούσει. Η βασική ιδέα είναι ότι αν ένας χρήστης έχει εκφράσει θετική προτίμηση για ένα αντικείμενο, είναι πολύ πιθανό να ενδιαφέρεται και για άλλα παρόμοια. Ο αλγόριθμος ξεκινά με τον υπολογισμό της ομοιότητας μεταξύ όλων των αντικειμένων, δημιουργώντας έναν πίνακα ομοιότητας μεταξύ αντικειμένων. Στη συνέχεια, το σύστημα εντοπίζει τα k πιο παρόμοια αντικείμενα με εκείνα που έχει ήδη αλληλεπιδράσει ο χρήστης και υπολογίζει την πιθανή αξιολόγηση, με βάση τις αξιολογήσεις του και την ομοιότητα των αντικειμένων. Ο παραπάνω υπολογισμός της αξιολόγησης, για έναν χρήστη u, το αντικείμενο-στόχο i, την αξιολόγηση του χρήστη u σε ένα γειτονικό αντικείμενο j ($r_{u,j}$) και την ομοιότητα του με αυτό ($sim(i,j)$), δίνεται από τον τύπο:

$$\hat{r}_{u,i} = \frac{\sum\limits_{j \in N_k(i)} sim(i,j) \cdot r_{u,j}}{\sum\limits_{j \in N_k(i)} |sim(i,j)|} \tag{3.3}$$

Η μέθοδος παρουσιάστηκε σε μεγάλη κλίμακα από τους Linden et al. [2] στο πλαίσιο του Amazon.com, όπου η σύσταση με βάση την ομοιότητα των αντικειμένων (item-to-item collaborative filtering) επέτρεψε την offline κατασκευή του πίνακα ομοιοτήτων και τη γρήγορη παραγωγή εξατομικευμένων προτάσεων. Σε περιβάλλοντα όπου οι χρήστες αλλάζουν συχνά αλλά τα αντικείμενα παραμένουν σταθερά, όπως σε πολλές μουσικές πλατφόρμες, η προσέγγιση αυτή είναι ιδιαίτερα αποδοτική. Οι Berkovsky et al. [10] αναφέρουν ότι οι τεχνικές βασισμένες στο αντικείμενο είναι κατάλληλες για περιβάλλοντα με συνεχώς εξελισσόμενη βάση χρηστών. Τέλος, παρόλο που ο ItemKNN έχει μεγαλύτερη σταθερότητα από τον UserKNN (αφού τα χαρακτηριστικά των αντικειμένων δεν αλλάζουν συχνά), εξαρτάται σημαντικά από τον αριθμό των αξιολογήσεων κάθε αντικειμένου.

3.1.3 Παραγοντοποίηση Πινάκων με SVD (Singular Value Decomposition)

Η παραγοντοποίηση πινάκων αποτελεί μία από τις πιο διαδεδομένες και επιτυχημένες προσεγγίσεις στο πεδίο του συνεργατικού φιλτραρίσματος, ειδικά σε περιβάλλοντα με μεγάλο όγκο δεδομένων και υψηλή αραίωση. Ο αλγόριθμος Singular Value Decomposition (SVD) συγκαταλέγεται στις τεχνικές αυτής της κατηγορίας, επιδιώκοντας να εντοπίσει τις λανθάνουσες διαστάσεις που διέπουν τις αλληλεπιδράσεις μεταξύ χρηστών και αντικειμένων. Αντί να συγκρίνει χρήστες ή αντικείμενα απευθείας, το μοντέλο «μαθαίνει» αφηρημένες αναπαραστάσεις για κάθε χρήστη και κάθε αντικείμενο, οι οποίες αντανακλούν πρότυπα προτίμησης, ενδιαφέροντος ή συσχέτισης, που δεν είναι ορατά στις αρχικές αξιολογήσεις. Για να το καταφέρει αυτό, ο αλγόριθμος ξεκινά διασπώντας τον αραιό πίνακα αλληλεπιδράσεων σε δύο υποπίνακες μικρότερων διαστάσεων, έναν για τους χρήστες και έναν για τα αντικείμενα. Στη συνέχεια, για ζεύγος χρήστη-αντικειμένου υπολογίζεται η πρόβλεψη μέσω του εσωτερικού γινομένου των διανυσμάτων τους. Η πρόβλεψη της αξιολόγησης ενός χρήστη u για ένα αντικείμενο i, όπου p_u το λανθάνον διάνυσμα που αναπαριστά τον χρήστη και q_i το λανθάνον διάνυσμα που αναπαριστά το αντικείμενο, δίνεται από τον τύπο:

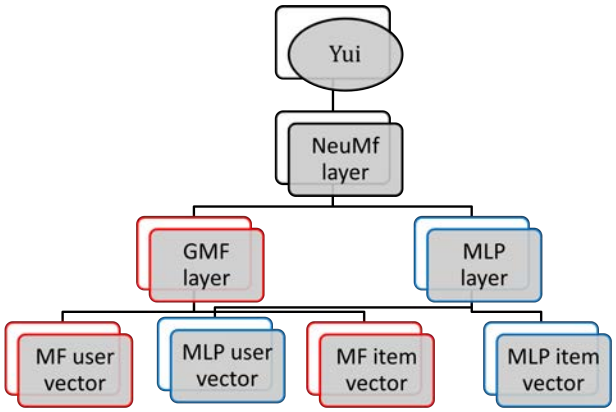
$$\hat{r}_{u,i} = p_u^\top q_i \tag{3.4}$$

Το βασικό πλεονέκτημα της SVD έγκειται στη δυνατότητά της να αποτυπώνει αυτές τις λανθάνουσες σχέσεις σε έναν συμπαγή χώρο χαρακτηριστικών, όπου οι προβλέψεις πραγματοποιούνται με βάση τη γεωμετρική συνάφεια ανάμεσα σε χρήστες και αντικείμενα. Το αποτέλεσμα είναι ένα μοντέλο που δεν εξαρτάται από άμεσες ομοιότητες, αλλά από πιο θεμελιώδεις συμπεριφορικές τάσεις, οι οποίες έχουν εξαχθεί από τα υπάρχοντα δεδομένα. Στη μελέτη των Koren et al. [3], η SVD εφαρμόζεται στο ευρύτερο πλαίσιο των latent factor μοντέλων και αποδεικνύεται ότι υπερτερεί σε ακρίβεια και γενικευσιμότητα σε σχέση με παραδοσιακές μεθόδους γειτνίασης. Οι συγγραφείς παρουσιάζουν επίσης επεκτάσεις που ενσωματώνουν πρόσθετες πληροφορίες, όπως οι μέσοι όροι χρηστών και αντικειμένων, αλλά και χρονικές δυναμικές, προκειμένου να κατανοηθεί πώς μεταβάλλονται οι προτιμήσεις με την πάροδο του χρόνου. Επιπλέον, προτείνεται η ενσωμάτωση τεχνικών regularization ώστε να αποφευχθεί η υπερπροσαρμογή, ενώ συζητούνται και διαφορετικοί τρόποι βελτιστοποίησης, όπως η χρήση στοχαστικής καθόδου βαθμίδωσης. Η μέθοδος SVD τυποποιήθηκε ευρέως

μετά την επιτυχία της στον διαγωνισμό Netflix Prize, όπου αποτέλεσε τον βασικό μηχανισμό για την παραγωγή εξαιρετικά ακριβών συστάσεων. Η χρήση του SVD επιτρέπει στο σύστημα να συλλαμβάνει πιο αφηρημένα, αλλά ουσιαστικά πρότυπα προτίμησης, οδηγώντας σε πιο προσωποποιημένες και αποτελεσματικές προτάσεις.

3.2 Νευρωνικό Συνεργατικό Φιλτράρισμα (Neural Collaborative Filtering)

Το μοντέλο Neural Collaborative Filtering, όπως παρουσιάστηκε από τους He et al. [11], αποτελεί εξέλιξη των τεχνικών παραγοντοποίησης πινάκων, συνδυάζοντας το γραμμικό εσωτερικό γινόμενο (GMF) των embedding διανυσμάτων με ένα πολυεπίπεδο νευρωνικό δίκτυο (MLP). Στην παρούσα εργασία η υλοποίηση του υβριδικού αλγορίθμου νευρωνικών δικτύων και παραγοντοποίησης πινάκων (NeuMF) βασίστηκε στη παραπάνω έρευνα. Η βασική ιδέα πίσω από το NeuMF είναι η εκμάθηση μη γραμμικών συσχετίσεων ανάμεσα στις αναπαραστάσεις χρηστών και αντικειμένων, αξιοποιώντας πλήρως συνδεδεμένα επίπεδα (fully connected layers) και μη γραμμικές συναρτήσεις ενεργοποίησης (όπως ReLU). Η σύνθεση αυτών των επιπέδων επιτρέπει την αυτόματη ανάδειξη πιο αφηρημένων και πολυδιάστατων προτύπων προτιμήσεων, τα οποία δύσκολα αποτυπώνονται με κλασικές μεθόδους. Το τελικό επίπεδο του NeuMF προβλέπει τη βαθμολογία αλληλεπίδρασης μεταξύ ενός χρήστη και ενός αντικειμένου, συνδυάζοντας τα αποτελέσματα του συνεργατικού φιλτραρίσματος μέσω παραγοντοποίησης πινάκων και αυτά του πολυεπίπεδου νευρωνικού δικτύου, ενώ η εκπαίδευση βασίζεται σε σχήματα implicit feedback, δηλαδή σχήματα στα οποία οι προτιμήσεις χρηστών υπονοούνται από συμπεριφορές (κλικ, προβολές, χρόνο παραμονής, κ.λπ.) αντί για ρητές βαθμολογίες. Το NeuMF αποτελεί χαρακτηριστικό παράδειγμα υβριδικής αρχιτεκτονικής, καθώς ενσωματώνει τόσο τη μαθηματική λογική της παραγοντοποίησης όσο και τις υπολογιστικές δυνατότητες της βαθιάς μάθησης, και έχει αποδειχθεί αποτελεσματικό σε πλήθος σεναρίων συστάσεων μουσικού περιεχομένου. Στο Σχήμα 3.1 παρουσιάζεται ένα απλοποιημένο διάγραμμα ροής σχετικά με τη διαδικασία πρόβλεψης του αλγορίθμου:



Σχήμα 3.1: Διάγραμμα ροής NeuMF

3.3 BERT και Μεγάλα Γλωσσικά Μοντέλα στην Επεξεργασία Περιεχομένου Συστάσεων

Η ραγδαία εξέλιξη των Μεγάλων Γλωσσικών Μοντέλων (Large Language Models - LLMs), όπως τα BERT, ChatGPT και LlamA, έχει διαμορφώσει ένα νέο υπόδειγμα στην ανάπτυξη συστημάτων συστάσεων, επιτρέποντας την αποτύπωση βαθύτερων σημασιολογικών σχέσεων ανάμεσα σε χρήστες και αντικείμενα. Η μετάβαση από παραδοσιακά, στατικά φίλτρα περιεχομένου σε συστήματα που βασίζονται σε προεκπαιδευμένα γλωσσικά μοντέλα ενισχύει τη δυνατότητα των συστημάτων να κατανοούν πολυσήμαντο και ανομοιογενές περιεχόμενο, όπως ετικέτες παραγώμενες από τον χρήστη, βιογραφικά καλλιτεχνών ή αποσπάσματα στίχων. Όπως επισημαίνουν οι Zhao et al. [12], τα LLMs χρησιμοποιούνται συχνά ως μετασχηματιστές χαρακτηριστικών (feature encoders), με μοντέλα όπως το BERT (Bidirectional Encoder Representations from Transformers) να αξιοποιούνται για την παραγωγή πλούσιων σημασιολογικών αναπαραστάσεων σε κειμενοκεντρικά σενάρια συστάσεων. Ιδιαίτερα όταν εφαρμόζονται τεχνικές fine-tuning ή prompting, τα μοντέλα αυτά μπορούν να μεταφερθούν και να προσαρμοστούν σε tasks προσωποποιημένων συστάσεων, διατηρώντας παράλληλα τις γλωσσικές τους ικανότητες. Στην εκτενή επισκόπηση του Said [13], αναλύεται πώς τα LLMs, συμπεριλαμβανομένων των μοντέλων τύπου BERT και Sentence-BERT, μπορούν να χρησιμοποιηθούν για την επεξήγηση των συστάσεων (explainability), αντιμετωπίζοντας το διαχρονικό ζήτημα της διαφάνειας. Οι Sentence-BERT αναπαραστάσεις, ειδικότερα, επιτρέπουν την χαρτογράφηση ερωτήσεων και τεκμηρίων σε κοινό σημασιολογικό χώρο, γεγονός που καθιστά τις απαντήσεις των συστημάτων πιο κατανοητές για τους

τελικούς χρήστες. Αξιοσημείωτη συμβολή στο πεδίο αποτελεί και η έρευνα των Yun και Lim [14], η οποία μεταφέρει το επίκεντρο από την τεχνική ακρίβεια στην εμπειρία του χρήστη (User Experience - UX) σε διαλογικά recommendation περιβάλλοντα που βασίζονται σε LLMs. Μέσω μιας εθνογραφικής μελέτης διάρκειας τριών εβδομάδων, διερευνώνται οι πραγματικές αλληλεπιδράσεις χρηστών με conversational agents που προτείνουν μουσική, όπως τα custom GPTs. Τα αποτελέσματα αποκαλύπτουν ότι οι LLM-based CRS (Conversational Recommender Systems) συμβάλλουν στη διαμόρφωση πιο στοχευμένων προτάσεων μέσω της ενίσχυσης της ικανότητας των χρηστών να διατυπώνουν τις προτιμήσεις τους με μεγαλύτερη σαφήνεια και να εξερευνούν άγνωστες πτυχές της μουσικής τους ταυτότητας. Συνοψίζοντας, οι σύγχρονες εφαρμογές LLMs σε συστήματα με βάση το περιεχόμενο ή διαλογικά συστήματα μουσικών συστάσεων προσφέρουν σημαντικές δυνατότητες: από την παραγωγή ισχυρών ενσωματώσεων (embeddings) σημασιολογικά πλούσιων περιγραφών μέσω Sentence-BERT, μέχρι τη βελτίωση της εξηγησιμότητας και της αλληλεπίδρασης με τον τελικό χρήστη. Ωστόσο, ακόμα και αυτά τα μοντέλα υστερούν σε περιβάλλοντα όπου τα δεδομένα δεν περιέχουν επαρκή πληροφορία. Οι πρακτικές και θεωρητικές προεκτάσεις αυτών των εξελίξεων αναμένεται να επαναπροσδιορίσουν τόσο τον τρόπο ανάπτυξης recommender systems όσο και τον ρόλο της φυσικής γλώσσας στην προσωποποιημένη πληροφορία.

3.4 Μετρικές Αξιολόγησης

Για την αξιολόγηση των συστημάτων συστάσεων, υιοθετήθηκαν μετρικές κατάταξης κατάλληλες για σενάρια *implicit feedback*. Οι κυριότερες ήταν οι *Precision@K*, *Recall@K* και *Normalized Discounted Cumulative Gain (nDCG)*, οι οποίες μετρούν την ικανότητα του συστήματος να φέρει σχετικά αντικείμενα στις πρώτες θέσεις μιας ταξινομημένης λίστας [10]. Οι μετρικές αυτές χρησιμοποιήθηκαν τόσο κατά τη φάση πειραματικής αξιολόγησης όσο και για τη σύγκριση μεταξύ διαφορετικών αλγοριθμικών προσεγγίσεων (CF, CB, SVD, NeuMF, BERT).

3.4.1 Precision@K

Η μετρική *Precision@K* χρησιμοποιείται ευρέως στην αξιολόγηση συστημάτων συστάσεων, ειδικά σε περιπτώσεις όπου τα δεδομένα χρήστη περιορίζονται σε *implicit feedback*. Εκφράζει το ποσοστό των σχετικών αντικειμένων ανάμεσα στα πρώτα K συσταθέντα αντι-

κείμενα, δηλαδή εκείνα που το σύστημα θεωρεί πιο πιθανό να ενδιαφέρουν τον χρήστη. Η υψηλή τιμή *Precision@K* σημαίνει ότι οι κορυφαίες K προτάσεις είναι ακριβείς και ταιριάζουν με τις πραγματικές προτιμήσεις του χρήστη. Στην εργασία των Berkovsky et al. [10], η *Precision@K* αξιολογήθηκε σε συνδυασμό με τεχνικές συνεργατικού φιλτραρίσματος, επιβεβαιώνοντας ότι είναι μια αποτελεσματική μετρική για την αξιολόγηση της ικανότητας του συστήματος να παρέχει εξατομικευμένες προτάσεις.

$$\text{Precision@K} = \frac{\#\{\text{σχετικά μέσα στα } K \text{ προτεινόμενα}\}}{K} \tag{3.5}$$

3.4.2 Recall@K

Η *Recall@K*, σε αντίθεση με την Precision, επικεντρώνεται στην ικανότητα του συστήματος να εντοπίσει όσο το δυνατόν περισσότερα σχετικά αντικείμενα, ανεξαρτήτως της θέσης τους στη λίστα. Πιο συγκεκριμένα, πρόκειται για το ποσοστό των σχετικών αντικειμένων που περιλαμβάνονται μέσα στα πρώτα K προτεινόμενα προς το σύνολο των σχετικών για τον χρήστη αντικειμένων. Στην πράξη, αυτό βοηθά στην εκτίμηση της «ευαισθησίας» του συστήματος. Όπως περιγράφεται από τους Berkovsky et al. [10], η *Recall@K* είναι ιδιαίτερα χρήσιμη σε σενάρια όπου η πληρότητα είναι πιο σημαντική από την ακρίβεια – για παράδειγμα, όταν ένας χρήστης αναζητά πλήθος επιλογών και όχι απαραίτητα μόνο τις κορυφαίες.

$$\text{Recall@K} = \frac{\#\{\text{σχετικά μέσα στα } K \text{ προτεινόμενα}\}}{\#\{\text{συνολικά σχετικά στο test}\}} \tag{3.6}$$

3.4.3 Normalized Discounted Cumulative Gain (nDCG)

Η *nDCG* αποτελεί μια πιο εκλεπτυσμένη μετρική αξιολόγησης που λαμβάνει υπόψη όχι μόνο αν τα προτεινόμενα αντικείμενα είναι σχετικά, αλλά και τη θέση στην οποία εμφανίζονται μέσα στη λίστα. Αυτό επιτρέπει την ποσοτικοποίηση της ποιότητας κατάταξης του συστήματος, δίνοντας μεγαλύτερη βαρύτητα στα πρώτα αποτελέσματα της λίστας. Ο υπολογισμός της βασίζεται στο *Cumulative Gain (CG)*, το οποίο προσαρμόζεται μέσω του συντελεστή *discount (DCG)*, και τελικά κανονικοποιείται ως *nDCG* ώστε να επιτρέπεται η σύγκριση μεταξύ διαφορετικών μοντέλων. Ο Schedl [7] τονίζει πως η *nDCG* αποτελεί αναπόσπαστο εργαλείο στην αξιολόγηση *content-based* συστημάτων μουσικών συστάσεων, καθώς επιτρέπει την ανάλυση της κατανομής της συνάφειας στο σύνολο της λίστας.

$$\text{DCG@K} = \sum_{i=1}^{K} \frac{\text{συνάφεια στη θέση } i}{\log_2(i+1)} \tag{3.7}$$

$$\text{nDCG@K} = \frac{\text{DCG@K}}{\text{IDCG@K}} \tag{3.8}$$

3.4.4 Hit Rate@K

Η *Hit Rate@K* είναι μια απλή αλλά χρήσιμη μετρική, η οποία υποδεικνύει αν το σύστημα έχει συμπεριλάβει τουλάχιστον ένα σχετικό αντικείμενο εντός των κορυφαίων K προτάσεων. Πρόκειται για δυαδική αξιολόγηση (hit ή miss) και χρησιμοποιείται συχνά για να δώσει μια γενική εικόνα της επιτυχίας ενός συστήματος συστάσεων. Όπως επισημαίνεται από τους S.Oramas et al. [15], η *Hit Rate* είναι ιδιαίτερα σημαντική σε *cold-start* σενάρια, όπου οι διαθέσιμες πληροφορίες είναι περιορισμένες και η ύπαρξη έστω και ενός επιτυχούς *recommendation* στις πρώτες θέσεις αποτελεί ένδειξη λειτουργικής ακρίβειας.

$$\text{HitRate@K} = \frac{\#\{\text{χρήστες με} \geq 1 \text{ σχετικό στα top-}K\}}{\#\{\text{συνολικοί χρήστες}\}} \tag{3.9}$$

3.4.5 Mean Reciprocal Rank (MRR)

Η *MRR (Mean Reciprocal Rank)* αποτελεί μια μετρική που επικεντρώνεται στην πρώτη εμφάνιση ενός σχετικού αντικειμένου. Συγκεκριμένα, υπολογίζει τον μέσο όρο του αντίστροφου της θέσης του πρώτου σχετικού αντικειμένου για όλους τους χρήστες. Η τιμή της είναι υψηλότερη όταν οι σχετικές προτάσεις εμφανίζονται στις πρώτες θέσεις της λίστας, επομένως ευνοεί τα μοντέλα που καταφέρνουν να «πιάσουν» γρήγορα την προτίμηση του χρήστη. Σύμφωνα με τους D.Liang et al. [16], η *MRR* προσφέρει μια ξεκάθαρη εικόνα για την αποτελεσματικότητα των πιθανοτικών (*probabilistic*) μοντέλων, όπως οι *Variational Autoencoders*, στο να προβλέπουν έγκαιρα την πρώτη σχετική επιλογή. Επιπλέον, οι Berkovsky et al. [10] τονίζουν τη χρήση της *MRR* ως χρήσιμη συμπληρωματική μετρική, σε συνδυασμό με την *Recall* και την *Precision*, για πλήρη ποιοτική αποτίμηση των αποτελεσμάτων.

$$\text{MRR} = \frac{1}{\#\{\text{χρήστες}\}} \sum_{u} \frac{1}{\text{θέση πρώτου σχετικού για τον χρήστη } u} \tag{3.10}$$

3.4.6 Coverage

Η *Coverage* αξιολογεί το εύρος των αντικειμένων που προτείνονται από το σύστημα συστάσεων, μετρώντας το ποσοστό των μοναδικών αντικειμένων από το συνολικό κατάλογο που εμφανίζονται τουλάχιστον μία φορά στις συστάσεις. Η μετρική αυτή έχει ιδιαίτερη σημασία για την ποικιλομορφία και την ικανότητα του συστήματος να προτείνει μη-δημοφιλή, αλλά εν δυνάμει σχετικά αντικείμενα. Στην εργασία των Dinnissen και Bauer [6], η *Coverage* αναδεικνύεται ως κρίσιμη μετρική για την εκτίμηση της «δικαιοσύνης» και της ισοκατανομής των συστάσεων. Υψηλή κάλυψη υποδηλώνει ότι το σύστημα μπορεί να αξιοποιήσει μεγάλο μέρος του συνόλου των αντικειμένων, αποφεύγοντας να εστιάζει μόνο στα δημοφιλή ή τα ήδη πολυπροβαλλόμενα. Αυτό είναι εξαιρετικά σημαντικό σε εφαρμογές μουσικής, όπου η ανακάλυψη νέου περιεχομένου αποτελεί βασικό ζητούμενο για την ενίσχυση της εμπειρίας του χρήστη.

$$\text{Coverage@K} = \frac{\#\{\text{μοναδικά αντικείμενα που προτάθηκαν στα top-}K\}}{\#\{\text{συνολικά αντικείμενα στον κατάλογο}\}} \tag{3.11}$$

Κεφάλαιο 4

Χαρακτηριστικά, Προεπεξεργασία και Διαχωρισμός Συνόλου Δεδομένων

4.1 Σύνολο Δεδομένων

Για τις ανάγκες της παρούσας εργασίας χρησιμοποιήθηκε το σύνολο δεδομένων HetRec2011 Last.fm 2K, το οποίο διατίθεται δωρεάν από την ερευνητική ομάδα Information Retrieval Group του Universidad Autónoma de Madrid στο πλαίσιο του 2ου Διεθνούς Εργαστηρίου HetRec 2011. Το σύνολο του διαδικτυακού συστήματος Last.fm περιλαμβάνει δεδομένα από 1.892 χρήστες, 17.632 καλλιτέχνες και 92.834 αλληλεπιδράσεις μεταξύ τους και περιέχει 9.749 πληροφορίες ετικετών (tags) και διάφορες πληροφορίες κοινωνικών σχέσεων, όπως φαίνεται και στο Σχήμα 4.1. Επιπλέον, στο Σχήμα 4.2 αποτυπώνεται η πυκνότητα του συνόλου δεδομένων, η οποία ανέρχεται μόλις στο 0,28%. Πρόκειται για ένα δημοφιλές benchmark dataset για την αξιολόγηση μουσικών συστημάτων σύστασης, καθώς συνδυάζει αλληλεπιδράσεις χρήστη-αντικειμένου με σημασιολογικά και κοινωνικά χαρακτηριστικά. Τα δεδομένα είναι οργανωμένα σε έξι αρχεία τύπου .dat, το καθένα από τα οποία περιλαμβάνει διαφορετικού τύπου πληροφορίες. Στην παρούσα εργασία χρησιμοποιήθηκαν τέσσερα από αυτά: το artists.dat που περιέχει βασικά στοιχεία των καλλιτεχνών, το user_artists.dat που καταγράφει τις αλληλεπιδράσεις των χρηστών με καλλιτέχνες, το tags.dat που περιέχει το λεξικό των ετικετών, και το user_taggedartists.dat που περιγράφει ποιες ετικέτες απέδωσε κάθε χρήστης σε συγκεκριμένους καλλιτέχνες. Τα υπόλοιπα δύο αρχεία (user_friends.dat και user_taggedartists-timestamps.dat) φορτώθηκαν για πληρότητα αλλά δεν χρησιμοποιήθηκαν στους αλγορίθμους σύστασης.

Σχήμα 4.1: Στατιστικά Δεδομένων προ επεξεραγασίας

Σχήμα 4.2: Πυκνότητα Δεδομένων προ επεξεργασίας

4.2 Προεπεξεργασία Δεδομένων

Στην ενότητα αυτή περιγράφονται αναλυτικά τα βήματα που ακολουθήθηκαν για τον καθαρισμό και την προετοιμασία του συνόλου HetRec2011–Last.fm 2K. Στόχος ήταν η εξαγωγή ενός καθαρού και αντιπροσωπευτικού υποσυνόλου που μπορεί να αξιοποιηθεί αποτελεσματικά από τους αλγορίθμους σύστασης. Για κάθε αρχείο του συνόλου δεδομένων γίνεται ξεχωριστή παρουσίαση των στηλών του και του τρόπου αξιοποίησής του. Στη συνέχεια παρουσιάζονται τα αντίστοιχα γραφήματα πυκνότητας στο Σχήμα 4.3 και στατιστικών δεδομένων στο Σχήμα 4.4 μετά την προεπεξεργασία τους και αναλύονται τα βήματα φιλτραρίσματος που οδήγησαν σε αυτά τα αποτελέσματα.

Σχήμα 4.3: Στατιστικά Δεδομένων μετά επεξεραγασίας

Σχήμα 4.4: Πυκνότητα Δεδομένων μετά επεξεργασίας

4.2.1 Αρχείο artists.dat

Το αρχείο αυτό περιέχει βασικές πληροφορίες για τους καλλιτέχνες:

- id: μοναδικό αναγνωριστικό καλλιτέχνη
- name: όνομα καλλιτέχνη
- url: σύνδεσμος προς το προφίλ του καλλιτέχνη στο Last.fm

Κατά τη φόρτωση πραγματοποιήθηκε έλεγχος εγκυρότητας των κωδικών artist_id και μετατροπή τους σε ακέραιους. Για τις περιπτώσεις όπου η κωδικοποίηση UTF-8 δεν ήταν συμβατή, χρησιμοποιήθηκε εναλλακτικά η latin1. Από τις πληροφορίες αυτές χρησιμοποιήθηκε κυρίως το πεδίο name, το οποίο αποτέλεσε τη βάση για τη δημιουργία περιγραφών στους αλγορίθμους με βάση το περιεχόμενο. Το πεδίο url αξιοποιήθηκε επικουρικά ως ένδειξη πληρότητας πληροφορίας.

4.2.2 Αρχείο user_artists.dat

Το συγκεκριμένο αρχείο καταγράφει τις αλληλεπιδράσεις των χρηστών με τους καλλιτέχνες:

- userid: μοναδικό αναγνωριστικό χρήστη
- artistid: μοναδικό αναγνωριστικό καλλιτέχνη
- weight: αριθμός ακροάσεων

Το αρχείο αποτέλεσε τον πυρήνα για τη δημιουργία του πίνακα αλληλεπιδράσεων. Πριν τη μετατροπή σε πίνακα πραγματοποιήθηκε φιλτράρισμα, ώστε να διατηρηθούν μόνο οι χρήστες και οι καλλιτέχνες με τουλάχιστον πέντε αλληλεπιδράσεις. Με τον τρόπο αυτό αφαιρέθηκαν σποραδικές εγγραφές που δεν προσέφεραν ουσιαστική πληροφορία ώστε να αυξηθεί η συχνότητα των αλληλεπιδράσεων. Στη συνέχεια κατασκευάστηκαν αντιστοιχίσεις (user_id → user_idx, artist_id → artist_idx) και δημιουργήθηκε αραιός πίνακας αλληλεπιδράσεων (csr_matrix), όπου οι γραμμές αντιστοιχούν σε χρήστες και οι στήλες σε καλλιτέχνες, με τιμές τον αριθμό ακροάσεων.

4.2.3 Αρχείο tags.dat

Το αρχείο αυτό περιέχει το λεξικό των διαθέσιμων ετικετών:

- tagid: μοναδικό αναγνωριστικό ετικέτας
- tagvalue: λεκτική περιγραφή της ετικέτας

Κατά τη φόρτωση έγινε μετατροπή των tag_id σε ακέραιους αριθμούς και έλεγχος εγκυρότητας. Το λεξικό αυτό χρησιμοποιήθηκε για την ενσωμάτωση σημασιολογικών πληροφοριών στα προφίλ καλλιτεχνών.

4.2.4 Αρχείο user_taggedartists.dat

Το αρχείο αυτό καταγράφει τις πράξεις ετικετοποίησης των χρηστών:

- userid: κωδικός χρήστη
- artistid: κωδικός καλλιτέχνη

- tagid: κωδικός ετικέτας
- day, month, year: χρονικά στοιχεία της ετικετοποίησης

Για τις ανάγκες της εργασίας χρησιμοποιήθηκαν κυρίως τα τρία πρώτα πεδία, τα οποία συνέβαλαν στη δημιουργία προφίλ χρηστών και καλλιτεχνών με βάση τα tags. Τα χρονικά πεδία φορτώθηκαν αλλά δεν αξιοποιήθηκαν στους αλγορίθμους. Για να μειωθεί ο θόρυβος, διατηρήθηκαν μόνο οι ετικέτες που εμφανίζονται τουλάχιστον δύο φορές. Μετά το φιλτράρισμα συντέθηκαν περιγραφικά κείμενα για κάθε καλλιτέχνη, συνδυάζοντας το όνομά του με τις πιο συχνές ετικέτες του. Το σύνολο δεδομένων υπέστη φιλτράρισμα ώστε να διατηρηθούν μόνο χρήστες και καλλιτέχνες με ικανοποιητική δραστηριότητα (≥5 αλληλεπιδράσεις). Η ενέργεια αυτή μείωσε σημαντικά το πλήθος των αντικειμένων αλλά αύξησε την πυκνότητα του πίνακα αλληλεπιδράσεων, γεγονός που οδηγεί σε πιο σταθερή εκπαίδευση των αλγορίθμων. Ο Πίνακας 4.1 παρουσιάζει συγκριτικά το μέγεθος του συνόλου δεδομένων πριν και μετά τον καθαρισμό:

Χαρακτηριστικό	**Πριν**	**Μετά**
Χρήστες	1.892	1.874
Καλλιτέχνες	17.632	2.828
Αλληλεπιδράσεις	92.834	71.411
Ετικέτες (χρήση)	9.749	6.790
Πυκνότητα	0,0028	0,0135
Αραιότητα	0,9972	0,9865

Πίνακας 4.1: Συγκριτικά αποτελέσματα πριν και μετά το φιλτράρισμα

4.3 Διαχωρισμός Δεδομένων σε Σύνολα Εκπαίδευσης και Αξιολόγησης

Στο πλαίσιο της ανάπτυξης συστημάτων σύστασης μουσικής, η επιλογή και η υλοποίηση κατάλληλης μεθόδου διαχωρισμού δεδομένων σε σύνολα εκπαίδευσης και αξιολόγησης αποτελεί κρίσιμο στοιχείο για την αξιόπιστη μέτρηση της απόδοσης των αλγορίθμων. Η παρούσα μελέτη παρουσιάζει μια εξειδικευμένη τεχνική διαχωρισμού που έχει σχεδιαστεί ειδικά για τις ιδιαιτερότητες των συστημάτων σύστασης, διασφαλίζοντας ότι κάθε χρήστης συμμετέχει

και στα δύο σύνολα δεδομένων και επιτρέποντας έτσι την πραγματική αξιολόγηση της ικανότητας του συστήματος να προτείνει νέα και σχετικά αντικείμενα. Η τεχνική διαχωρισμού που υιοθετείται αποτελεί μια παραμετροποιημένη εκδοχή της μεθόδου «leave-one-out». Αντί για την αφαίρεση ενός μόνο στοιχείου ανά χρήστη – όπως γίνεται στην κλασική μορφή της μεθόδου – η παρούσα προσέγγιση αφαιρεί ένα καθορισμένο ποσοστό αλληλεπιδράσεων από κάθε χρήστη, ώστε να προκύψουν πιο ρεαλιστικά σενάρια αξιολόγησης. Το ποσοστό αυτό έχει οριστεί στο 20 τοις εκατό επί του συνολικού αριθμού αλληλεπιδράσεων, με εξαίρεση τις περιπτώσεις χρηστών που διαθέτουν μόνο ένα στοιχείο, οπότε δεν πραγματοποιείται καμία μεταφορά, διατηρώντας έτσι τη συνοχή του συνόλου εκπαίδευσης. Παράλληλα, εξαλείφεται και η επικάλυψη μεταξύ του συνόλου εκπαίδευσης και του συνόλου αξιολόγησης, γεγονός που μειώνει σημαντικά τον κίνδυνο διαρροής πληροφορίας, προσδίδοντας αξιοπιστία και εγκυρότητα στα αποτελέσματα της αξιολόγησης.

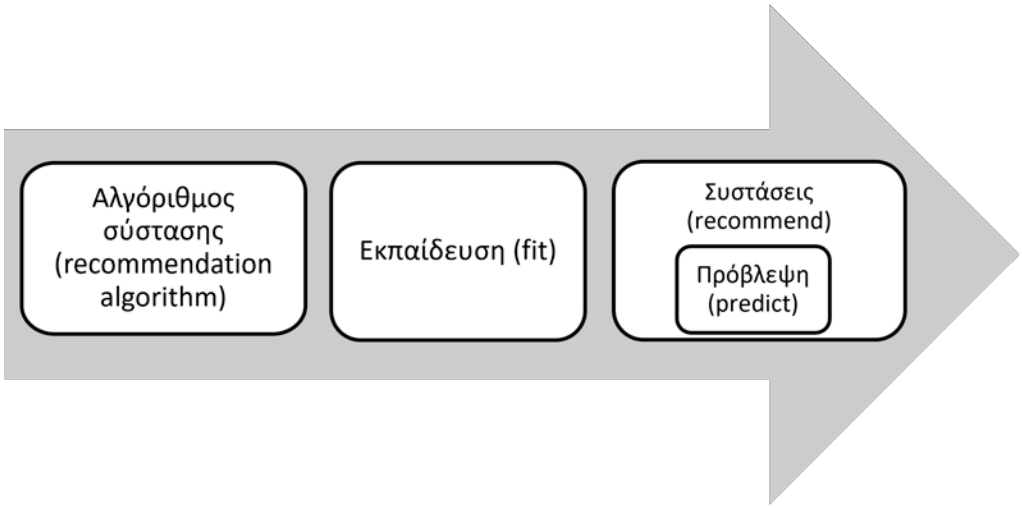
Κεφάλαιο 5

Μεθολογία και Αξιολόγηση Αλγορίθμων

5.1 Αρχιτεκτονική και Βασική Διεπαφή

Η αρχιτεκτονική του συστήματος σύστασης που αναπτύχθηκε και είναι διαθέσιμη στη πλατφόρμα «GitHub» [17], βασίζεται σε μία συνεκτική και επεκτάσιμη δομή, η οποία οργανώνει τις λειτουργίες σε διακριτά επίπεδα. Με τον τρόπο αυτό επιτυγχάνεται σαφής διαχωρισμός αρμοδιοτήτων και ταυτόχρονα εξασφαλίζεται η ευκολία στη διαλειτουργικότητα διαφορετικών τεχνικών. Ο σχεδιασμός επιτρέπει την ενιαία ενσωμάτωση συνεργατικών, νευρωνικών και προσεγγίσεων βασισμένων στο περιεχόμενο μέσα από μια κοινή πλατφόρμα, διατηρώντας σταθερή τη ροή από την είσοδο των δεδομένων έως την παραγωγή των τελικών προτάσεων. Η λογική οργάνωσης ακολουθεί αρχές αρθρωτής ανάπτυξης, όπου κάθε δομικό στοιχείο επιτελεί έναν διακριτό και αυτόνομο ρόλο. Αυτό καλύπτει όλη τη διαδικασία, από την προεπεξεργασία των δεδομένων μέχρι την εξαγωγή χαρακτηριστικών και την κατάταξη των αντικειμένων. Έτσι παρέχεται η δυνατότητα αντικατάστασης ή βελτίωσης μεμονωμένων τμημάτων χωρίς να επηρεάζεται η συνολική λειτουργία, γεγονός που είναι ουσιώδες για πειραματισμούς και μελλοντικές επεκτάσεις. Στο επίπεδο του συνεργατικού φιλτραρίσματος περιλαμβάνονται τρία μοντέλα: το UserKNN, το οποίο υπολογίζει προτιμήσεις με βάση την ομοιότητα μεταξύ χρηστών, το ItemKNN που εστιάζει στην ομοιότητα μεταξύ αντικειμένων και το SVD που αξιοποιεί παραγοντοποίηση πινάκων για την ανίχνευση λανθανουσών σχέσεων στις αλληλεπιδράσεις. Στο επίπεδο των νευρωνικών προσεγγίσεων εντάσσεται το NeuMF, ένα μοντέλο νευρωνικού συνεργατικού φιλτραρίσματος που συνδυάζει στοιχεία παραγοντοποίησης με βαθιά νευρωνικά δίκτυα ώστε να αποτυπώνει μη γραμμικές και πιο πολύπλοκες σχέσεις ανάμεσα σε χρήστες και αντικείμενα. Τέλος, στο επίπεδο του περιεχομένου

αξιοποιείται ένας αλγόριθμος βασισμένος στο Sentence-BERT, ο οποίος επεξεργάζεται τα κειμενικά μεταδεδομένα των καλλιτεχνών και δημιουργεί σημασιολογικές αναπαραστάσεις, βελτιώνοντας την κατανόηση των χαρακτηριστικών του μουσικού περιεχομένου και ενισχύοντας την ποιότητα των συστάσεων. Στο επίκεντρο της αρχιτεκτονικής βρίσκεται μια κοινή διεπαφή, της οποίας οι βασικές λειτουργίες αποτυπώνονται στο Σχήμα 5.1, που καθορίζει τον τρόπο με τον οποίο επικοινωνούν τα μοντέλα συνεργατικού φιλτραρίσματος με το υπόλοιπο σύστημα. Η διεπαφή αυτή ορίζει ενιαίες αρχές για την εκπαίδευση, την πρόβλεψη και τη δημιουργία λιστών συστάσεων, διασφαλίζοντας συνεπή συμπεριφορά και συγκρισιμότητα μεταξύ διαφορετικών μεθόδων. Τα νευρωνικά μοντέλα και τα μοντέλα βασισμένα στο περιεχόμενο διαθέτουν ξεχωριστές διεπαφές, γεγονός που αντανακλά τις ιδιαίτερες απαιτήσεις τους, αλλά παραμένουν ενσωματωμένα σε μία κοινή αρχιτεκτονική που υποστηρίζει την παράλληλη αξιολόγηση και την εύκολη επέκταση με νέες προσεγγίσεις.



Σχήμα 5.1: Ροή Βασικής διεπαφής

5.2 Ανάλυση Τεχνικών Βελτιστοποίησης Συστάσεων

Μια σειρά από τεχνικές βελτιστοποίησης εφαρμόζονται στο σύστημα σύστασης, με στόχο τη βελτίωση της ακρίβειας, της σταθερότητας και της ποικιλίας των αποτελεσμάτων. Οι τεχνικές αυτές λειτουργούν τόσο σε επίπεδο προεπεξεργασίας όσο και κατά την παραγωγή των συστάσεων.

Κεντράρισμα (Centering): Η τεχνική του κεντραρίσματος εφαρμόζεται στους αλγορίθμους UserKNN και ItemKNN όταν χρησιμοποιείται η μετρική Pearson ή όταν η παράμετρος normalization είναι ενεργοποιημένη. Η παράμετρος normalization λειτουργεί ως εναλλακτι-

κός τρόπος για την ενεργοποίηση του κεντραρίσματος. Βασικός της στόχος είναι η εξάλειψη της προκατάληψης που προκαλείται από τις διαφορετικές συνήθειες αξιολόγησης των χρηστών. Για να το πετύχει αυτό, το σύστημα αφαιρεί τον μέσο όρο κάθε χρήστη από τις επιμέρους βαθμολογίες του (π.χ. πόσο περισσότερο ή λιγότερο ακούει κάποιον καλλιτέχνη σε σχέση με το μέσο όρο του). Για παράδειγμα αν ένας χρήστης τείνει να βαθμολογεί τους καλλιτέχνες που έχει ακούσει με υψηλά σκόρ (3 και 4) τότε για την εξάλειψη αυτής της προκατάληψης αφαιρείται ο μέσος όρος (3,5) από όλες τις βαθμολογίες του. Έτσι, οι προβλέψεις βασίζονται σε σχετικές αποκλίσεις και όχι σε απόλυτες τιμές, με αποτέλεσμα η σύγκριση μεταξύ χρηστών να γίνεται πιο αξιόπιστη.

Ποινή Δημοτικότητας (Popularity Penalty / Long-tail Boosting): Στα περισσότερα συστήματα σύστασης παρατηρείται τάση υπερπροώθησης των πιο δημοφιλών αντικειμένων, μειώνοντας τη διαφοροποίηση και οδηγώντας σε φτωχότερες εμπειρίες χρήστη. Η ποινή δημοτικότητας μειώνει τη βαθμολογία ενός αντικειμένου ανάλογα με τη συνολική του δημοτικότητα, με τη χρήση ενός ρυθμιζόμενου εκθετικού συντελεστή. Ο υπολογισμός της δημοτικότητας των αντικειμένων αποτελεί προαπαιτούμενο για την εφαρμογή της ποινής δημοτικότητας. Κατά την εκπαίδευση, για κάθε αντικείμενο καταγράφεται ο συνολικός αριθμός των χρηστών που έχουν αλληλεπιδράσει μαζί του. Οι τιμές αυτές αυξάνονται κατά μία μονάδα για να αποφευχθεί διαίρεση με το μηδέν. Στόχος είναι η ενίσχυση της παρουσίας λιγότερο γνωστών, αλλά πιθανώς πολύ σχετικών, αντικειμένων στη λίστα συστάσεων. Η τεχνική αυτή έχει ιδιαίτερη σημασία σε πεδία όπως η μουσική σύσταση, όπου η ανακάλυψη νέων καλλιτεχνών αποτελεί ζητούμενο. Άρα αν η βαθμολογία ενός αντικειμένου από έναν χρήστη είναι $s_{u,i}$, η συνολική δημοτικότητα του αντικειμένου είναι pop_i και ο εκθετικός συντελεστής είναι ε (με τιμές από 0 έως 1), τότε η τελική βαθμολογία δίνεται από τον τύπο:

$$pop_i = \sum_{u \in U} (interaction_{u,i} > 0) + 1 \tag{5.1}$$

$$s'_{u,i} = \frac{s_{u,i}}{(pop_i)^{\epsilon}} \tag{5.2}$$

Οι τεχνικές αυτές διαμορφώνουν ένα συνεκτικό πλαίσιο βελτιστοποίησης των συστάσεων, παρέχοντας καλύτερη ισορροπία ανάμεσα στην ακρίβεια, την κάλυψη και την ικανότητα γενίκευσης των μοντέλων.

5.3 Αλγόριθμοι συνεργατικού φιλτραρίσματος

Στο πλαίσιο της παρούσας διπλωματικής εργασίας υλοποιήθηκαν και μελετήθηκαν τρεις αλγόριθμοι συνεργατικού φιλτραρίσματος, οι οποίοι εφαρμόζονται σε δεδομένα αλληλεπιδράσεων μεταξύ χρηστών και μουσικών αντικειμένων. Οι συστάσεις βασίζονται στην υπόθεση ότι είτε χρήστες με παρόμοιες προτιμήσεις είτε αντικείμενα με παρόμοιο προφίλ αλληλεπιδράσεων παρουσιάζουν κοινά χαρακτηριστικά, και επομένως μπορούν να χρησιμοποιηθούν για την πρόβλεψη προτιμήσεων. Συγκεκριμένα, εξετάζονται τρεις αλγόριθμοι: ο αλγόριθμος UserKNN , ItemKNN, καθώς και ο SVD.

5.3.1 Εύρεση βέλτιστου αριθμού γειτόνων (k)

Για την επιλογή αριθμού γειτόνων k στους αλγορίθμους γειτνίασης UserKNN και ItemKNN πραγματοποιήθηκε μια σειρά πειραμάτων με αυξανόμενο k (5,10,20). Για κάθε μία από τις τιμές του k εξετάστηκαν όλες οι μετρικές για ενδεικτικό πλήθος συστάσεων 5 και 10. Όπως φαίνεται στους Πίνακες 5.1, 5.2, 5.3 τις καλύτερες επιδόσεις εμφάνισε το πείραμα για k=5, το οποίο και χρησιμοποιήθηκε στη συνέχεια.

Μοντέλο	Precision@5	Recall@5	NDCG@5	Hit_Rate@5	Precision@10	Recall@10	NDCG@10	Hit_Rate@10	MRR	Coverage
UserKNN	0,0370	0,0264	0,0371	0,1631	0,0393	0,0544	0,0464	0,3000	0,1156	0,9965
ItemKNN	0,0201	0,0156	0,0204	0,0952	0,0182	0,0288	0,0241	0,1636	0,0680	0,9968

Πίνακας 5.1: Αποτελέσματα αξιολόγησης για $k = 5$

Μοντέλο	Precision@5	Recall@5	NDCG@5	Hit_Rate@5	Precision@10	Recall@10	NDCG@10	Hit_Rate@10	MRR	Coverage
UserKNN	0,0232	0,0173	0,0233	0,1059	0,0255	0,0359	0,0301	0,2118	0,0839	0,9812
ItemKNN	0,0127	0,0105	0,0139	0,0610	0,0133	0,0222	0,0179	0,1214	0,0513	0,9965

Πίνακας 5.2: Αποτελέσματα αξιολόγησης για $k = 10$

Μοντέλο	Precision@5	Recall@5	NDCG@5	Hit_Rate@5	Precision@10	Recall@10	NDCG@10	Hit_Rate@10	MRR	Coverage
UserKNN	0,0121	0,0085	0,0111	0,0583	0,0148	0,0217	0,0166	0,1337	0,0526	0,9421
ItemKNN	0,0086	0,0070	0,0089	0,0417	0,0071	0,0111	0,0099	0,0674	0,0328	0,9950

Πίνακας 5.3: Αποτελέσματα αξιολόγησης για $k = 20$

5.3.2 Αλγόριθμος KNN με βάση τον χρήστη (UserKNN)

Ο αλγόριθμος αυτός αποσκοπεί στην παραγωγή συστάσεων για κάθε χρήστη, βασιζόμενος στη συμπεριφορά άλλων χρηστών που παρουσιάζουν παρόμοιες προτιμήσεις. Η βασική

του λειτουργία στηρίζεται στη λογική ότι χρήστες που έχουν αντιδράσει με παρόμοιο τρόπο στο παρελθόν, είναι πιθανό να δείξουν ενδιαφέρον για τα ίδια αντικείμενα και στο μέλλον. Ο αλγόριθμος εφαρμόζεται πάνω σε δεδομένα αλληλεπιδράσεων και δεν απαιτεί πληροφορίες για το περιεχόμενο των αντικειμένων.

Βήματα αλγορίθμου

Αρχικά, κατασκευάζεται ένας αραιός πίνακας αλληλεπιδράσεων χρηστών–αντικειμένων, όπου σε κάθε κελί αποθηκεύεται ο αριθμός αλληλεπιδράσεων ενός χρήστη με ένα αντικείμενο. Σε αυτό το σημείο υπολογίζεται η δημοτικότητα κάθε αντικειμένου, η οποία θα χρησιμοποιηθεί αργότερα κατά την παραγωγή συστάσεων. Στη συνέχεια υπολογίζεται, για όλα τα ζεύγη χρηστών, ο βαθμός ομοιότητάς τους βάσει των κοινών αντικειμένων με τα οποία έχουν αλληλεπιδράσει. Η ομοιότητα προκύπτει με τη χρήση ομοιότητας συνημιτόνου, ενώ προαιρετικά μπορεί να ενεργοποιηθεί κανονικοποίηση/κεντράρισμα των τιμών αλληλεπίδρασης ώστε να μειώνεται η μεροληψία χρηστών που τείνουν να εμφανίζουν συστηματικά υψηλότερα ή χαμηλότερα επίπεδα δραστηριότητας. Εφόσον υπολογιστούν οι ομοιότητες, για κάθε χρήστη εντοπίζονται οι k πλησιέστεροι γείτονες, δηλαδή οι πιο όμοιοι χρήστες, με βάση τις καταγεγραμμένες προτιμήσεις τους. Εξετάζεται αν αυτοί έχουν αλληλεπιδράσει με κάποιο αντικείμενο που ο χρήστης δεν έχει δει. Αν υπάρχουν τέτοιοι χρήστες, ο αλγόριθμος υπολογίζει μία προβλεπόμενη τιμή ενδιαφέροντος για το αντικείμενο, λαμβάνοντας υπόψη τις τιμές αλληλεπίδρασης των γειτόνων και το πόσο όμοιοι είναι. Αν δεν υπάρχουν γείτονες που να έχουν σχετική αλληλεπίδραση, τότε χρησιμοποιείται ο μέσος όρος τιμών αλληλεπίδρασης του χρήστη ή του συνόλου. Αφού ολοκληρωθούν οι προβλέψεις για όλα τα αντικείμενα που ο χρήστης δεν έχει ήδη δει, εφαρμόζεται μια ποινή δημοτικότητας ώστε να ενισχυθούν λιγότερο προβεβλημένες επιλογές και να αυξηθεί η ποικιλομορφία στις προτάσεις. Τέλος, τα αντικείμενα ταξινομούνται κατά φθίνουσα σειρά με βάση την εκτιμώμενη προτίμηση και ο αλγόριθμος επιστρέφει τα κορυφαία αποτελέσματα ως συστάσεις.

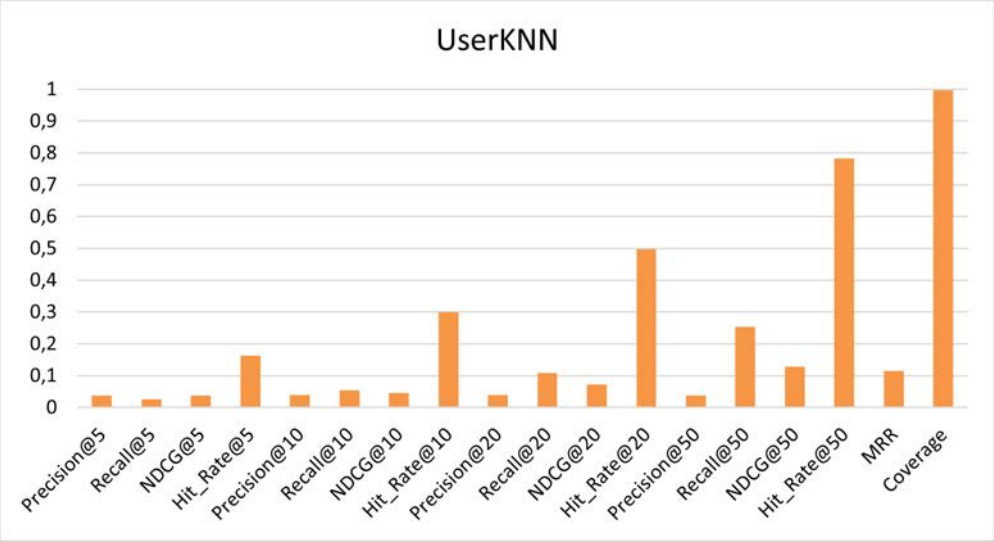
Αποτελέσματα

Στον Πίνακα 5.4 παρατίθενται τα αποτελέσματα του αλγορίθμου BERT σε όλες τις μετρικές για αυξανόμενο αριθμό συστάσεων. Επιπλέον, στο Σχήμα 5.2 εμφανίζονται τα ίδια αποτελέσματα και σε μορφή γραφήματος γραμμής για την οπτικοποίηση και ευκολότερη κατανόησή τους.

Μετρική	@5	@10	@20	@50
Precision	0,0370	0,0393	0,0392	0,0371
Recall	0,0264	0,0544	0,1079	0,2541
NDCG	0,0371	0,0464	0,0722	0,1277
Hit_Rate	0,1631	0,3000	0,4968	0,7824

Μετρική	Τιμή
MRR	0,1156
Coverage	0,9965

Πίνακας 5.4: Αποτελέσματα αξιολόγησης UserKNN



Σχήμα 5.2: Γραφική αναπαράσταση μετρικών UserKNN

Παρατηρήσεις

Ο αλγόριθμος UserKNN εμφανίζει μέτριες επιδόσεις στις περισσότερες μετρικές, με περιορισμένη ακρίβεια στις κορυφαίες θέσεις. Ωστόσο, παρατηρείται ότι καθώς αυξάνεται το πλήθος των συστάσεων, η αποτελεσματικότητα του αλγορίθμου βελτιώνεται σημαντικά. Ιδιαίτερα αξιοσημείωτο είναι το γεγονός ότι η κάλυψη φτάνει σε υψηλά επίπεδα (πάνω από 99%), ξεπερνώντας αρκετούς άλλους αλγόριθμους και δείχνοντας την ικανότητά του να εξερευνά ένα ευρύ φάσμα καλλιτεχνών. Αντίστοιχα, το Hit Rate σε μεγαλύτερες λίστες ξεπερνά τα τρία τέταρτα των χρηστών, γεγονός που αναδεικνύει ότι, αν και ο UserKNN υστερεί στην ακρίβεια, μπορεί να εντοπίζει σχετικές προτάσεις όταν του δίνεται μεγαλύτερο περιθώριο. Η

μέθοδος αυτή παραμένει απλή και εύκολα ερμηνεύσιμη, με καλύτερη απόδοση όταν υπάρχει επαρκές ιστορικό χρήσης. Παράλληλα, η απόδοσή του επηρεάζεται έντονα από τον βαθμό αραιότητας των δεδομένων, ενώ όπως και στα περισσότερα συνεργατικά συστήματα, το πρόβλημα της ψυχρής εκκίνησης παραμένει σημαντικός περιορισμός. Για να ελεγχθεί αν βελτιώνονται οι επιδόσεις στην αντίθετη κατάσταση, δηλαδή όταν η ομοιότητα μεταφέρεται από το επίπεδο των χρηστών σε αυτό των αντικειμένων εφαρμόζεται στη συνέχεια ο ItemKNN.

5.3.3 Αλγόριθμος KNN με βάση το Αντικείμενο (ItemKNN)

Ο αλγόριθμος ItemKNN προτείνει αντικείμενα σε κάθε χρήστη, όχι με βάση τις προτιμήσεις άλλων χρηστών, αλλά με βάση την ομοιότητα μεταξύ των αντικειμένων που έχει ήδη αξιολογήσει και άλλων που δεν έχει ακόμα δει. Η βασική ιδέα στηρίζεται στην υπόθεση ότι αντικείμενα που έχουν προσελκύσει χρήστες με παρόμοιες μουσικές προτιμήσεις είναι πιθανό να είναι παρόμοια μεταξύ τους.

Βήματα αλγορίθμου

Αρχικά, κατασκευάζεται ένας αραιός πίνακας αλληλεπιδράσεων αντικειμένων-χρηστών, όπου σε κάθε κελί αποθηκεύεται ο αριθμός αλληλεπιδράσεων ενός αντικειμένου με έναν χρήστη. Σε αυτό το σημείο υπολογίζεται η δημοτικότητα κάθε αντικειμένου, η οποία θα χρησιμοποιηθεί αργότερα κατά την παραγωγή συστάσεων. Στη συνέχεια υπολογίζεται, για όλα τα ζεύγη αντικειμένων, ο βαθμός ομοιότητάς τους βάσει των κοινών χρηστών οι οποίοι έχουν αλληλεπιδράσει με αυτά. Η ομοιότητα προκύπτει με τη χρήση ομοιότητας συνημιτόνου, ενώ προαιρετικά μπορεί να ενεργοποιηθεί κανονικοποίηση/κεντράρισμα των τιμών αλληλεπίδρασης ώστε να μειώνεται η μεροληψία χρηστών που τείνουν να εμφανίζουν συστηματικά υψηλότερα ή χαμηλότερα επίπεδα δραστηριότητας.

Έπειτα για κάθε χρήστη, ο αλγόριθμος εντοπίζει τα αντικείμενα με τα οποία έχει αλληλεπιδράσει. Για καθένα από αυτά τα αντικείμενα, εντοπίζονται τα k πιο παρόμοια (μεγίστης ομοιότητας) αντικείμενα στο σύνολο. Κατόπιν, υπολογίζεται το εκτιμώμενο score (τιμή ενδιαφέροντος) για κάθε υποψήφιο αντικείμενο που ο χρήστης δεν έχει δει, βάσει των τιμών αλληλεπίδρασης των παρόμοιων αντικειμένων και το βαθμό ομοιότητας. Αν δεν υπάρχουν διαθέσιμα παρόμοια αντικείμενα που να έχει αξιολογήσει ο χρήστης, τότε η πρόβλεψη βασίζεται στον μέσο όρο τιμών αλληλεπίδρασης του αντικειμένου ή στον συνολικό μέσο όρο του dataset. Αφού ολοκληρωθεί το παραπάνω βήμα εφαρμόζεται μια ποινή δημοτικότητας

ώστε να ενισχυθούν λιγότερο προβεβλημένες επιλογές και να αυξηθεί η ποικιλομορφία στις προτάσεις. Τέλος, τα αντικείμενα που δεν έχει δει ο χρήστης ταξινομούνται κατά φθίνον εκτιμώμενο score, και προτείνονται τα κορυφαία ως συστάσεις.

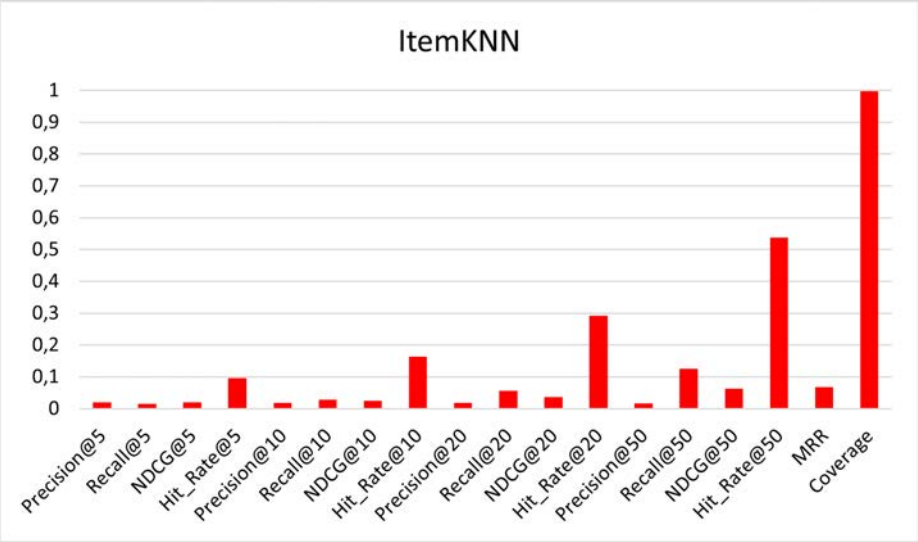
Αποτελέσματα

Στον Πίνακα 5.5 παρατίθενται τα αποτελέσματα του αλγορίθμου BERT σε όλες τις μετρικές για αυξανόμενο αριθμό συστάσεων. Επιπλέον, στο Σχήμα 5.3 εμφανίζονται τα ίδια αποτελέσματα και σε μορφή γραφήματος γραμμής για την οπτικοποίηση και ευκολότερη κατανόησή τους.

Μετρική	**@5**	**@10**	**@20**	**@50**
Precision	0,0201	0,0182	0,0182	0,0170
Recall	0,0156	0,0288	0,0560	0,1247
NDCG	0,0204	0,0241	0,0368	0,0625
Hit_Rate	0,0952	0,1636	0,2925	0,5380

Μετρική	**Τιμή**
MRR	0,0680
Coverage	0,9968

Πίνακας 5.5: Αποτελέσματα αξιολόγησης ItemKNN



Σχήμα 5.3: Γραφική αναπαράσταση μετρικών ItemKNN

Παρατηρήσεις

Ο αλγόριθμος ItemKNN παρουσιάζει χαμηλή ακρίβεια στις κορυφαίες θέσεις, κάτι που είναι αναμενόμενο λόγω της μεγάλης αραιότητας του πίνακα αλληλεπιδράσεων. Ωστόσο, ξεχωρίζει σε δύο σημεία: πρώτον, στο πολύ υψηλό Coverage που φτάνει σχεδόν το 100%, την υψηλότερη τιμή ανάμεσα σε όλους τους αλγορίθμους, γεγονός που αποδεικνύει την ικανότητά του να εξερευνά σχεδόν ολόκληρο το σύνολο των καλλιτεχνών. Δεύτερον, το Hit Rate σε μεγαλύτερες λίστες (50 συστάσεων) ξεπερνά το 50%, ένδειξη ότι, αν και οι πρώτες προτάσεις είναι περιορισμένες, σε εκτενέστερες λίστες οι μισοί χρήστες λαμβάνουν τουλάχιστον μία σχετική σύσταση. Ωστόσο, η χαμηλή ακρίβεια στις κορυφαίες θέσεις καταδεικνύει τον περιορισμό του όταν πρόκειται για την παροχή απολύτως συναφών συστάσεων. Στη συνέχεια εξετάζεται ο αλγόριθμος SVD, για να εξερευνηθεί η επιρροή ύπαρξης μοτίβων στις αλληλεπιδράσεις των χρηστών στη παραγωγή πιο σχετικών συστάσεων από το σύστημα.

5.3.4 Αλγόριθμος Παραγοντοποίησης Πινάκων (SVD)

Ο αλγόριθμος SVD αξιοποιεί τεχνικές παραγοντοποίησης πίνακα (matrix factorization) για την παραγωγή εξατομικευμένων συστάσεων, βασιζόμενος στη λογική των λανθάντων παραγόντων (latent factors) . Σε αντίθεση με τους αλγορίθμους KNN, οι οποίοι στηρίζονται στην ομοιότητα μεταξύ χρηστών ή αντικειμένων, ο SVD στοχεύει στην ανακάλυψη μοτίβων στο σύνολο των αλληλεπιδράσεων, αναλύοντας τον αρχικό πίνακα χρήστη–αντικειμένου σε μικρότερες διαστάσεις. Η μέθοδος αυτή έχει τη δυνατότητα να καταγράψει λανθάνουσες προτιμήσεις και ιδιότητες που δεν είναι άμεσα παρατηρήσιμες, οδηγώντας σε προβλέψεις υψηλής ποιότητας ακόμη και σε περιβάλλοντα με αραιά δεδομένα.

Βήματα αλγορίθμου

Η διαδικασία ξεκινά με τη δημιουργία ενός αραιού πίνακα αλληλεπιδράσεων χρηστών–αντικειμένων, όπου κάθε τιμή αποθηκεύει το πλήθος των αλληλεπιδράσεων ενός χρήστη με ένα αντικείμενο. Υπολογίζεται η δημοτικότητα κάθε αντικειμένου, η οποία θα χρησιμοποιηθεί κατά την παραγωγή συστάσεων και ο πίνακας περνά απευθείας στον αλγόριθμο SVD. Στη συνέχεια, εφαρμόζεται παραγοντοποίηση πίνακα με χρήση της μεθόδου Singular Value Decomposition (SVD). Η διαδικασία αυτή διασπά τον πίνακα σε δύο μικρότερους πίνακες, οι οποίοι αντιστοιχούν σε προφίλ χρηστών και ιδιότητες αντικειμένων. Μετά την εκπαίδευση του μοντέλου, υπολογίζονται οι τιμές ενδιαφέροντος (scores) για όλα τα ζεύγη αντικειμέ-

νων-χρηστών μέσω του εσωτερικού γινομένου των πινάκων που έχουν προκύψει από τη παραπάνω διαδικασία διάσπασης. Για κάθε χρήστη, επιλέγονται τα αντικείμενα που δεν έχει ακόμα αλληλεπιδράσει και εφαρμόζεται ποινή δημοτικότητας, ώστε να ενισχυθούν λιγότερο προβεβλημένες επιλογές και να αυξηθεί η ποικιλομορφία στις προτάσεις. Τέλος, κατατάσσονται κατά φθίνουσα σειρά σύμφωνα με το score πρόβλεψης. Ο αλγόριθμος επιστρέφει τα κορυφαία αποτελέσματα ως προτεινόμενα αντικείμενα.

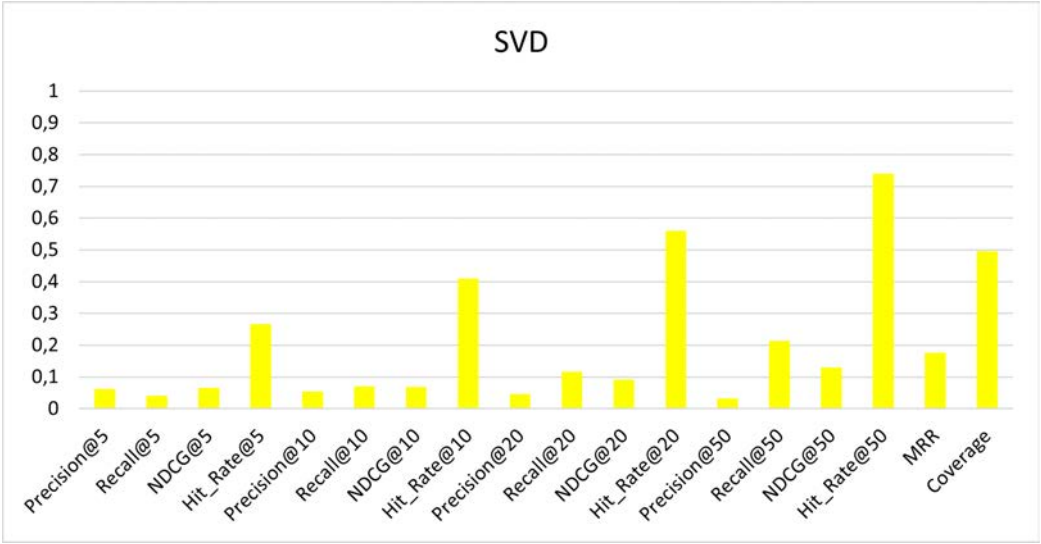
Αποτελέσματα

Στον Πίνακα 5.6 παρατίθενται τα αποτελέσματα του αλγορίθμου BERT σε όλες τις μετρικές για αυξανόμενο αριθμό συστάσεων. Επιπλέον, στο Σχήμα 5.4 εμφανίζονται τα ίδια αποτελέσματα και σε μορφή γραφήματος γραμμής για την οπτικοποίηση και ευκολότερη κατανόησή τους.

Μετρική	**@5**	**@10**	**@20**	**@50**
Precision	0,0627	0,0540	0,0452	0,0330
Recall	0,0406	0,0706	0,1171	0,2136
NDCG	0,0660	0,0684	0,0917	0,1292
Hit_Rate	0,2674	0,4102	0,5610	0,7406

Μετρική	**Τιμή**
MRR	0,1758
Coverage	0,4953

Πίνακας 5.6: Αποτελέσματα αξιολόγησης SVD



Σχήμα 5.4: Γραφική αναπαράσταση μετρικών SVD

Παρατηρήσεις

Ο αλγόριθμος παραγοντοποίησης πίνακα SVD υπερέχει σαφώς έναντι των μεθόδων γειτνίασης κυρίως στα μικρότερα πλήθη συστάσεων, καταγράφοντας εντυπωσιακά αποτελέσματα στις βασικές μετρικές. Στις πρώτες πέντε συστάσεις, η ακρίβεια φτάνει περίπου στο 6,3%, ενώ η ανάκληση ξεπερνά το 4% και ο δείκτης NDCG βρίσκεται στο 0,066. Το ποσοστό χρηστών που λαμβάνουν τουλάχιστον μία σχετική πρόταση ανέρχεται στο 26,7%, δηλαδή περισσότεροι από ένας στους τέσσερις, επίδοση που αποτελεί άλμα σε σχέση με τον UserKNN και τον ItemKNN. Η ανοδική πορεία συνεχίζεται όσο αυξάνεται ο αριθμός των συστάσεων, με το Hit Rate να ξεπερνά το 74% στις 50 κορυφαίες προτάσεις, γεγονός που αποδεικνύει την ικανότητα του μοντέλου να εντοπίζει μεγάλο μέρος των πραγματικών προτιμήσεων των χρηστών. Επιπλέον, ο δείκτης MRR (0,176) υποδηλώνει ότι η πρώτη επιτυχημένη σύσταση εμφανίζεται κατά μέσο όρο στην 6η θέση, στοιχείο που αναδεικνύει την ιδιαίτερα υψηλή αποτελεσματικότητα του SVD. Το μοναδικό πεδίο στο οποίο υστερεί είναι το Coverage, το οποίο περιορίζεται περίπου στο 49%, δείχνοντας ότι το μοντέλο τείνει να προτείνει ένα πιο περιορισμένο υποσύνολο καλλιτεχνών, εστιάζοντας όμως σε αυτούς με τη μεγαλύτερη σχετικότητα. Ο αλγόριθμος SVD αποδεικνύεται ιδιαίτερα αποτελεσματικός σε περιβάλλοντα με επαρκή όγκο δεδομένων, καθώς μπορεί να αποτυπώσει πολύπλοκες λανθάνουσες σχέσεις που δεν γίνονται εύκολα αντιληπτές από μεθόδους βασισμένες στην ομοιότητα. Η παραγοντοποίηση συμβάλλει επίσης στη μείωση του θορύβου και περιορίζει το πρόβλημα της υπερπροσαρμογής, συμπιέζοντας την πληροφορία σε λανθάνοντες παράγοντες. Ωστόσο, η

απόδοση του εξαρτάται έντονα από την ποιότητα των δεδομένων, ενώ το πρόβλημα της ψυχρής εκκίνησης παραμένει, καθώς νέοι χρήστες ή καλλιτέχνες χωρίς ιστορικό δεν μπορούν να ενταχθούν εύκολα στο μοντέλο. Τέλος, η επιλογή του αριθμού λανθανουσών διαστάσεων και των παραμέτρων κανονικοποίησης είναι κρίσιμη, καθώς επηρεάζει άμεσα την ισορροπία ανάμεσα στην ακρίβεια και στη γενίκευση των συστάσεων. Στην επόμενη ενότητα θα εξεταστεί κατά πόσο μια υβριδική προσέγγιση, η οποία συνδυάζει τα πλεονεκτήματα του SVD μαζί με αυτά της βαθιάς μάθησης για την ανακάλυψη μη γραμμικών συσχετίσεων, μπορεί να βελτιώσει τα αποτελέσματα των συστάσεων.

5.4 Αλγόριθμοι Βαθιάς Μάθησης: Neural Matrix Factorization (NeuMF)

Ο αλγόριθμος NeuMF συνδυάζει δύο συμπληρωματικές αρχιτεκτονικές νευρωνικών δικτύων — τη Γενικευμένη Παραγοντοποίηση Μήτρας (Generalized Matrix Factorization) και ένα Πολυεπίπεδο Perceptron (Multi-Layer Perceptron) — ώστε να εκμεταλλεύεται τόσο γραμμικές όσο και μη γραμμικές συσχετίσεις ανάμεσα σε χρήστες και αντικείμενα. Η προσέγγιση αυτή επιτρέπει την αυτόματη εκμάθηση σύνθετων προτύπων προτίμησης χωρίς την ανάγκη χειροποίητων χαρακτηριστικών.

Βήματα αλγορίθμου

Αρχικά, ο πίνακας αλληλεπιδράσεων χρηστών–καλλιτεχνών μετατρέπεται σε έναν αραιό δυαδικό χώρο αναπαράστασης, όπου κάθε θετική εγγραφή δηλώνει ότι ο χρήστης έχει αλληλεπιδράσει τουλάχιστον μία φορά με τον καλλιτέχνη. Για την ουσιαστική εκπαίδευση ενός δυαδικού ταξινομητή απαιτούνται και αρνητικά δείγματα· επομένως πραγματοποιείται αρνητικός δειγματισμός (negative sampling), όπου για κάθε θετικό ζεύγος (u,i) επιλέγονται τυχαία τέσσερα αντικείμενα που ο χρήστης δεν έχει ακούσει, σχηματίζοντας ισορροπημένα mini-batches. Στο τμήμα GMF, κάθε χρήστης και αντικείμενο αποκτά ένα διανύσμα (embedding) χαμηλής διάστασης. Η συνάρτηση παλινδρόμησης υλοποιείται ως στοιχειώδες γινόμενο των δύο embedding, το οποίο αιχμαλωτίζει γραμμικούς συσχετισμούς. Παράλληλα, στο τμήμα MLP τα ίδια embedding διέρχονται από σειρά πλήρως συνδεδεμένων στρωμάτων με συνάρτηση ενεργοποίησης ReLU και Dropout, προκειμένου να μοντελοποιηθούν μη-γραμμικές αλληλεπιδράσεις. Τα δύο διανύσματα εξόδου ενοποιούνται μέσω απλής συνένωσης και περ-

νούν από ένα τελικό νευρώνα sigmoid που αποδίδει την πιθανότητα ο χρήστης να αλληλεπιδράσει με το αντικείμενο. Η εκπαίδευση εκτελείται για 100 εποχές με βελτιστοποίηση Adam (ρυθμός μάθησης 0,001) και συνάρτηση απώλειας δυαδικής διασταύρωσης. Όπως και στις προηγούμενες προσεγγίσεις, έτσι και εδώ εφαρμόζεται ποινή δημοτικότητας, η οποία διαιρεί το τελικό σκορ προτίμησης με την συχνότητα αλληλεπιδράσεων του αντικειμένου υψομένη στο συντελεστή δημοτικότητας (0,3) , ώστε να μειωθεί η κυριαρχία εξαιρετικά διάσημων καλλιτεχνών και να δοθεί έμφαση στη μουσική της «μακράς ουράς». Κατά τη φάση σύστασης, ο NeuMF υπολογίζει τις πιθανότητες για όλα τα αντικείμενα που δεν έχει ήδη ακούσει ο χρήστης. Τα αντικείμενα ταξινομούνται κατά φθίνουσα πιθανότητα και προτείνονται τα κορυφαία k. Οι προηγούμενες αλληλεπιδράσεις μηδενίζονται ώστε να μην επανεμφανιστούν στα αποτελέσματα.

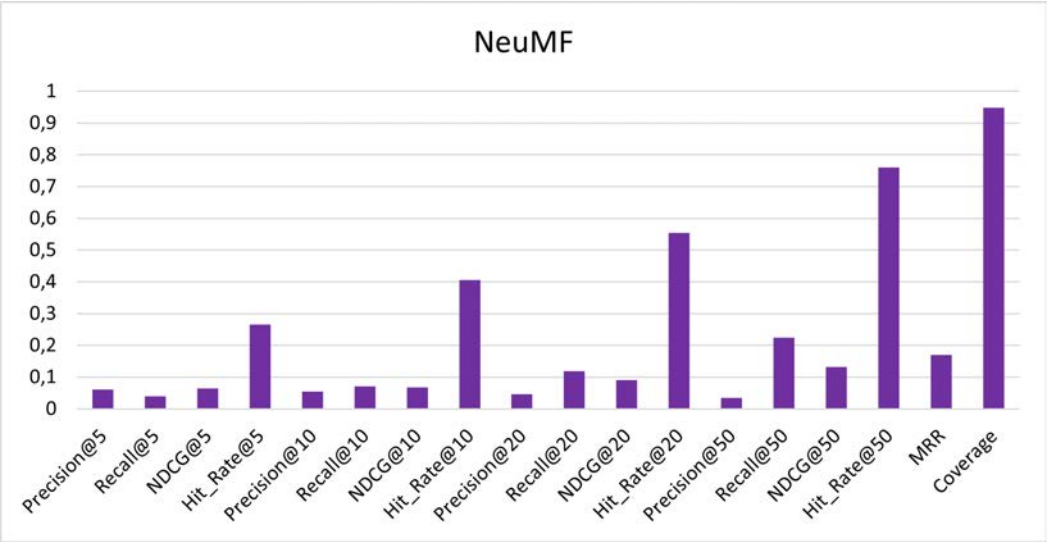
Αποτελέσματα

Στον Πίνακα 5.7 παρατίθενται τα αποτελέσματα του αλγορίθμου BERT σε όλες τις μετρικές για αυξανόμενο αριθμό συστάσεων. Επιπλέον, στο Σχήμα 5.5 εμφανίζονται τα ίδια αποτελέσματα και σε μορφή γραφήματος γραμμής για την οπτικοποίηση και ευκολότερη κατανόησή τους.

Μετρική	**@5**	**@10**	**@20**	**@50**
Precision	0,0614	0,0539	0,0454	0,0342
Recall	0,0399	0,0705	0,1187	0,2237
NDCG	0,0635	0,0671	0,0910	0,1316
Hit_Rate	0,2647	0,4059	0,5545	0,7594

Μετρική	**Τιμή**
MRR	0,1703
Coverage	0,9486

Πίνακας 5.7: Αποτελέσματα αξιολόγησης NeuMF



Σχήμα 5.5: Γραφική αναπαράσταση μετρικών NeuMF

Παρατηρήσεις

Ο αλγόριθμος NeuMF καταγράφει εξαιρετικές επιδόσεις, εφάμιλλες με εκείνες του SVD, ενώ σε ορισμένες μετρικές τις ξεπερνά. Στις πέντε πρώτες συστάσεις σημειώνει υψηλή ακρίβεια, με Precision άνω του 6% και NDCG στο 0,063, τιμές που βρίσκονται πολύ πάνω από τις αντίστοιχες των μεθόδων γειτνίασης. Η απόδοσή του διατηρείται ισχυρή και σε μεγαλύτερες λίστες: το Hit Rate προσεγγίζει το 76% στις 50 συστάσεις, αποδεικνύοντας ότι η πλειονότητα των χρηστών λαμβάνει τουλάχιστον μία σχετική πρόταση. Παράλληλα, ο δείκτης MRR (0,17) δείχνει ότι οι πρώτες επιτυχημένες συστάσεις εμφανίζονται γρήγορα στις λίστες κατάταξης. Ιδιαίτερα αξιοσημείωτο είναι το Coverage, που αγγίζει σχεδόν το 95% γεγονός που φανερώνει την ικανότητα του NeuMF να καλύπτει σχεδόν ολόκληρο το φάσμα των καλλιτεχνών. Ο NeuMF αξιοποιεί ταυτόχρονα τη λογική της παραγοντοποίησης και τα πλεονεκτήματα των βαθιών νευρωνικών δικτύων, κάτι που του επιτρέπει να αποτυπώνει μη γραμμικές σχέσεις και να παρέχει πιο ακριβείς προβλέψεις. Η υψηλή του κάλυψη σε σύγκριση με τον SVD υποδηλώνει ότι δεν περιορίζεται σε λίγους δημοφιλείς καλλιτέχνες αλλά εξερευνά ευρύτερο χώρο επιλογών. Η σταθερά καλή απόδοση στις βασικές μετρικές τον καθιστά έναν από τους πιο αξιόπιστους αλγορίθμους, ιδιαίτερα σε σύνολα δεδομένων με επαρκή όγκο αλληλεπιδράσεων. Παρά τα πλεονεκτήματά του, η πολυπλοκότητα του μοντέλου συνεπάγεται υψηλότερο υπολογιστικό κόστος και μεγαλύτερη ανάγκη για παραμετροποίηση σε σχέση με απλούστερες μεθόδους. Τέλος, παρουσιάζεται ο Sentence-BERT που ανήκει στο φάσμα των μεγάλων γλωσσικών μοντέλων, ώστε η μελέτη να περιλαμβάνει και τις πιο σύγχρονες

προσεγγίσεις.

5.5 Αλγόριθμος με βάση το περιεχόμενο: Sentence-BERT με Last.fm Tags

Ο αλγόριθμος Sentence-BERT με Last.fm Tags υλοποιεί ένα σύστημα σύστασης με βάση το περιεχόμενο που βασίζεται στα κειμενικά μεταδεδομένα του HetRec2011 Last.fm dataset. Αξιοποιεί προ-εκπαιδευμένο Sentence-BERT για την παραγωγή υψηλού επιπέδου ενσωματώσεων (embeddings) καλλιτεχνών και υπολογίζει συστηνόμενα αντικείμενα μέσω ομοιότητας συνημιτόνου. Η προσέγγιση αυτή επιτρέπει την αυτόματη εκμάθηση σημασιολογικών συσχετίσεων μεταξύ καλλιτεχνών χωρίς την ανάγκη χειροποίητων χαρακτηριστικών.

Βήματα αλγορίθμου

Αρχικά, δημιουργείται ένα κειμενικό σώμα (corpus) για κάθε καλλιτέχνη, ο οποίος συνδυάζει το όνομα του καλλιτέχνη με τις πιο συχνές ετικέτες που του έχουν ανατεθεί από τους χρήστες. Για κάθε καλλιτέχνη, εξάγονται οι ετικέτες με συχνότητα μεγαλύτερη ή ίση του κατώτατου ορίου που ορίζεται σε 2, οι οποίες στη συνέχεια μετατρέπονται σε κείμενο με τη μορφή «όνομα καλλιτέχνη. tags: ετικέτα1, ετικέτα2, ...». Ετικέτες με συχνότητα κάτω του κατώτατου ορίου αγνοούνται για αποφυγή θορύβου και βελτίωση της ποιότητας των ενσωματώσεων. Στη συνέχεια, το κείμενο που διαμορφώθηκε παραπάνω τροφοδοτείται στο προεκπαιδευμένο μοντέλο Sentence-BERT (sentence-transformers/all-mpnet-base-v2) για την παραγωγή ενσωματώσεων. Το μοντέλο υπολογίζει τα embeddings σε παρτίδες και τα αποθηκεύει σε cache αρχείο για αποφυγή επανυπολογισμού σε μελλοντικές εκτελέσεις. Επιπλέον, υποστηρίζεται fine-tuning του Sentence-BERT στο Last.fm domain, όπου για κάθε χρήστη με τουλάχιστον δύο αλληλεπιδράσεις δημιουργούνται θετικά παραδείγματα από ζεύγη καλλιτεχνών που έχει ακούσει. Τα embeddings κανονικοποιούνται σε μοναδιαία διανύσματα για απλοποίηση του υπολογισμού ομοιότητας συνημιτόνου. Επίσης, υπολογίζεται η δημοτικότητα κάθε καλλιτέχνη για εφαρμογή ποινής δημοτικότητας κατά τη φάση σύστασης. Κατά τη φάση σύστασης, για κάθε χρήστη δημιουργείται ένα προφίλ υπολογίζοντας το μέσο όρο των κανονικοποιημένων embeddings όλων των καλλιτεχνών με τους οποίους έχει αλληλεπιδράσει. Η ομοιότητα μεταξύ του προφίλ χρήστη και κάθε καλλιτέχνη υπολογίζεται ως το εσωτερικό γινόμενο των κανονικοποιημένων διανυσμάτων. Έπειτα εφαρμόζεται η ποινή δη-

μοτικότητας, όπου το τελικό σκορ διαιρείται με την ύψωση της δημοτικότητας του καλλιτέχνη σε δύναμη ίση με την ποινή δημοτικότητας, ώστε να μειωθεί η κυριαρχία εξαιρετικά διάσημων καλλιτεχνών και να δοθεί έμφαση στη μουσική της «μακράς-ουράς». Οι προηγούμενες αλληλεπιδράσεις μηδενίζονται ώστε να μην επανεμφανιστούν στα αποτελέσματα. Τα αντικείμενα ταξινομούνται κατά φθίνουσα ομοιότητα και προτείνονται τα κορυφαία k.

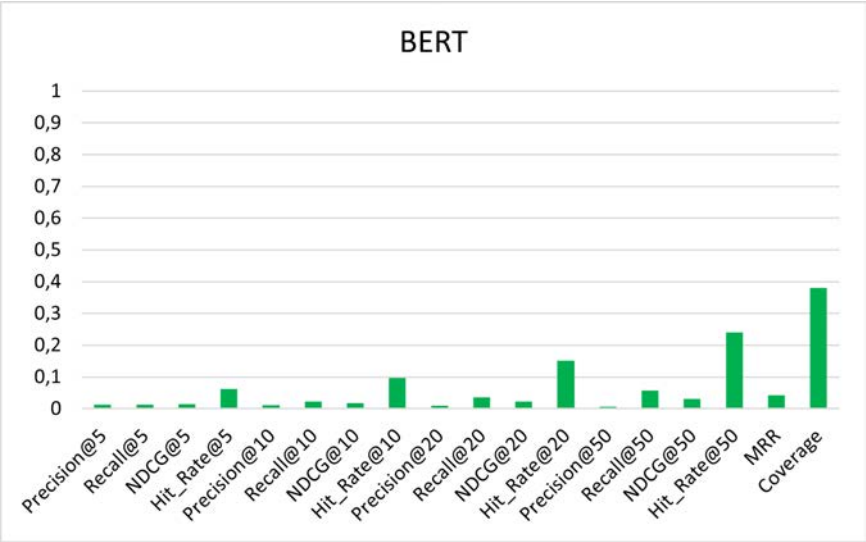
Αποτελέσματα

Στον Πίνακα 5.8 παρατίθενται τα αποτελέσματα του αλγορίθμου BERT σε όλες τις μετρικές για αυξανόμενο αριθμό συστάσεων. Επιπλέον, στο Σχήμα 5.6 εμφανίζονται τα ίδια αποτελέσματα και σε μορφή γραφήματος γραμμής για την οπτικοποίηση και ευκολότερη κατανόησή τους.

Μετρική	**@5**	**@10**	**@20**	**@50**
Precision	0,0134	0,0107	0,0091	0,0062
Recall	0,0135	0,0219	0,0355	0,0571
NDCG	0,0151	0,0175	0,0233	0,0309
Hit_Rate	0,0626	0,0973	0,1508	0,2396

Μετρική	**Τιμή**
MRR	0,0417
Coverage	0,3811

Πίνακας 5.8: Αποτελέσματα αξιολόγησης BERT



Σχήμα 5.6: Γραφική αναπαράσταση μετρικών BERT

Παρατηρήσεις

Ο αλγόριθμος BERT καταγράφει επιδόσεις που δεν φτάνουν τα επίπεδα των υπόλοιπων αλγορίθμων. Όπως και οι υπόλοιποι αλγόριθμοι, έτσι και ο BERT, όσο αυξάνεται ο αριθμός των συστάσεων, τόσο αυξάνονται και οι επιδόσεις του. Το Coverage περιορίζεται σε περίπου 38%, αναδεικνύοντας την τάση του αλγορίθμου να εστιάζει σε μικρότερο υποσύνολο καλλιτεχνών. Η χρήση σημασιολογικών αναπαραστάσεων μέσω του BERT επιτρέπει την καλύτερη κατανόηση του μουσικού περιεχομένου, παρόλο που, λόγω του περιορισμένου όγκου του συνόλου δεδομένων και της έλλειψης σημασιολογικής πληροφορίας, περιορίζεται η απόδοση του. Παράλληλα, η χαμηλότερη κάλυψη του δείχνει ότι το μοντέλο τείνει να επικεντρώνεται σε πιο «πυκνά» και αναγνωρίσιμα μοτίβα περιεχομένου, χάνοντας μέρος της εξερεύνησης του συνολικού χώρου.

5.6 Συγκριτική επίδοση των αλγορίθμων

Η παρούσα ενότητα εστιάζει στη συστηματική αξιολόγηση των αλγορίθμων που ενσωματώθηκαν στο σύστημα συστάσεων που αναπτύχθηκε, μέσω εφαρμογής τους σε δεδομένα χρήσης της μουσικής πλατφόρμας Last.fm. Πραγματοποιείται συγκριτική μελέτη των αλγορίθμων UserKNN, ItemKNN, SVD, NeuMF και BERT βάσει ποικίλων μετρικών αξιολόγησης (Precision, Recall, NDCG, Hit Rate, MRR, Coverage) σε πολλαπλά μεγέθη top N συστάσεων. Η ανάλυση αποσκοπεί όχι μόνο στην ανάδειξη των επιδόσεων κάθε μεθόδου, αλλά και

στην επισήμανση των ενδογενών περιορισμών και των δυνατοτήτων επέκτασης μελλοντικά, ιδίως υπό συνθήκες σπανιότητας δεδομένων και ανάγκης για υψηλή προσωποποίηση.

Μοντέλο	**Precision@5**	**Recall@5**	**NDCG@5**	**Hit_Rate@5**
UserKNN	0,0370	0,0264	0,0371	0,1631
ItemKNN	0,0201	0,0156	0,0204	0,0952
SVD	**0,0627**	**0,0406**	**0,0660**	**0,2674**
NeuMF	0,0614	0,0399	0,0635	0,2647
BERT	0,0134	0,0135	0,0151	0,0626

Πίνακας 5.9: Συγκριτικά αποτελέσματα στα @5

Στον Πίνακα 5.9 εμφανίζονται οι μετρικές κάθε αλγορίθμου για πλήθος συστάσεων N=5. Παρατηρείται ότι ο αλγόριθμος παραγοντοποίησης πινάκων (SVD) σημειώνει υψηλότερες τιμές και στις 4 μετρικές που αφορούν την ακρίβεια των συστάσεων με μικρή διαφορά από τον NeuMF. Οι υπόλοιπες προσεγγίσεις κυμαίνονται σε αρκετά χαμηλά επίπεδα. Αυτό συμβαίνει λόγω της δυνατότητας των πιο σύνθετων μοντέλων (SVD, NeuMF) να ανακαλύπτουν λανθάνουσες και μη-γραμμικές συσχετίσεις, ιδίως σε περιορισμένο αριθμό συστάσεων.

Μοντέλο	**Precision@10**	**Recall@10**	**NDCG@10**	**Hit_Rate@10**
UserKNN	0,0393	0,0544	0,0464	0,3000
ItemKNN	0,0182	0,0288	0,0241	0,1636
SVD	**0,0540**	**0,0706**	**0,0684**	**0,4102**
NeuMF	0,0539	0,0705	0,0671	0,4059
BERT	0,0107	0,0219	0,0175	0,0973

Πίνακας 5.10: Συγκριτικά αποτελέσματα στα @10

Παρόμοια αποτελέσματα παρατηρούνται και στις 10 συστάσεις όπως φαίνεται στον Πίνακα 5.10. Ο SVD καταφέρνει να ξεπεράσει όλους τους αλγόριθμους σε όλες τις μετρικές. Ωστόσο, αξίζει να σημειωθεί ότι η απόκλιση του από τον NeuMF σχετικά με τις μετρικές Precision και Recall αρχίζει να μειώνεται σημαντικά. Οι υπόλοιποι αλγόριθμοι παραμένουν σε χαμηλά επίπεδα σημειώνοντας μικρή αύξηση.

Μοντέλο	**Precision@20**	**Recall@20**	**NDCG@20**	**Hit_Rate@20**
UserKNN	0,0392	0,1079	0,0722	0,4968
ItemKNN	0,0182	0,0560	0,0368	0,2925
SVD	0,0452	0,1171	**0,0917**	**0,5610**
NeuMF	**0,0454**	**0,1187**	0,0910	0,5545
BERT	0,0091	0,0355	0,0233	0,1508

Πίνακας 5.11: Αποτελέσματα αξιολόγησης στα @20

Στον παραπάνω Πίνακα 5.11 παρατηρείται ότι όπως και στις 10 συστάσεις, έτσι και στις 20 οι αλγόριθμοι SVD και NeuMF παράγουν τις καλύτερες μετρικές και παραμένουν κοντά μεταξύ τους. Ακολουθεί με μικρή διαφορά ο UserKNN, ο οποίος όσο αυξάνονται οι συστάσεις τόσο καλύτερα φαίνεται να τα πηγαίνει στις μετρικές Precision, Recall και HitRate.

Μοντέλο	**Precision@50**	**Recall@50**	**NDCG@50**	**Hit_Rate@50**
UserKNN	**0,0371**	**0,2541**	0,1277	**0,7824**
ItemKNN	0,0170	0,1247	0,0625	0,5380
SVD	0,0330	0,2136	0,1292	0,7406
NeuMF	0,0342	0,2237	**0,1316**	0,7594
BERT	0,0062	0,0571	0,0309	0,2396

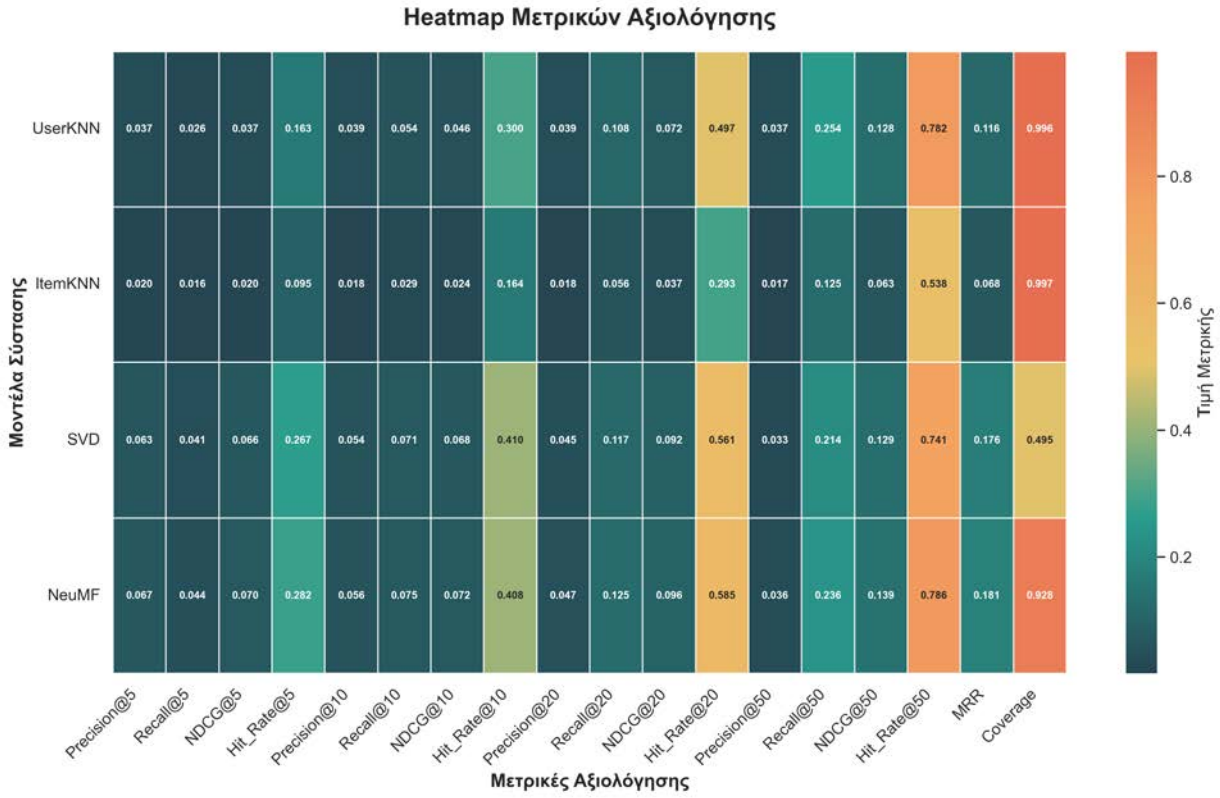
Πίνακας 5.12: Αποτελέσματα αξιολόγησης στα @50

Στον Πίνακα 5.12 ο αλγόριθμος NeuMF καταφέρνει πλέον να ξεπεράσει τον SVD και να αποδώσει καλύτερα σε όλες τις μετρικές. Ωστόσο, για πρώτη φορά ο UserKNN ξεπερνά τους υπόλοιπους αλγορίθμους, το οποίο είναι αναμενόμενο, καθώς το σύνολο δεδομένων το οποίο χρησιμοποιείται στην παρούσα εργασία δεν έχει επαρκή όγκο και είναι αρκετά αραιό, αποδυναμώνοντας τα πιο σύνθετα μοντέλα σε μεγαλύτερο αριθμό συστάσεων.

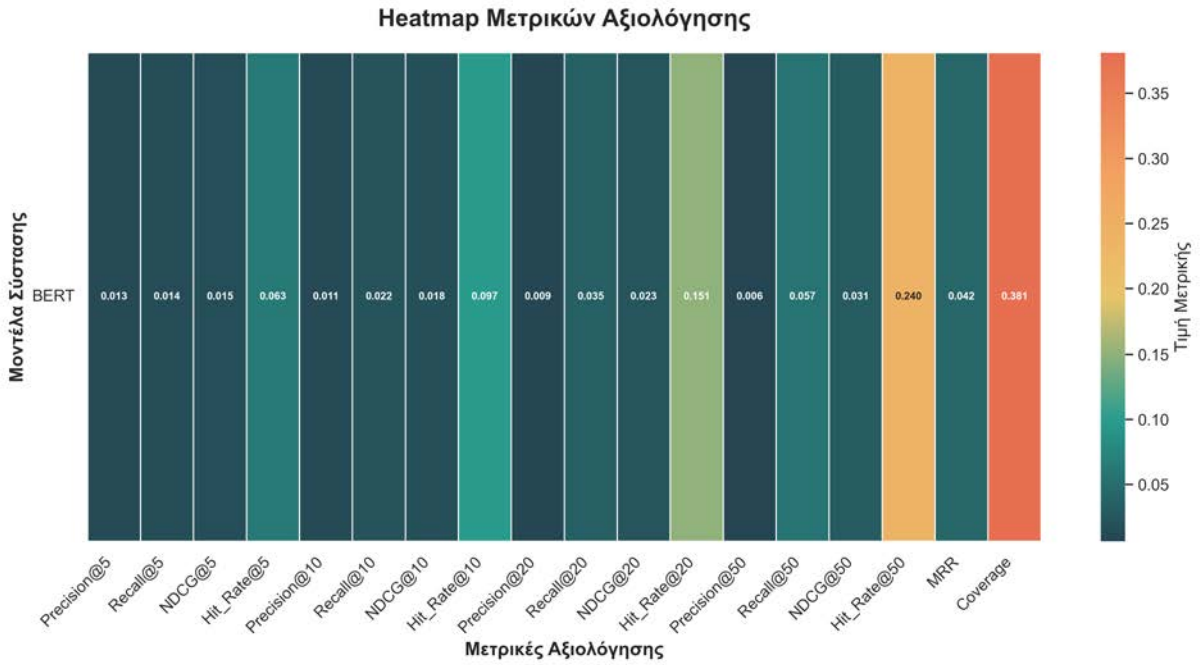
Μοντέλο	**MRR**	**Coverage**
UserKNN	0,1156	0,9965
ItemKNN	0,0680	**0,9968**
SVD	**0,1758**	0,4953
NeuMF	0,1703	0,9486
BERT	0,0417	0,3811

Πίνακας 5.13: Αποτελέσματα αξιολόγησης με MRR και Coverage

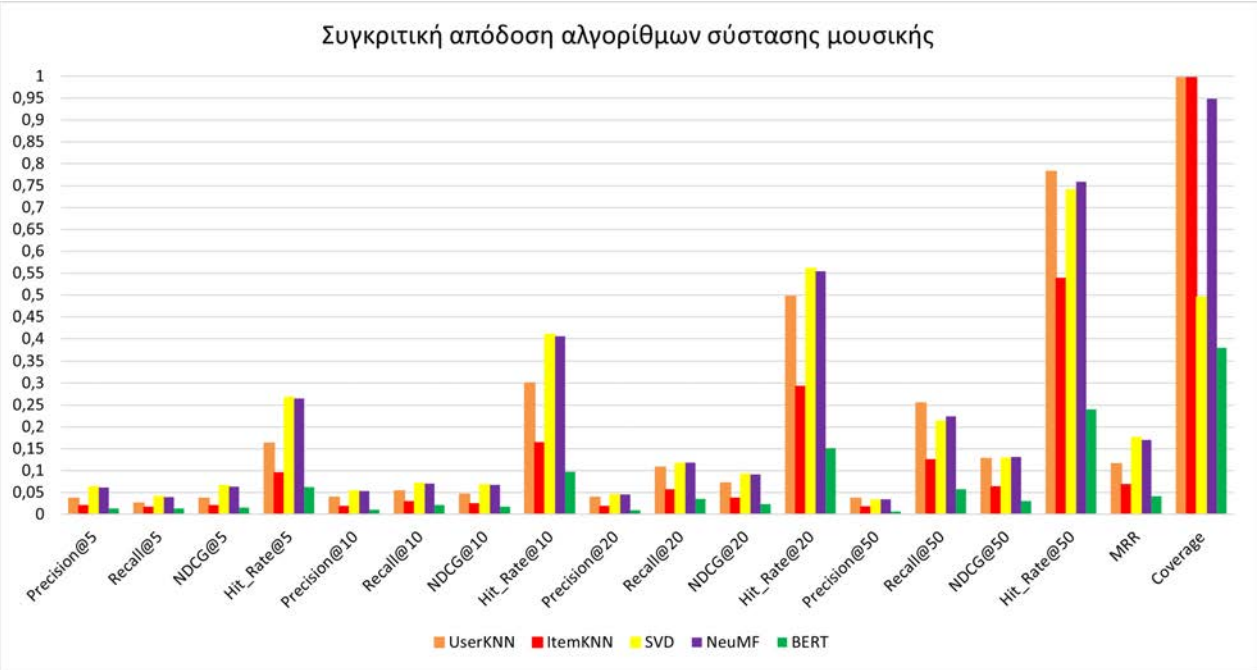
Στον τελευταίο Πίνακα 5.13 παρατηρείται για άλλη μια φορά η κυριαρχία του SVD και στη μετρική MRR αναδεικνύοντας την ικανότητά του να επιστρέφει σχετικές συστάσεις ψηλά στη λίστα, ακολουθούμενος στενά από τον NeuMF. Ωστόσο, η επίδοση του SVD πέφτει σημαντικά στη μετρική Coverage όπου οι ItemKNN, UserKNN και NeuMF αγγίζουν σχεδόν την τελειότητα. Παρακάτω παρουσιάζονται τα αποτελέσματα όλων των αλγορίθμων με 2 είδη γραφημάτων: Heatmap (Σχήματα 5.7 και 5.8) και Bar chart (Σχήμα 5.9), για την καλύτερη οπτικοποίηση των αποτελεσμάτων και την αμεσότερη σύγκρισή τους.



Σχήμα 5.7: Heatmap μετρικών UserKNN, ItemKNN, SVD, NeuMF



Σχήμα 5.8: Heatmap μετρικών BERT



Σχήμα 5.9: Bar Chart μετρικών UserKNN, ItemKNN, SVD, NeuMF, BERT

Κεφάλαιο 6

Συμπεράσματα και Μελλοντικές Επεκτάσεις

6.1 Συμπεράσματα

Η παρούσα εργασία ανέδειξε τις δυνατότητες και τους περιορισμούς διαφορετικών κατηγοριών αλγορίθμων μουσικών συστάσεων, μέσα από ένα συνεκτικό και ενιαίο πλαίσιο αξιολόγησης. Η ανάπτυξη ενός συστήματος μουσικών συστάσεων που αξιολογείται με διαφορετικές μετρικές, προσέφερε σημαντικά συμπεράσματα για την αποδοτικότητα και την καταλληλότητα κάθε μεθόδου, ιδιαίτερα σε συνθήκες όπου τα δεδομένα είναι περιορισμένα ή ανομοιογενή. Η αρχιτεκτονική του συστήματος αποδείχθηκε σταθερή και ευέλικτη, καθώς υποστήριξε την ενσωμάτωση τόσο κλασικών συνεργατικών μεθόδων όσο και πιο σύνθετων μοντέλων, όπως τα νευρωνικά δίκτυα και οι μέθοδοι βασισμένες στο περιεχόμενο. Η χρήση μιας κοινής προγραμματιστικής διεπαφής διευκόλυνε την εφαρμογή των διαφορετικών αλγορίθμων και επέτρεψε τη συγκριτική τους αξιολόγηση υπό ενιαίο πλαίσιο, χωρίς να περιορίζεται η εσωτερική τους λειτουργία. Από τα αποτελέσματα προέκυψε ότι οι παραδοσιακές μέθοδοι UserKNN και ItemKNN, αν και χρήσιμες ως σημείο αναφοράς, εμφάνισαν σχετικά περιορισμένες επιδόσεις στις μετρικές ακριβείας (Precision, Recall) λόγω των αραιών αλληλεπιδράσεων, με τον UserKNN να παρουσιάζει καλύτερη σχετικότητα (HitRate), ενώ και οι δύο φτάνουν σε πολύ υψηλά επίπεδα κάλυψης, πάνω από 99%. Ο αλγόριθμος SVD, που βασίζεται στη διάσπαση του πίνακα αλληλεπιδράσεων, παρουσίασε βελτιωμένα αποτελέσματα σε όλες τις μετρικές καθώς χρησιμοποιεί λανθάνουσες αναπαραστάσεις και αναγνωρίζει μοτίβα αυξάνοντας τη σχετικότητα των συστάσεων, πλήν αυτή της κάλυψης στην οποία

παρουσιάζει πτώση λόγω του ότι επικεντρώνεται στη σύσταση πιο δημοφιλών αντικειμένων. Ο NeuMF, που συνδυάζει παραγοντοποίηση με νευρωνικά δίκτυα, σημείωσε εξαιρετικά αποτελέσματα σε όλες τις μετρικές παρόμοια με αυτά του SVD, αποτέλεσμα αναμενόμενο καθώς ο NeuMF εμπεριέχει και τον SVD. Η επίδοση του τον καθιστά τον πιο σταθερό αλγόριθμο συνολικά, καθώς προσφέρει υψηλή ακρίβεια και εκτενή διερεύνηση. Επιπλέον, στον Sentence-BERT, που δεν αξιοποιεί συνεργατικές αλληλεπιδράσεις αλλά μόνο σημασιολογικές περιγραφές των καλλιτεχνών, παρατηρήθηκε πτώση των μετρικών, γεγονός που κατά πάσα πιθανότητα οφείλεται στη μη αξιοποίηση δεδομένων συνεργατικών αλληλεπιδράσεων καθώς και στο περιορισμένο σε σημασιολογικά χαρακτηριστικά σύνολο δεδομένων. Η χρήση τεχνικών όπως η ποινή δημοτικότητας και το κεντράρισμα των αξιολογήσεων βοήθησε στη σταθερότητα των προβλέψεων και στη βελτίωση της ποικιλομορφίας των συστάσεων, χωρίς να μειώνεται σημαντικά η ακρίβεια. Τέλος, τα αποτελέσματα δείχνουν ότι δεν υπάρχει απόλυτα βέλτιστη λύση για κάθε περίπτωση, και πώς η επιλογή του κατάλληλου αλγορίθμου εξαρτάται από τις ανάγκες του εκάστοτε σεναρίου — όπως η κάλυψη, η ακρίβεια, ή η αντιμετώπιση της αραιότητας.

6.2 Μελλοντικές Επεκτάσεις

Η υλοποίηση που παρουσιάστηκε στη παρούσα διπλωματική αποτελεί ένα πλήρως λειτουργικό πρόγραμμα συγκριτικής αξιολόγησης για συστήματα μουσικών συστάσεων. Ωστόσο, αρκετές σημαντικές επεκτάσεις θα μπορούσαν να αξιοποιηθούν σε μελλοντική εργασία, τόσο ως προς τη βελτίωση της ακρίβειας όσο και τη λειτουργική επεκτασιμότητα και την ερευνητική εμβάθυνση. Πρωταρχικά, μία σημαντική επέκταση αφορά την ενσωμάτωση χρονικής πληροφορίας. Στην παρούσα εργασία, οι αλληλεπιδράσεις μεταξύ χρήστη και αντικειμένων θεωρούνται στατικές, χωρίς να λαμβάνεται υπόψη η χρονική τους διάσταση. Η αξιοποίηση τεχνικών όπως οι Recurrent Neural Networks (RNNs) ή οι μηχανισμοί αυτοπροσοχής (self-attention) θα μπορούσε να επιτρέψει την ανάλυση μεταβαλλόμενων μουσικών προτιμήσεων, οδηγώντας σε πιο δυναμικά και προσωποποιημένα μοντέλα σύστασης. Ένα άλλο ερευνητικό μονοπάτι σχετίζεται με την βελτιστοποίηση του Sentence-BERT μέσα από καλιμπράρισμα στο domain του Last.fm για την παραγωγή καλύτερων μετρικών τόσο στην ακρίβεια όσο και στη σχετικότητα. Η χρήση ελεγχόμενης μάθησης (π.χ. Multiple Negatives Ranking Loss σε θετικά ζεύγη καλλιτεχνών που συνυπάρχουν σε playlists) δύναται να βελτιώσει δραστικά την

ποιότητα των ενσωματώσεων, ενισχύοντας τη σχετικότητα των συστάσεων σε σημασιολογικό επίπεδο. Επιπλέον, μια ενδιαφέρουσα προέκταση αφορά τη χρήση εναλλακτικών τύπων περιεχομενικών αναπαραστάσεων, όπως βιογραφικά καλλιτεχνών, αποσπάσματα στίχων ή πολυτροπικά embeddings από οπτικά εξώφυλλα δίσκων και clips. Η ενσωμάτωσή τους σε πολυεπίπεδες νευρωνικές αρχιτεκτονικές μπορεί να εμπλουτίσει τη γνωστική βάση του συστήματος, επιτρέποντας την καλύτερη κατανόηση της μουσικής ταυτότητας των αντικειμένων. Σε επίπεδο συνεργατικών μοντέλων, θα μπορούσε να διερευνηθεί η εφαρμογή τεχνικών συστολής (shrinkage), με στόχο τη βελτίωση της σταθερότητας, ειδικά σε χρήστες ή αντικείμενα με χαμηλή συχνότητα. Η χρήση shrinkage στις μετρικές ομοιότητας – όπως η σταθμισμένη Pearson – μπορεί να εξομαλύνει ακραίες τιμές και να προσφέρει πιο ρεαλιστικές τιμές ομοιότητας σε αλληλεπιδράσεις με περιορισμένη πληροφορία. Μια ακόμη δυνητική κατεύθυνση είναι η διασύνδεση του συστήματος συστάσεων με πραγματικό user interface (GUI ή web), που θα μετέτρεπε το έργο από αμιγώς ερευνητικό σε λειτουργικά εφαρμόσιμο, ανοίγοντας δρόμους για εφαρμογή σε πραγματικές συνθήκες χρήσης. Τέλος, προτείνεται η χρήση γραφημάτων νευρωνικών δικτύων (graph neural networks - GNNs) για την αναπαράσταση του κοινωνικού δικτύου (φιλικές συνδέσεις) ως γράφου. Μέσω GNN-based encoders, μπορεί να μοντελοποιηθεί πληρέστερα η τοπολογία και η επιρροή των κοινωνικών κύκλων, ενσωματώνοντας ταυτόχρονα τόσο διασυνδέσεις όσο και περιεχομενικές πληροφορίες κόμβων. Συνολικά, το υφιστάμενο σύστημα παρέχει μια σταθερή βάση, πάνω στην οποία μπορούν να αναπτυχθούν ισχυρότερες, επεκτάσιμες και πιο εξελιγμένες εκδοχές.

Βιβλιογραφία

[1] G. Research, "Hetrec 2011 datasets (movielens, imdb/rotten tomatoes, last.fm, delicious)," https://grouplens.org/datasets/hetrec-2011/, 2011.

[2] G. Linden, B. Smith, and J. York, "Amazon.com recommendations: item-to-item collaborative filtering," *IEEE Internet Computing*, vol. 7, no. 1, pp. 76–80, Jan. 2003.

[3] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, Aug. 2009.

[4] Y. Deldjoo, M. Schedl, and P. Knees, "Content-driven music recommendation: Evolution, state of the art, and challenges," *arXiv preprint arXiv:2107.11803*, Jan. 2023.

[5] A. Bellogín, I. Cantador, and P. Castells, "A study of heterogeneity in recommendations for a social music service," in *Proc. 1st Int. Workshop on Information Heterogeneity and Fusion in Recommender Systems*. Barcelona, Spain: ACM, Sep. 2010, pp. 1–8.

[6] K. Dinnissen and C. Bauer, "Fairness in music recommender systems: A stakeholder-centered mini review," *Frontiers in Big Data*, vol. 5, p. 913608, Jul. 2022.

[7] M. Schedl, "Deep learning in music recommendation systems," *Frontiers in Applied Mathematics and Statistics*, vol. 5, p. 44, Aug. 2019.

[8] B. Amiri, N. Shahverdi, A. Haddadi, and Y. Ghahremani, "Beyond the trends: Evolution and future directions in music recommender systems research," *IEEE Access*, vol. 12, pp. 51 500–51 522, 2024.

[9] A. Majumdar, "Music recommendations based on implicit feedback and social circles: The last fm data set," [Unpublished technical report].

[10] S. Berkovsky, I. Cantador, and D. Tikk, *Collaborative Recommendations: Algorithms, Practical Challenges and Applications*. World Scientific, 2018.

[11] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural collaborative filtering," *arXiv preprint arXiv:1708.05031*, Aug. 2017.

[12] Z. Zhao *et al.*, "Recommender systems in the era of large language models (llms)," *IEEE Transactions on Knowledge and Data Engineering*, vol. 36, no. 11, pp. 6889–6907, Nov. 2024.

[13] A. Said, "On explaining recommendations with large language models: a review," *Frontiers in Big Data*, vol. 7, p. 1505284, Jan. 2025.

[14] S. Yun and Y. Lim, "User experience with llm-powered conversational recommendation systems: A case of music recommendation," in *Proc. 2025 CHI Conf. on Human Factors in Computing Systems*, Apr. 2025, pp. 1–15.

[15] S. Oramas, O. Nieto, M. Sordo, and X. Serra, "A deep multimodal approach for cold-start music recommendation," in *Proc. 2nd Workshop on Deep Learning for Recommender Systems*, Aug. 2017, pp. 32–37.

[16] D. Liang, R. G. Krishnan, M. D. Hoffman, and T. Jebara, "Variational autoencoders for collaborative filtering," *arXiv preprint arXiv:1802.05814*, Feb. 2018.

[17] A. Roussas, "Music recommendation systems comparison study," https://github.com/ChrisForis/Music_Recommendation_Systems_Comparison_Study, 2025.