



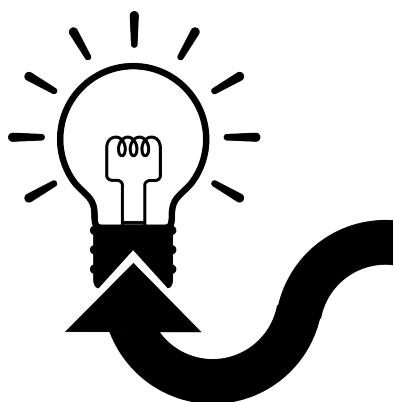
ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ "ΕΠΙΣΤΗΜΗ ΔΕΔΟΜΕΝΩΝ & ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ"

**Υλοποίηση Εξατομικευμένου  
Συστήματος Συστάσεων Βασιζόμενο σε  
Αρνητικά Δείγματα με Στόχο τη Ταχεία  
Προσαρμογή σε Πραγματικό Χρόνο**

**ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

του

**ΝΙΚΟΛΑΚΟΠΟΥΛΟΥ Κ. ΑΝΑΣΤΑΣΙΟΥ**



**Επιβλέπων:** Γιώργος Στάμου  
Αναπληρωτής Καθηγητής

Αθήνα, Μάρτιος 2021





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ "ΕΠΙΣΤΗΜΗ ΔΕΔΟΜΕΝΩΝ & ΜΗΧΑΝΙΚΗ ΜΑ-  
ΘΗΣΗ"

**Υλοποίηση Εξατομικευμένου Συστήματος  
Συστάσεων Βασιζόμενο σε Αρνητικά  
Δείγματα με Στόχο τη Ταχεία Προσαρμογή  
σε Πραγματικό Χρόνο**

**ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

του

**ΝΙΚΟΛΑΚΟΠΟΥΛΟΥ Κ. ΑΝΑΣΤΑΣΙΟΥ**

**Επιβλέπων:** Γιώργος Στάμου  
Αναπληρωτής Καθηγητής

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 26 Μαρτίου 2021.

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

..... ..... .....  
Γιώργος Στάμου Ανδρέας-Γεώργιος Σταφυλοπάτης Στέφανος Κόλλιας  
Αναπληρωτής Καθηγητής Καθηγητής Καθηγητής

Αθήνα, Μάρτιος 2021





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ "ΕΠΙΣΤΗΜΗ ΔΕΔΟΜΕΝΩΝ & ΜΗΧΑΝΙΚΗ ΜΑ-  
ΘΗΣΗ"

Copyright © – All rights reserved. Με την επιφύλαξη παντός δικαιώματος.  
Αναστάσιος Νικολακόπουλος, 2021.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα.

Το περιεχόμενο αυτής της εργασίας δεν απηχεί απαραίτητα τις απόψεις του Τμήματος, του Επιβλέποντα, ή της επιτροπής που την ενέκρινε.



# Περίληψη

---

Από την αρχαιότητα, οι άνθρωποι βρίσκουν συντροφιά και απόλαυση στο άκουσμα μουσικής και όμορφης μελωδίας. Με τη πάροδο των ετών, η μουσική έχει εξελιχθεί σε τεράσιο βαθμό. Μάλιστα, οι τελευταίες δεκαετίες βρήκαν το παγκόσμιο μουσικό στερέωμα να αλλάζει για πάντα, καθώς πέρασε στη ψηφιακή εποχή. Σήμερα, η πρόσθιαση σε μουσική είναι ευκολότερη από ποτέ, μέσω των έξυπνων ηλεκτρονικών συσκευών. Ένας χρήστης μπορεί να δημιουργήσει μια προσωπική του λίστα αναπαραγωγής, η οποία θα αποτελείται από αμέτρητα μουσικά κομμάτια διαφόρων ειδών, απολαμβάνοντας τις μελωδίες της επιλογής του. Μέσω της διαδικασίας τυχαίας αναπαραγωγής, ο χρήστης μπορεί να "χαθεί" στις σκέψεις, τα συναισθήματα και τη φαντασία του. Τι γίνεται όμως στη περίπτωση που, κάποια χρονική στιγμή, ο ίδιος χρήστης επιθυμήσει να ακούσει μόνο κομμάτια κάποιου συγκεκριμένου είδους από τη λίστα αναπαραγωγής του; Η παρούσα διπλωματική εργασία θα επιχειρήσει να δώσει απάντηση σε αυτό ακριβώς το ερώτημα. Θα εξετάσει αν είναι εφικτό να γίνουν αντιληπτές οι μουσικές προτιμήσεις ενός ατόμου μια χρονική στιγμή, μέσω των κομματίων που προσπερνάει (*skips*) κατά τη διάρκεια ακρόασης του. Θα γίνει χρήση και αξιοποίηση συστημάτων συστάσεων και μοντέλων μηχανικής μάθησης, ώστε να απαντηθεί το αν γίνεται μια λίστα αναπαραγωγής να "παιζει" μόνο τα μουσικά κομμάτια που θέλει ο χρήστης κάποια συγκεκριμένη χρονική στιγμή, σε πραγματικό χρόνο, με μοναδική τροφοδότηση τα αρνητικά του δείγματα, δηλαδή τα τραγούδια που ο ίδιος προσπερνάει.

## Λέξεις Κλειδιά

Μουσική, Ψηφιακή Μουσική, Λίστες Αναπαραγωγής, Υπηρεσίες Ροής Μουσικής, Τυχαία Αναπαραγωγή, Μηχανική Μάθηση, Συστήματα Συστάσεων, Αρνητικά Δείγματα



# **Abstract**

---

From antiquity, humans have used to find company and pleasure by listening to music and beautiful melodies. Over the years, music has evolved big time. Indeed, these past decades have changed the global industry forever, by passing to the digital age of music. Nowadays, access to music had never been easier, especially through smart electronic devices. A user can create a personal playlist, containing countless music tracks from all kinds of genres, and enjoy the melodies he/she selected. Furthermore, through the playlist shuffle feature, the user can relax and "travel" through his/her thoughts, feelings or fantasies. However, what if, at some particular moment, this user wished to listen to tracks of specific genre/s from his/her playlist? This Master Thesis will attempt to provide an answer for the aforementioned question. It will examine whether it is possible (or not) to detect the music preferences of a user, at a particular moment in time, through the music tracks that are skipped during his/her listening session. Recommendation System models & Machine Learning algorithms shall be used and tested, in order to figure out if a playlist can indeed "play" just the music tracks that a user wishes to listen to in a specific moment and in real-time, using only the negative samples (that is, the tracks being skipped).

## **Keywords**

Music, Digital Music, Playlists, Music Streaming, Shuffling, Machine Learning, Recommendation Systems, Negative Sampling



*...στους γονείς μου*



# Ευχαριστίες

---

Θα ήθελα καταρχάς να ευχαριστήσω τον καθηγητή κ. Γιώργο Στάμου για την επίβλεψη της παρούσας διπλωματικής εργασίας, καθώς και για την ευκαιρία που μου έδωσε να ερευνήσω & να αναπτύξω ένα δικό μου θέμα. Επίσης ευχαριστώ ιδιαίτερα τον Υποψήφιο Διδάκτορα Γιώργο Φιλανδριανό για τη βοήθεια, τη καθοδήγηση του και την εξαιρετική συνεργασία που είχαμε. Τέλος, θα ήθελα φυσικά να ευχαριστήσω τους γονείς μου για τη συμπαράσταση που μου προσέφεραν στο διάθα των ετών.

Αθήνα, Μάρτιος 2021

*Αναστάσιος Νικολακόπουλος*



# Περιεχόμενα

---

<b>Περίληψη</b>	<b>1</b>
<b>Abstract</b>	<b>3</b>
<b>Ευχαριστίες</b>	<b>7</b>
<b>Πρόλογος</b>	<b>17</b>
<b>1 Εισαγωγή</b>	<b>19</b>
1.1 Αντικείμενο της διπλωματικής . . . . .	28
1.2 Οργάνωση του τόμου . . . . .	29
<b>I Θεωρητικό Μέρος</b>	<b>31</b>
<b>2 Shuffle Play, Συστήματα Συστάσεων &amp; Μηχανική Μάθηση</b>	<b>33</b>
2.1 Shuffle Play . . . . .	33
2.1.1 Γενικός Ορισμός . . . . .	33
2.2 Συστήματα Συστάσεων . . . . .	35
2.2.1 Ορισμός . . . . .	35
2.2.2 Είδη & Κατηγορίες . . . . .	41
2.3 Μηχανική Μάθηση . . . . .	46
2.3.1 Ορισμός . . . . .	46
2.3.2 Προσεγγίσεις . . . . .	49
2.3.3 Αρνητικά Δείγματα ως Δεδομένα Δειγματοληψίας . . . . .	52
<b>3 Περιγραφή θέματος</b>	<b>53</b>
3.1 Η Κεντρική Ιδέα . . . . .	53
3.2 Σχετικές Υλοποιήσεις . . . . .	59
3.2.1 Υλοποιήσεις Συστημάτων Συστάσεων . . . . .	59
3.2.2 Υλοποιήσεις Μηχανικής Μάθησης . . . . .	64

<b>II Πρακτικό Μέρος</b>	<b>69</b>
<b>4 Ανάλυση Υλοποίησης</b>	<b>71</b>
4.1 Το Σύνολο Δεδομένων . . . . .	71
4.1.1 Γνώστα Σύνολα Διαθέσιμα στο Διαδίκτυο . . . . .	71
4.1.2 Spotify Dataset . . . . .	73
4.2 Ταξινομητές Μηχανικής Μάθησης . . . . .	82
4.3 Content-Based Σύστημα Συστάσεων . . . . .	87
<b>5 Υλοποίηση</b>	<b>93</b>
5.1 Κώδικας Δημιουργίας Συνόλου Δεδομένων . . . . .	93
5.2 Κώδικας Δημιουργίας Convolutional Neural Network . . . . .	97
5.3 Κώδικας Δημιουργίας Συστήματος Συστάσεων . . . . .	99
<b>6 Αποτελέσματα</b>	<b>105</b>
6.1 Μηχανική Μάθηση . . . . .	105
6.2 Σύστημα Συστάσεων . . . . .	111
<b>III Επίλογος</b>	<b>117</b>
<b>7 Συμπεράσματα &amp; Επεκτάσεις</b>	<b>119</b>
7.1 Συμπεράσματα . . . . .	119
7.2 Μελλοντικές Επεκτάσεις . . . . .	120
<b>Παραρτήματα</b>	<b>123</b>
<b>Βιβλιογραφία</b>	<b>130</b>

# Κατάλογος Σχημάτων

---

1.1	Με τη πάροδο των ετών, τα είδη μουσικής έχουν πολλαπασιαστεί, δημιουργώντας νέες βασικές κατηγορίες και πολλά παρακλάδια αυτών [1]. . . . .	20
1.2	Από τα Walkmans που έφεραν την επανάσταση, στα iPods που τη συνέχισαν, συνδυάζοντας τη μετάβαση στον ψηφιακό κόσμο της μουσικής. 21	
1.3	Όλη η σειρά μοντέλων iPod Shuffle της Apple που κυκλοφόρησαν από το 2005 έως και το 2017. Να σημειωθεί πως η Apple θεωρείται μια από τις εταιρείες που εδραιώσαν τη σύγχρονη ψηφιακή μορφή της μουσικής στο χώρο. . . . .	23
1.4	Μερικές από τις δημοφιλέστερες εταιρείες υπηρεσιών ροής μουσικής παγκοσμίως. Η βιομηχανία του streaming έχει γιγαντωθεί τα τελευταία χρόνια, ενώ αξίζει να σημειωθεί πως έχει επεκταθεί / εδραιωθεί και στον χώρο του κινηματογράφου. . . . .	25
1.5	Τα συστήματα συστάσεων είναι έτσι κατασκευασμένα, ώστε να προβλέπουν τις προτιμήσεις του χρήστη με βάση υπάρχουσες πληροφορίες γύρω από εκείνον ή γύρω από σχετικούς με αυτόν χρήστες [2]. . . . .	27
2.1	Το καθολικά "αναγνωρισμένο" σύμβολο της τυχαίας αναπαραγωγής. Η πλειοψηφία των αναπαραγών πολυμέσων χρησιμοποιούν το συγκεκριμένο σύμβολο των διασταυρωμένων βελών για να υποδείξουν στο χρήστη τη δυνατότητα του shuffle play. . . . .	34
2.2	Οι "commercial recommenders" παίζουν πλέον βασικό ρόλο στη προσπάθεια των εταιρειών να προσελκύσουν πελάτες, παρέχοντας τους όσο το δυνατόν καλύτερες προτάσεις προϊόντων που να αντιστοιχούν στις προτιμήσεις τους [3]. . . . .	37
2.3	Οι εσερχόμενες πληροφίες στα συστήματα συστάσεων, όπως και οι κατηγορίες αυτών, είναι αχανείς. Θα μπορούσαν να παρουσιαστούν σε έναν γράφο εισερχόμενων στοιχείων γνώσης διαφορετικών ειδών. Μέσα από όλα αυτά, οι recommenders καλούνται να εξάγουν χρήσιμα συμπεράσματα και ασφαλείς προβλέψεις για χρήστες [4]. . . . .	41

2.4	Σήμερα, τρεις εκ των προαναφερθεισών κατηγοριών recommenders χαίρουν υιοθέτησης και ενδιαφέροντος. Τα content-based συστήματα, η δημοφιλής collaborative filtering τεχνική και ο φουτουριστικός hybrid συνδυασμός υλοποιήσεων είναι οι κατηγορίες που έχουν ξεχωρίσει [5]. . . . .	45
2.5	Αποτύπωση της δημοφιλούς άποψης πως ο κλάδος της τεχνητής νοημοσύνης εμπεριέχει εκείνους της μηχανικής μάθησης και της νεότερης τάσης στο χώρο, το deep learning [6]. . . . .	47
2.6	Αποτύπωση μιας αισθητά διαφορετικής προσέγγισης, στην οποία η μηχανική μάθηση μοιράζεται αρκετά κοινά στοιχεία με τον τομέα της τεχνητής νοημοσύνης, αλλά η μία δεν εμπεριέχεται στην άλλη. Σε κάθε περίπτωση, το deep learning αποτελεί στοιχείο εμβάθυνσης της μηχανικής μάθησης [7]. . . . .	47
2.7	Απεικόνιση της πολυπλοκότητας που χαρακτηρίζει τα νέα συστήματα εκπαίδευσης υπολογιστών για απόκτηση ευφυΐας, αλλά και τη διασύνδεση μεταξύ τους. Η μηχανική μάθηση, η τεχνητή νοημοσύνη και το "νεότερο" deep learning έχουν άμεση σχέση στην επιτυχημένη λειτουργία ενός τεχνητού νευρώνα [8]. . . . .	48
2.8	Οι κυριότερες προσεγγίσεις της μηχανικής μάθησης (εξαιρώντας τη semi-supervised περίπτωση), μαζί με τα είδη χρήσεων και τις εφαρμογές τους [9]. . . . .	51
3.1	Η επιλογή κατηγοριοποίησης των μουσικών κομματιών ανά είδος, όπως φαίνεται στα αριστερά της εφαρμογής Apple Music, σε μια συσκευή Apple iPad. Όμως, δεν υπάρχει η ίδια δυνατότητα και για τα κομμάτια εντός των λιστών αναπαραγωγής [10]. . . . .	54
3.2	Μία αρχιτεκτονική απεικόνιση του αρχικού θεωρητικού συστήματος που πραγματεύεται και ερευνά η διπλωματική εργασία. Δοκιμή μοντέλων μηχανικής μάθησης και content-based συστημάτων συστάσεων για εντοπισμό μουσικών προτιμήσεων χρήστη. Θα εξεταστούν μεμονομένα και αν τα αποτελέσματα είναι αξιόλογα, θα γίνει δοκιμή συνδυασμού αυτών. . . . .	57
3.3	Η αρχιτεκτονική της προαναφερθείσας δημοσίευσης, που παρουσιάζει ένα προηγμένο σύστημα συστάσεων, το οποίο μάλιστα λειτουργεί σε πραγματικό χρόνο [11]. Αξιοποιεί πληροφορίες από ψηφιακές κοινότητες (κοινωνικών δικτύων) χρηστών ώστε εν συνεχείᾳ να τους προτείνει σχετικά αγαθά & προϊόντα. . . . .	60

3.4	Τι μπορεί να βρίσκεται στο επίκεντρο των μουσικών συναισθημάτων & προτιμήσεων ενός χρήστη κάθε φορά ; Αυτό είναι μια πραγματική πρόκληση που καλούνται να αντιμετωπίσουν τα μελλοντικά συστήματα συστάσεων [12]. Η ιδέα της παρούσας διπλωματικής εργασίας πειραματίζεται με το προαναφερθέν ερώτημα. . . . .	62
3.5	Η χρήση μοντέλων μηχανικής μάθησης με σκοπό τη πρόβλεψη μουσικών ειδών και κατηγοριοποίηση κομματιών με βάση αυτά, είναι ενα ιδιαίτερα ενδιαφέρον σκέλος έρευνας [13]. Έχει ουσιαστικό νόημα η χρήση του machine learning ή πρόκειται για ένα πρόβλημα που χρήζει άλλης λύσης ; . . . . .	65
4.1	Μια λίστα με τα δημοφιλέστερα μουσικά σύνολα δεδομένων, διαθέσιμα μέσω διαδικτύου [14]. Το dataset του Spotify δεν περιλαμβάνεται στη λίστα, πιθανότατα λόγω του γεγονότος πως το δημιουργεί κανείς μόνος του, δυναμικά, μέσω του παρεχόμενου API. . . . .	72
4.2	Ένα παράδειγμα χρήσης του Spotify Web API μέσω της επίσημης κονσόλας που παρέχει για δοκιμές η υπηρεσία. Στο συγκεκριμένο λειτουργεί για λήψη πληροφοριών ενός συγκεκριμένου καλλιτέχνη. Στο πεδίο id γίνεται εισαγωγή του μοναδικού χαρακτηριστικού κωδικού ενός καλλιτέχνη, ενώ στο αμέσως επόμενο πεδίο γίνεται εισαγωγή ενός στοιχείου αυθεντικοποίησης για τη λειτουργία του συστήματος. Δεξιά φαίνεται το request, καθώς και ένα κομμάτι της απάντησης. Το id που χρησιμοποείται αφορά τον καλλιτέχνη Pitbull. . . . .	74
4.3	Danceability, κατανομή ανά μουσικά είδη. . . . .	77
4.4	Energy, κατανομή ανά μουσικά είδη. . . . .	77
4.5	Loudness, κατανομή ανά μουσικά είδη. . . . .	78
4.6	Speechiness, κατανομή ανά μουσικά είδη. . . . .	78
4.7	Acousticness, κατανομή ανά μουσικά είδη. . . . .	78
4.8	Instrumentalness, κατανομή ανά μουσικά είδη. . . . .	79
4.9	Live ness, κατανομή ανά μουσικά είδη. . . . .	79
4.10	Valence, κατανομή ανά μουσικά είδη. . . . .	79
4.11	Tempo, κατανομή ανά μουσικά είδη. . . . .	80
4.12	Η διαδικασία απόκτησης των μουσικών ειδών για ένα κομμάτι. Αφού πρώτα συλλεχθούν τα μελωδικά χαρακτηριστικά του, έπειτα γίνεται αναζήτηση πληροφοριών στο Web API με βάση τον καλλιτέχνη του, και εν συνεχεία συλλέγονται τα είδη που τον χαρακτηρίζουν. . . . .	81
4.13	Το γνωστό και ιδιαίτερα δημοφιλές Θεώρημα Bayes. . . . .	83
4.14	Η διαφορά μεταξύ της γραμμικής (αριστερά) και λογιστικής (δεξιά) παλινδρόμησης, ως προς τη κατανομή των τιμών που τείνουν να πάρουν σε ένα δυαδικό μοντέλο [15]. . . . .	83

4.15	Απεικόνιση της λειτουργίας του αλγορίθμου kNN. Ανάλογα με τον καθορισμό του $k$ , το μοντέλο θα ταξινομήσει το δείγμα (πράσινη κουκίδα) στα κόκκινα τρίγωνα ή τα μπλέ τετράγωνα [16]. . . . .	84
4.16	Η αρχιτεκτονική και ο τρόπος λειτουργίας του Random Forest Classifier. . . . .	84
4.17	Τα δέντρα αποφάσεων και η λειτουργία μέσω της οποίας προχωρούν στην ταξινόμηση των χαρακτηριστικών [17]. . . . .	85
4.18	Απεικόνιση του τρόπου με τον οποίο δρα ο XGBoost, ενισχύοντας την απόδοση των κλασικών gradient boosting μοντέλων και βελτιώνοντας το τελικό αποτέλεσμα [18]. . . . .	85
4.19	Η αποτύπωση της διάταξης των επιπέδων στο δημιουργηθέν 2d convolutional νευρωνικό δίκτυο για το μουσικό σύνολο δεδομένων μας. .	86
4.20	Η μέθοδος εμπλουτισμού των δεδομένων και η δημιουργία του νέου συνόλου, μέσω χρήσης του endpoint εντοπισμού παρόμοιων καλλιτεχνών από το Spotify Web API. Γίνεται αναζήτηση σχετικών καλλιτεχνών και των μουσικών ειδών τους. Τελικά, ένα νέο σύνολο προκύπτει με τα εμπλουτισμένα είδη και ομαδοποιημένους τους καλλιτέχνες ανά είδος. Έτσι, ένα μουσικό κομμάτι στο αρχικό σύνολο, το οποίο θα αντιστοιχεί σε καποια είδη, θα έχει επιπρόσθετη αντιστοίχηση σε πληροφορίες (καλλιτέχνες), μέσω αυτών των ειδών. . . . .	89
4.21	Η "κεντρική" μαθηματική εξίσωση μέσω της οποίας εξάγει τις μετρικές του ο αλγόριθμος του cosine similarity. . . . .	90
4.22	Οι μαθηματικές εξισώσεις που οδηγούν στον τελικό υπολογισμό της μετρικής tf-idf. . . . .	91
4.23	Η τελική και ολοκληρωμένη αρχιτεκτονική του συστήματος που παρουσιάζει η παρούσα διπλωματική εργασία. Συνδυασμός των επιλογών του Spotify Web API, των δυνατοτήτων της μηχανικής μάθησης, της λειτουργίας του content-based recommendation και φυσικά της αγαπημένης λειτουργίας του shuffling. Όλα αυτά, σε προσωποποιημένο επίπεδο, για έναν συγκεκριμένο χρήστη, μια συγκεκριμένη χρονική στιγμή και σε πραγματικό χρονο. . . . .	92
6.1	Πίνακας συσχετίσεων μελωδικών χαρακτηριστικών του κεντρικού dataset. . . . .	105
6.2	Ιστόγραμμα σπουδαιότητας μελωδικών χαρακτηριστικών του κεντρικού dataset, μέσω του κατηγοριοποιητή XGBoost. . . . .	106
6.3	Ο πλήρης πίνακας αποτελεσμάτων ακρίβειας των υλοποιηθέντων μοντέλων μηχανικής μάθησης, με χρήση του πλήρους dataset και εναλλαγή στον αριθμό των μουσικών ειδών (labels) προς κατηγοριοποίηση. . . . .	107
6.4	Διάγραμμα Συνολικών Αναπαραγωγών "μικρού" συνόλου. . . . .	111
6.5	Διάγραμμα Συνολικών Αναπαραγωγών "μεσαίου" συνόλου. . . . .	112

6.6	Διάγραμμα Συνολικών Skips “μικρού” συνόλου.	113
6.7	Διάγραμμα Συνολικών Skips “μεσαίου” συνόλου.	113
6.8	Διάγραμμα Play / Skip Ratio Ανά Ώρα (PSh) “μικρού” συνόλου.	114
6.9	Διάγραμμα Play / Skip Ratio Ανά Ώρα (PSh) “μεσαίου” συνόλου.	114
6.10	Διάγραμμα Πέντε Συνεχόμενων Αναπαραγωγών “μικρού” συνόλου.	115
6.11	Διάγραμμα Πέντε Συνεχόμενων Αναπαραγωγών “μεσαίου” συνόλου.	115



## Πρόλογος

---

Η παρούσα διπλωματική εργασία ερευνήθηκε, αναπτύχθηκε και συντάχθηκε στην Αθήνα, το χειμερινό εξάμηνο του ακαδημαϊκού έτους 2020 - 2021. Αξιοποιεί γνώσεις, μεθόδους και τεχνικές που αποκομήθηκαν από τον κύκλο σπουδών του μεταπτυχιακού προγράμματος "Επιστήμη Δεδομένων & Μηχανική Μάθηση". Παρά τη δυσκολία της περιόδου, λόγω της πανδημίας από την COVID-19 που έπληξε την ανθρωπότητα το 2020, η διπλωματική εργασία ολοκληρώθηκε εντός ενός εξαμήνου, χωρίς να λάβει πρόσθετη χρονική παράταση.



## Κεφάλαιο 1

### Εισαγωγή

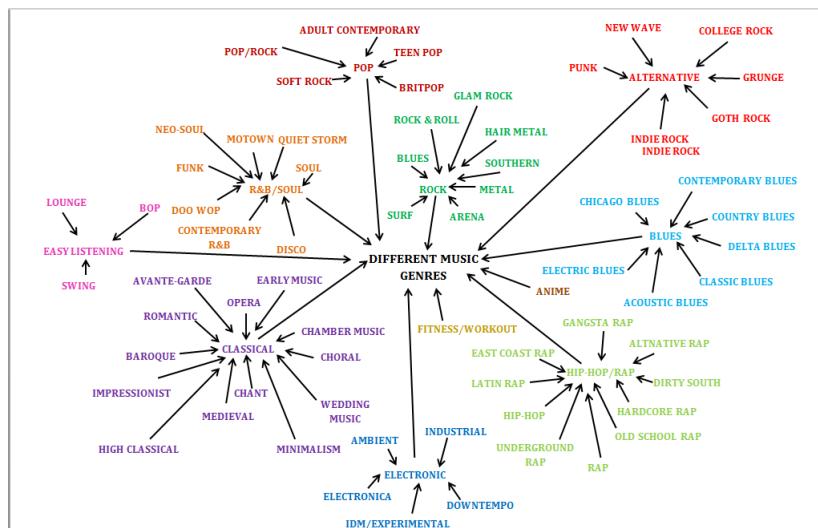
---

**Η** μουσική συνοδεύει τον άνθρωπο εδώ και χιλιάδες χρόνια. Από την αρχαιότητα, οι λαοί χρησιμοποιούσαν τη μελωδία ως έναν τρόπο ψυχαγωγίας, έκφρασης, αλλά και επικοινωνίας. Μάλιστα, χαρακτηριστική είναι και η πληθώρα μουσικών οργάνων που χρονολογούνται χιλιάδες χρόνια πριν από τη κοινή χρονολογία (προ Χριστού). Έχουν ανακαλυφθεί συσκευές που παραπέμπουν σε πρώτες μορφές μουσικών οργάνων και οι οποίες χρονολογούνται (τουλάχιστον) από το 60000 π.Χ [19]. Είναι προφανές πως ο άνθρωπος, εκ φύσεως, έχει τη τάση της δημιουργίας μελωδίας, ενός συνδυασμού ήχων που συνδυάζονται αρμονικά και παράγουν ένα όμορφο αποτέλεσμα, το οποίο γίνεται αντιληπτό φυσικά μέσω της ακοής. Άλλωστε, η χώρα στην οποία ζούμε, η Ελλάδα, έχει μακρά ιστορία στη μουσική. Οι "αοιδοί" και οι "ραψωδοί" της αρχαίας ελληνικής ιστορίας, πρόσωπα που απήγγηλαν έπη & ιστορίες μπροστά σε κοινό, συνόδευαν το αφήγημα τους με μελωδίες κιθάρας, φόρμιγγας ή άλλου οργάνου της εποχής. Η πρόοδος των θεωρητικών επιστημών και δή της ιστορίας έχει αποδείξει περίτρανα πως η μουσική ήταν αναπόσπαστο κομμάτι των αρχαίων πολιτισμών σε κάθε γωνιά του πλανήτη.

Καθώς τα χρόνια περνούσαν, οι άνθρωποι ανακάλυπταν όλο και περισσότερες μεθόδους με τις οποίες μπορούσαν να παράξουν μελωδίες και να συνθέσουν μουσική. Αυτό φυσικά ώθησε στη δημιουργία διαφορετικών ειδών μελωδίας. Αργά αλλά σταθερά, η μουσική εξελίχθηκε μαζί με τις τάσεις, τις προτιμήσεις και τα γούστα των ατόμων που ασχολούνταν ενεργά με εκείνη. Κάθε πολιτισμός ανέπτυσσε μια δικιά του φιλοσοφία γύρω από τους όμορφους τόνους και τις μελωδίες των δικών του μουσικών οργάνων, παρουσιάζοντας ένα ή και περισσότερα μουσικά είδη. Καθώς οι αιώνες περνούσαν και η μεταφορά των ανθρώπων σε διάφορα μέρη της υφηλίου έγινε πιο ευρεία & εύκολη (εξερευνήσεις), η διάδοση των πολλών και διαφορετικών ειδών μουσικής από πολιτισμό σε πολιτισμό έμοιαζε αναπόφεκτη. Πλέον, γινόταν αντιληπτό το αχανές φάσμα και η ποικιλία των ειδών μουσικής που ήδη υπήρχε, αλλά και μου έμελλε να ανακαλυφθεί. Άλλωστε, δεν είναι υπερβολή να πιστεύει κανείς σήμερα πως οι μελωδίες μπορούν να συνδυάζονται επ' αόριστον, δημιουργώντας ατελείωτα παράγοντα είδη. Πόσο μάλλον για τους μουσικούς των προηγούμενων αιώνων, που

με τη σειρά τους λάβαιναν τα σκύπτρα των προκατόχων τους και έφερναν πρόοδο στις μελωδίες. Η Ευρώπη έπαιξε καθοριστικό ρόλο στην εδραίωση αρκετών ειδών, όπως η κλασική μουσική, μέσα από θρυλικούς συνθέτες (Mozart, Bach, Beethoven, Strauss, Brahms και πολλοί άλλοι), η πλειοψηφία των οποίων εντοπίζεται στη κεντρική/θόρεια έκταση της Γηραιάς Ήπειρου.

Καθώς οδηγούμαστε στο σήμερα, η μουσική έχει εξελιχθεί σε τεράστιο βαθμό. Σε αυτό φυσικά συνέβαλε και η ανακάλυψη / κατασκευή αρκετών νέων οργάνων (πχ. Συνθεσάιζερ - Ηλεκτρικό Αρμόνιο), αλλά και η "αναβάθμιση" / διαφοροποίηση ήδη υπαρχόντων (πχ. ηλεκτρική κιθάρα), προσδίδοντας τους νέες δυνατότητες. Φυσικά, μεγάλο ρόλο σε αυτό έπαιξε και η άνθιση της τεχνολογίας, κυρίως από τον 20ο αιώνα και έπειτα. Η επανάσταση που έφεραν οι ηλεκτρονικοί υπολογιστές στη μουσική, ενδεχομένως να μην έχει προηγούμενο. Ουσιαστικά, τώρα πια είναι εφικτή η παραγωγή οποιασδήποτε μελωδίας, από οποιοδήποτε μουσικό όργανο, μέσω ενός H/Y. Αν και κάτι τέτοιο δεν αντικαθιστά τους αυθεντικούς ήχους των οργάνων, μαρτυρά πως σήμερα η μουσική μπορεί να αναπαραχθεί με πολλούς τρόπους που προηγουμένως δεν ήταν εφικτοί. Επίσης, οι ίδιοι οι υπολογιστές έχουν οδηγήσει στη δημιουργία νέων κατηγοριών από είδη, των οποίων τα μελωδικά στοιχεία παράγονται κυρίως (ή αποκλειστικά) μέσα από έναν H/Y. Πλέον, έχει δημιουργηθεί ένας αχανής γράφος αποτελούμενος από αμέτρητα διαφορετικά μουσικά είδη, πολλά εκ των οποίων είναι μικρότερες υπο-κατηγορίες "βασικών" ειδών. Αυτό δημιουργεί μια τεράστια γκάμα επιλογών για τους ακροατές, που είναι σε θέση να επιλέξουν είδη τα οποία "αγγίζουν" τα συναισθήματα και συμφωνούν με τις ιδιαίτερες προτιμήσεις τους. Σήμερα, περισσότερο από ποτέ, οι άνθρωποι μπορούν να βρουν ένα είδος μουσικής που απευθύνεται σχεδόν αποκλειστικά σε εκείνους. Ποτέ άλλοτε δεν ήταν τόσο μεγάλη η συλλογή μουσικών ειδών!



Σχήμα 1.1: Με τη πάροδο των ετών, τα είδη μουσικής έχουν πολλαπασιαστεί, δημιουργώντας νέες βασικές κατηγορίες και πολλά παρακλάδια αυτών [1].

Φυσικά, η τεράστια ποικιλία μουσικών ειδών έχει οδηγήσει τη σύγχρονη επιστήμη στην ανάγκη δημιουργίας συστημάτων που θα είναι σε θέση να αντιληφθούν τις "ακουστικές" προτιμήσεις ενός ανθρώπου και εν συνεχεία να μπορούν να του προτείνουν σχετικές μελωδίες / τραγούδια. Γιατί; Ας κάνουμε ενα βήμα πίσω και ας θυμηθούμε την ιστορία της δυνατότητας ατομικής ακρόασης μουσικής. Οι άνθρωποι μπορούσαν να απολαύσουν τη μουσική ως ομάδες από την αρχαιότητα. Όμως, αυτό που θα ήθελε ο οποιοσδήποτε αγαπάει τη μουσική, είναι να βρίσκεται σε θέση να απολαύσει τις μελωδίες της προτίμησης του σε ατομικό επίπεδο. Η επινόηση των ατομικών και φορητών audio players, δηλαδή συσκευών ήχου που μπορούσαν να μεταφέρονται από το χρήστη με ευκολία και ταυτόχρονα να του δίνουν τη δυνατότητα αναπαραγωγής ήχου σε ατομικό επίπεδο, έγινε στα μέσα του προηγούμενου αιώνα (δεκαετία 1950). Όμως, η πραγματική επανάσταση έγινε μέσω των Walkman συσκευών της Sony, οι οποίες πρωτο-εμφανίστηκαν το 1979 [20]. Τα Walkman κατέστησαν εφικτή τη φορητή και ατομική αναπαραγωγή κασετών, δίσκων CD (για πάνω από 10 χρόνια, τα Walkman CD players της Sony είχαν την ονομασία Discman) και αργότερα ηλεκτρονικών ηχητικών κομματιών. Ο χρήστης φορούσε ενα ζευγάρι ακουστικά και εκείνα μετέδιδαν τον ήχο απευθείας στα αυτιά του μέσω καλωδίου συνδεδεμένου με τη Walkman συσκευή. Με τη πάροδο του χρόνου και την δημοτικότητα που απέκτησαν τα Walkman προϊόντα της Ιαπωνικής εταιρείας, περισσότεροι κατασκευαστές κυκλοφόρησαν φορητά audio players. Συνεχιστής των Walkmans στη θέση των πιο επιτυχημένων προϊόντων ήταν τα iPods της Apple, τα οποία πλέον έκαναν αποκλειστικά αναπαραγωγή ψηφιακών μουσικών κομματιών. Η Sony (και φυσικά μαζί της πολλοί άλλοι κατασκευαστές) συνέχισαν στη κατασκευή ψηφιακών media players, όμως οι σχεδιαστικά ιδιαίτερες και ποιοτικές προτάσεις της Apple υπερίσχυσαν του ανταγωνισμού. Όπως είναι λογικό (ενώ αναφέρθηκε και προηγουμένως), η ψηφιοποίηση του ήχου άλλαξε για πάντα τη μουσική βιομηχανία. Πολλοί έχουν χαρακτηρίσει αυτή τη μετάβαση (ψηφιοποίηση) σε μάστιγα για τη μουσική, άλλοι πάλι τη θεωρούν ένα μεγάλο βήμα μπροστά. Το σίγουρο είναι πως κάθε μεγάλη αλλαγή έχει τόσο τα θετικά της, όσο και τα αρνητικά της.



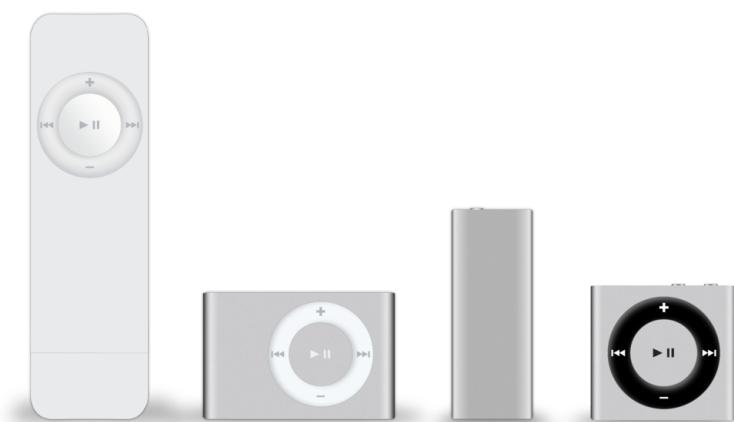
**Σχήμα 1.2:** Από τα Walkmans που έφεραν την επανάσταση, στα iPods που τη συνέχισαν, συνδυάζοντας τη μετάβαση στον ψηφιακό κόσμο της μουσικής.

Όμως, είναι γεγονός πως τα θετικά χαρακτηριστικά είναι εκείνα που υπερι-

σχύουν. Εμείς θα εξετάσουμε μόλις μερικά εξ’ αυτών, συγκεκριμένα τα κυριότερα, ώστε τελικά να οδηγηθούμε στην κατανόηση και απάντηση του ερωτήματος που αφέθηκε αναπάντητο παραπάνω και αφορά την ανάγκη δημιουργίας συστημάτων που να αντιλαμβάνονται τις προτιμήσεις των χρηστών. Ποια είναι αυτά τα κύρια πλεονεκτήματα της ψηφιακής εποχής στη μουσική; Αυτή η επανάσταση στη βιομηχανία, έδωσε στους ακροατές τη δυνατότητα συλλογής μουσικών κομματιών μεμονομένα, από πολλούς διαφορετικούς καλλιτέχνες & άλμπουμ, χωρίς την ανάγκη προμήθειας ολόκληρου του άλμπουμ ενός καλλιτέχνη (όπως συνέβαινε με τις κασέτες ή τα CD). Τα προηγούμενα χρόνια, αν κάποιος ήθελε να αποκτήσει ένα συγκεκριμένο τραγούδι ενός καλλιτέχνη, έπρεπε (κατά κύριο λόγο) να προμηθευτεί τη κασέτα ή το δίσκο CD που το περιείχε, κάτι που σημαίνει πως καλούνταν να ακούσει και άλλα μουσικά κομμάτια που είχαν εγγραφεί στο μέσο. Η ψηφιακή επανάσταση όμως έδωσε τη δυνατότητα επιλογής συγκεκριμένων μουσικών κομματιών (αφού πλέον ένας χρήστης μπορούσε να αποκτήσει πρόσθιαση μεμονομένα στο κομμάτι που ήθελε, χωρίς να πρέπει να αγοράσει ολόκληρο το ηλεκτρονικό άλμπουμ) και δημιουργίας μιας προσωπικής συλλογής εκείνου, αποτελούμενης από (ενδεχομένως) πολλά κομμάτια από διαφορετικά είδη. Για άλλη μια φορά, να σημειώσουμε πως πριν από μερικές δεκαετίες, για να έχει κάποιος συλλογή κομματιών, έπρεπε ταυτόχρονα να κάνει και συλλογή κασετών, δίσκων CD ή φυσικά βινυλίου. Ήταν ουσιαστικά αδύνατο να συγκεντρώσει τα αγαπημένα του κομμάτια σε ένα μέσο, ή τουλάχιστον αρκετά δύσκολο, αφού μπορούσε μεν να γίνει εγγραφή νέων κασετών / CDs που περιείχαν τα κομμάτια επιλογής ενός ατόμου, αλλά δεν ήταν τόσο απλή διαδικασία. Ανάλογα με την όρεξη, τη διάθεση και τη συναισθηματική κατάσταση, ο χρήστης τοποθετούσε το εκάστοτε μέσο εγγραφής στο σύστημα αναπαραγωγής και απολάμβανε τη μουσική του. Αν ήθελε να αλλάξει είδος μουσικής ή καλλιτέχνη, έπρεπε να αφαιρέσει το μέσο αυτό και να τοποθετήσει κάποιο άλλο στο σύστημα αναπαραγωγής. Με λίγα λόγια, έκανε ένα δικό του “ανακάτεμα” (shuffling), αν και αξίζει να σημειωθεί πως μερικά CD players είχαν τη δυνατότητα τυχαίας αναπαραγωγής κομματιών του CD ή των CD, αφού κάποια είχαν υποδοχές για περισσότερους από έναν οπτικούς δίσκους. Αντιλαμβανόμαστε όμως πως η δημιουργία προσωπικής συλλογής από επιλεγμένα κομμάτια και η τυχαία αναπαραγωγή αυτών ανάλογα με τη διάθεση ήταν μια διαδικασία που απέκτησε ουσιαστικό νόημα & αξία με τη ψηφιακή επανάσταση...!

Στον “νέο” κόσμο της ψηφιακής μουσικής, ο χρήστης μπορούσε να δημιουργήσει όσες λίστες αναπαραγωγής (playlists) ήθελε, γεμάτες από επιλεγμένα κομμάτια (ή φυσικά ολόκληρα άλμπουμ) που συνήθως με τη σειρά τους αντιπροσώπευαν πολλά διαφορετικά μουσικά είδη. Η ατομική ακρόαση / απόλαυση της μουσικής απέκτησε έναν πολύ πιο “προσωπικό” χαρακτήρα, αφού πλέον οι χρήστες μπορούσαν να δημιουργήσουν συλλογές “φτιαγμένες και ραμμένες” για εκείνους. Φυσικά, η εναλλαγή των κομματιών ήταν πια μια απλούστατη διαδικασία, αφού λάθαινε χώρα με το πάτημα ενός πλήκτρου. Μάλιστα, σύντομα μετά τη κυκλοφορία ψηφιακών audio

players, προστέθηκε η δυνατότητα τυχαίας αναπαρωγής (shuffling ή ανακάτεμα των μουσικών κομματιών), κάτι που έκανε τη χαλάρωση και την απόλαυση της μουσικής ακόμα πιο εύκολη & άμεση. Ένας χρήστης μπορούσε να απολαμβάνει αμέτρητα μουσικά κομμάτια και να εναλάσσεται τυχαία μεταξύ, μουσικών ειδών με ευκολία που όμοια της δεν είχε εφευρεθεί μέχρι τώρα... Απλώς επέλεγε τη δυνατότητα τυχαίας αναπαραγωγής και κάθε νέο κομμάτι στη σειρά αυτής (της αναπαραγωγής) ερχόταν τυχαία από τη playlist του. Για να γίνει αυτό πριν μερικές δεκαετίες, έπρεπε να αλλάζει συνεχώς δίσκο ή κασέτα! Αξίζει να τονιστεί πως οι λίστες μουσικών κομματιών που ήταν πλέον σε θέση να δημιουργήσει ο χρήστης μπορούσαν να αποτελούνται από εκατοντάδες ή και χιλιάδες κομμάτια. Αυτό γιατί άλλο ένα μεγάλο πλεονέκτημα της ψηφιοποίησης είναι η δυνατότητα μείωσης του μεγέθους των μουσικών κομματιών σε μερικά Megabytes, ώστε να είναι εφικτή η δημιουργία μεγάλων συλλογών ακόμα και σε συσκευές με χαμηλή χωρητικότητα. Η δυνατότητα του shuffling ήταν τόσο σημαντική για το κοινό, που πολλοί κατασκευαστές το χρησιμοποίησαν ως βασικό στοιχείο στα πλαίσια της προώθησης των audio players τους. Η Apple κυκλοφόρησε για μια περίοδο ετών (2005 - 2017) τη σειρά iPod Shuffle, μια γκάμα μικρών audio players που εξ' αρχής αναπαρήγαγαν τυχαία τα μουσικά κομμάτια που ήταν αποθηκευμένα εντός τους. Δεν έφεραν ψηφιακή οθόνη ώστε ο χρήστης να βλέπει ποιο κομμάτι έπαιζε την εκάστοτε στιγμή. Απλώς έκαναν τυχαία αναπαραγωγή. Αυτή η απλή λειτουργία, παρά τις βασικές ελλείψεις των iPod Shuffle σε τεχνικά χαρακτηριστικά συγκριτικά με μοντέλα του ανταγωνισμού, τα έκαναν ιδιαίτερα αγαπητά στο καταναλωτικό κοινό. Αντιλαμβάνεται κανείς πόσο σημαντική ήταν η δυνατότητα της τυχαίας αναπαραγωγής, στα πλαίσια των πλεονεκτημάτων της ψηφιακής επανάστασης στη μουσική βιομηχανία.

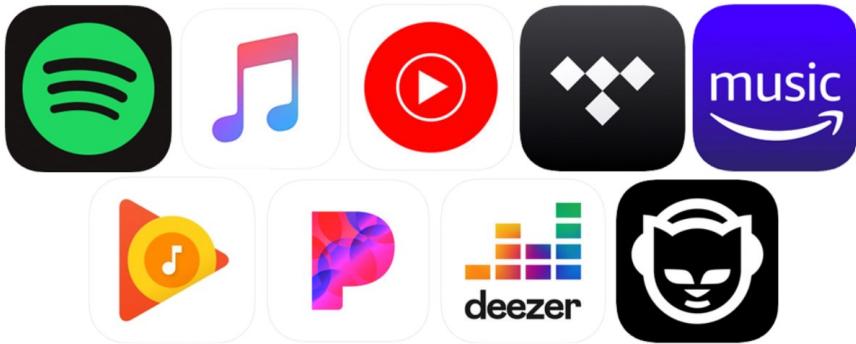


**Σχήμα 1.3:** Όλη η σειρά μοντέλων iPod Shuffle της Apple που κυκλοφόρησαν από το 2005 έως και το 2017. Να σημειωθεί πως η Apple θεωρείται μια από τις εταιρείες που εδραιώσαν τη σύγχρονη ψηφιακή μορφή της μουσικής στο χώρο.

Σήμερα, τα προαναφερθέντα πλεονεκτήματα (και φυσικά πολλά παραπάνω) της ψηφιακής εποχής της μουσικής είναι εμφανή καθημερινά. Η συντριπτική πλειο-

ψηφία του πληθυσμού χρησιμοποιεί το κινητό του τηλέφωνο, τον ηλεκτρονικό υπολογιστή, τη ταμπλέτα ή το έξυπνο ρολόι του ως μέσα αναπαραγωγής μουσικής. Λίγοι είναι εκείνοι που προμηθεύονται CD για να ακούσουν μουσική, ενώ μια μικρή μερίδα ανθρώπων, κυρίως συλλεκτών, εξακολουθεί να έχει μια ιδιαίτερη αγάπη ως προς την "αναλογική" αναπαραγωγή ήχου, μέσω δίσκων βινυλίου. Κατά γενική ομολογία όμως, η ψηφιακή μουσική είναι το "standard" που θα παραμείνει για τις μελλοντικές γενιές. Το ίδιο συμβαίνει και σε επαγγελματικό επίπεδο. Τα άτομα που ασχολούνται με τη μουσική βιομηχανία επαγγελματικά, έχουν υιοθετήσει τις ψηφιακές δυνατότητες ως μέρος της δουλειάς τους. Χαρακτηριστικά παραδείγματα είναι οι DJ, τους οποίους μάλιστα μπορούμε να πούμε πως αφέλησε ιδιαίτερα η δυνατότητα ταχείας εναλλαγής μουσικών κομματιών, αλλά και η δυνατότητα σύνθεσης / δημιουργίας αυτών αποκλειστικά μέσα από έναν υπολογιστή. Οι ηχολήπτες, οι υπεύθυνοι επεξεργασίας, οι ίδιοι οι καλλιτέχνες, ολόκληρες οι δισκογραφικές εταιρείες ουσιαστικά όλοι οι κλάδοι στη μουσική βιομηχανία έχουν "εισέλθει" στον ψηφιακό κόσμο.

Μια άλλη κατηγορία που ουσιαστικά δημιουργήθηκε μέσα από τη νέα εποχή της μουσικής, είναι εκείνη των υπηρεσιών ροής μουσικής (streaming). Πρόκειται για εταιρείες που παρέχουν τη δυνατότητα πρόσθασης σε μια τεράστια γκάμα μουσικών τραγουδιών, εκατομμυρίων σε αριθμό, από καλλιτέχνες ολόκληρου του κόσμου. Οι χρήστες έχουν τη δυνατότητα αναπαραγωγής οποιουδήποτε μουσικού κομματιού ή άλμπουμ επιθυμούν, τη δημιουργία προσωπικών τους λιστών αναπαραγωγής, τη προώθηση συνδέσμων που παραπέμπουν σε κομμάτια, καλλιτέχνες, λίστες ή άλμπουμ σε φίλους μέσω κοινωνικών δικτύων, αλλά και πολλές άλλες επιλογές. Για να έχουν πρόσθαση στην υπηρεσία της εταιρείας, οι χρήστες καλούνται να πληρώσουν μια μηνιαία, πολύμηνη ή ετήσια συνδρομή. Μπορεί κανείς να δοκιμάσει τις υπηρεσίες δωρεάν, όμως οι δυνατότητες που παρέχονται σε εκείνον / εκείνη είναι περιορισμένες. Το μεγάλο "ατού" των υπηρεσιών ροής μουσικής - streaming σε σχέση με την αναπαραγωγή κομματιών μέσω του Youtube (δηλαδή της απλής πλοήγησης στον ιστότοπο του, διότι το Youtube παρέχει δική του streaming υπηρεσία) ή με τη πειρατική πρόσθαση σε αυτά, είναι πως οι υπηρεσίες παρέχουν το περιεχόμενο απολύτως νόμιμα, σε συνεργασία με τους καλλιτέχνες, ενώ επίσης το διαθέτουν σε υψηλή ποιότητα και με τη δυνατότητα λήψης των κομματιών, ώστε ο χρήστης να μπορεί να απολαύσει τη μουσική του ακόμα και όταν δεν είναι συνδεδεμένος στο διαδίκτυο. Ακόμα, οι υπηρεσίες διαθέτουν αμέτρητες έτοιμες λίστες αναπαραγωγής, που ο χρήστης μπορεί να διαλέξει ανάλογα με τη διάθεση και την όρεξη του κάθε φορά. Δεν είναι τυχαίο πως οι υπηρεσίες streaming έχουν κερδίσει την εμπιστοσύνη εκατοντάδων εκατομμυρίων καταναλωτών παγκοσμίως. Δημοφιλείς εταιρείες streaming στον χώρο της μουσικής βιομηχανίας είναι, μεταξύ άλλων, οι Spotify, Apple Music, Tidal, Youtube Music, Soundcloud, Pandora, Deezer, Amazon Music, Google Play Music και Napster.



**Σχήμα 1.4:** Μερικές από τις δημοφιλέστερες εταιρείες υπηρεσιών ροής μουσικής παγκοσμίως. Η βιομηχανία του streaming έχει γιγαντωθεί τα τελευταία χρόνια, ενώ αξίζει να σημειωθεί πως έχει επεκταδεί / εδραιωθεί και στον χώρο του κινηματογράφου.

Ήρθε λοιπόν η ώρα να επιστρέψουμε στο αρχικό μας ζήτημα, που αφήσαμε αναπάντητο και μη ορθά ορισμένο για να κάνουμε μια μικρή ιστορική αναδρομή στους φορητούς audio players αλλά και για να δούμε κάποια σημαντικά πλεονεκτήματα της ψηφιακής εποχής στη μουσική. Όπως προαναφέρθηκε λοιπόν, η εξέλιξη της μουσικής βιομηχανίας, ο ερχομός νέων και αμέτρητων δυνατοτήτων, μαζί με την σημαντική αύξηση των μουσικών ειδών έφεραν (μεταξύ άλλων) μια "πρωτοφανή" ανάγκη. Ποια είναι αυτή; Πρόκειται για την ανάγκη δημιουργίας συστημάτων που θα αντιλαμβάνονται τις προτιμήσεις των ακροατών, ώστε να προσαρμόζουν το ηχητικό περιεχόμενο τους. Όμως, γιατί είναι κάτι τέτοιο σημαντικό; Η σημερινή κοινωνία εξελίσσεται ταχύτατα. Κινητήρια δύναμη μιας κοινωνίας είναι οι άνθρωποι. Οπότε, καταλήγουμε στο λογικό συμπέρασμα πως οι ίδιοι οι άνθρωποι είναι εκείνοι που εξελίσσονται ταχύτατα σήμερα.

Η διαρκής ανάπτυξη στην επιστήμη και τη τεχνολογία, η γιγάντωση των εννοιών της παγκοσμιοποίησης και του καταναλωτισμού, καθώς και τα ασταμάτητα σημαντικά γεγονότα που λαβαίνουν χώρα καθημερινά στην υφήλιο, μαζί με τις συνέπειες που αυτά φέρνουν σε τοπικό / ιπειρωτικό / παγκόσμιο επίπεδο έχουν οδηγήσει τη σύγχρονη ανθρωπότητα στην τάση της διαρκούς εξέλιξης. Κατά μια έννοια, οι άνθρωποι είναι "νομάδες" της καθημερινότητας τους. Αυτό φυσικά έχει άμεσο αντίκτυπο στη ψυχαγωγία και τα μέσα χαλάρωσης τους. Συγκεκριμενοποιώντας τη σκέψη αυτή στο κομμάτι της μουσικής ακρόασης, ένα άτομο σήμερα ενδέχεται να επιθυμεί την ακρόαση εντελώς διαφορετικών μουσικών ειδών την ίδια χρονική περίοδο εντός της ημέρας. Για παράδειγμα, μπορεί αρχικά να αρέσκεται στην ακρόαση μερικών κομματιών Rock, αλλά στη πορεία να θελήσει να συνεχίσει με Jazz μουσική. Επίσης, ενδέχεται να επιθυμεί συγκεκριμένα κομμάτια ανεξαρτήτως είδους, εναλλάσσοντας τις μουσικές κατηγορίες σε κάθε νέο τραγούδι. Επιπρόσθετα, ενδέχεται να επιθυμεί μόνο ένα συγκεκριμένο είδος από τη λίστα αναπαραγωγής που έχει δημιουργήσει, η οποία περιέχει εκατοντάδες ή χιλιάδες τραγούδια από διαφορετικά είδη. Σε κάθε περίπτωση, οι άνθρωποι ακουν μουσική και ενδέχεται να επιθυμούν ταχείς εναλλαγές στις ακουστικές προτιμήσεις και τα είδη που θα επιλέξουν στο άμεσο μέλλον.

Με κάποιον τρόπο, θα πρέπει να αντιμετωπιστεί η ανάγκη αυτή.

Η σύγχρονη επιστήμη έχει αφιερώσει ολόκληρους τομείς, ενώ μάλιστα έχει δημιουργήσει αποκλειστικούς κλάδους που εξετάζουν τα μουσικά μοτίβα και τις προτιμήσεις των ακροατών. Σε συνδυασμό με τεχνολογίες ανάλυσης δεδομένων & μηχανικής μάθησης, δανειζόμενες από το πεδίο των Η/Υ, οι επιστήμονες επιχειρούν εδώ και μερικά χρόνια να δώσουν στοχευμένες λύσεις στο "ζήτημα" της αντίληψης και πρόβλεψης των "θέλω" του χρήστη. Φυσικά, η επιστημονική αυτή μελέτη δεν αφορά μόνο το χώρο της μουσικής, αλλά και άλλες κατηγορίες, όπως η αγορά, ο κινηματογράφος, η πλοήγηση στο διαδίκτυο κλπ. Άλλωστε, η γνώση των μοτίβων & προτιμήσεων ενός χρήστη, αποτελεί πολύτιμο αγαθό για τις σημερινές εταιρείες, έχοντας ως στόχο το κέρδος και προωθώντας σε εκείνον τα αγαθά / προϊόντα / υπηρεσίες που θεωρούν πως τον αφορούν, με βάση τα στοιχεία που έχουν στη κατοχή τους. Βέβαια, η συγκεκριμένη ανάλυση αποτελεί ένα αρκετά ξεχωριστό κεφάλαιο που δεν θα συζητηθεί και αναλυθεί στη παρούσα διπλωματική εργασία. Εμείς θα εστιάσουμε στις δυνατότητες πρόβλεψης των προτιμήσεων ενός ατόμου στα πλαίσια της μουσικής, και θα επιχειρήσουμε να εξετάσουμε μια νέα μέθοδο που αφορά τη ταχεία πρόβλεψη των μουσικών προτιμήσεων του χρήστη σε μια συγκεκριμένη χρονική στιγμή. Τελικά όμως, ποια είναι η ονομασία των συστημάτων που προβλέπουν τις προτιμήσεις των χρηστών σε μια κατηγορία / κλάδο:

Ο λόγος για τα Συστήματα Συστάσεων, γνωστά ευρέως ως "Recommender Systems". Πρόκειται για μοντέλα αλγορίθμων λογισμικού που τροφοδοτούνται με δεδομένα από έναν ή περισσότερους χρήστες (ανάλογα με το είδος του συστήματος) και παράγουν μια σειρά από συμπεράσματα ή απευθείας προτιμήσεις για εκείνον / εκείνους. Τα συμπεράσματα και οι προτιμήσεις αποτελούν τις προβλέψεις του συστήματος για τον συγκεκριμένο χρήστη (ή χρήστες, αν είναι πολλοί). Φυσικά, είναι εφικτός ο συνδυασμός διαφορετικών ειδών συστημάτων συστάσεων, ώστε οι τελικές προβλέψεις να είναι πολυδιάστατες ή επικυρωμένες από περισσότερα στοιχεία. Όπως είναι λογικό, οι recommenders χρησιμοποιούνται ευρέως στον χώρο της μουσικής βιομηχανίας. Πέραν των υπηρεσιών παροχής μουσικής (ανάμεσα τους και οι προαναφερθείσες υπηρεσίες streaming) που έχουν "ανάγκη" τα συγκεκριμένα συστήματα για τη καλύτερη δυνατή παροχή των υπηρεσιών τους στους χρήστες, δημιουργώντας προτάσεις "στα μέτρα" τους, οι recommenders υιοθετούνται και από DJs, εταιρείες που επιθυμούν αναπαραγώγη συγκεκριμένης μουσικής στους χώρους τους (συνεργασία με εταιρείες παροχής μουσικών πακέτων προερχόμενα από recommenders) κ.α.. Αρκεί κανείς να αναλογιστεί το όφελος της χρήσης συστημάτων συστάσεων για να κατανοήσει το πόσο ουσιαστική είναι η ύπαρξη τους και πόσο όφελος είναι σε θέση να αποφέρει στη βιομηχανία της μουσικής. Άλλωστε, δεκαετίες τώρα οι άνθρωποι έχουν υποθέσει πως "αν μπορούσαμε να γνωρίζουμε το μέλλον, θα είμασταν πλούσιοι". Οπότε, οποιαδήποτε απόπειρα κατανόησης των μοτίβων των μελλοντικών αποφάσεων εκατοντάδων χιλιάδων ανθρώπων (ακόμα και εκατομμυρίων), μπορεί να

---

αποφέρει τεράστια κέρδη. Φυσικά, οι recommenders δεν χρησιμοποιούνται αποκλειστικά και μόνο για οικονομικό όφελος. Τα συμπεράσματα που εξάγουμε από ένα σύστημα συστάσεων μπορούν να βοηθήσουν την επιστήμη να κατανοήσει πολλά, τόσο για τη προσωπικότητα και τη ψυχοσύνθεση των ανθρώπων, όσο και για τη πρόσδοτης ανάλυσης δεδομένων / μηχανικής μάθησης / τεχνητής νοημοσύνης. Εν ολίγοις, πρόκειται για εξαγωγή συμπερασμάτων που αφορά διάφορα επιστημονικά πεδία, κάτι το ιδιαίτερα χρήσιμο.



**Σχήμα 1.5:** Τα συστήματα συστάσεων είναι έτσι κατασκευασμένα, ώστε να προβλέπουν τις προτιμήσεις του χρήστη με βάση υπάρχουσες πληροφορίες γύρω από εκείνου ή γύρω από σχετικούς με αυτόν χρήστες [2].

Κάνοντας λοιπόν ένα γρήγορο συλλογισμό, μπορεί κανείς εύκολα να αντιληφθεί τη πληθώρα εφαρμογών που θα είχαν (και έχουν) οι recommender systems πάνω στα είδη και τις κατηγορίες μουσικής σήμερα. Μία ενδιαφέρουσα ιδέα μάλιστα θα ήταν ο συνδυασμός των συστημάτων συστάσεων με τους αλγορίθμους τυχαίας αναπαραγωγής μουσικών κομματιών. Πως είναι εφικτό κάτι τέτοιο; Με μια πρώτη σκέψη, οι δυο έννοιες μοιάζουν να αλληλοκαλύπτονται. Άλλωστε, ποιο το νόημα της τυχαίας αναπαραγωγής κομματιών για ένα χρήστη, αν ο στόχος είναι να προβλεφθούν οι προτιμήσεις των ειδών και η συγκεκριμενοποίηση των ακόλουθων τραγουδιών; Ουσιαστικά, αυτό είναι το μεγαλύτερο κομμάτι του αντικειμένου της παρούσας διπλωματικής εργασίας. Οι recommenders μπορούν να αξιοποιηθούν με πολλούς και διαφορετικούς τρόπους. Ένας εξ αυτών είναι και η αρμονική συνύπαρξη με μοντέλα τυχαίας αναπαραγωγής σε (παραδείγματος χάρη) μια υπηρεσία ροής μουσικής (streaming) όπως είναι το Spotify. Το νόημα και η χρησιμότητα αυτής της

ιδέας φυσικά θα αναλυθούν παρακάτω. Το σίγουρο είναι πως, η ιδέα που παρουσιάζεται μέσω της διπλωματικής εργασίας, αποτελεί άλλη μια απόδειξη πως η ψηφιακή εποχή της μουσικής βιομηχανίας έχει προσδώσει αμέτρητες νέες δυνατότητες που διευκολύνουν (ή ενδέχεται να διευκολύνουν) τη ζωή των χρηστών, συμβάλοντας στην ακόμα μεγάλυτερη απόλαυση των μελωδιών και τη βελτίωση της έννοιας της ψυχαγωγίας στη μουσική.

## 1.1 Αντικείμενο της διπλωματικής

Στόχος της παρούσας διπλωματικής εργασίας είναι να εξετάσει τη δυνατότητα συνδυασμού ενός συστήματος συστάσεων και της δυνατότητας τυχαίας αναπαραγωγής μουσικών κομματιών, με σκοπό να αντιληφθεί τις προτιμήσεις του χρήστη και να προβλέψει κομμάτια που εκείνος επιθυμεί. Τα κομμάτια θα εξακολουθούν να είναι μέρος της λίστας τυχαίας αναπαραγωγής, όμως τα μουσικά τους είδη θα είναι “σύμφωνα” με τις προτιμήσεις του χρήστη εκείνη τη χρονική στιγμή. Όμως, η συγκεκριμένη ιδέα φέρει και κάποιες ιδιαιτερότητες που τη καθιστούν ξεχωριστή από παρόμοιες υλοποιήσεις ή γενικότερες υπάρχουσες προτάσεις recommenders στη μουσική. Οι ιδιαιτερότητες έχουν ως εξής:

1. Η κύρια λειτουργία του recommender θα πρέπει να λαβαίνει χώρα σε πραγματικό χρόνο. Το σύστημα θα πρέπει να αντιλαμβάνεται τις προτιμήσεις του χρήστη μια δεδομένη χρονική στιγμή ακριβώς, χωρίς να έχει προηγουμένως πληροφορίες σχετικά με τα είδη μουσικής που δεν προτιμάει
2. Οι πληροφορίες που θα λαβαίνει το σύστημα συστάσεων θα αφορούν αποκλειστικά αρνητικά δείγματα. Εν ολίγοις, όταν ο χρήστης θα “προσπερνάει” (κάνει “skip”) ένα κομμάτι που θα εμφανίζεται από τη τυχαία λίστα αναπαραγωγής, εκείνο θα αποτελεί πληροφορία για τον recommender (ως αρνητικό δείγμα) και εκείνος θα καλείται εν συνεχείᾳ να εξάγει αποτελέσματα
3. Λόγω της φύσης του προβλήματος, ο recommender θα πρέπει να είναι σε θέση να έχει διαθέσιμες προβλέψεις όσο πιο άμεσα γίνεται. Αυτό σημαίνει πως θα πρέπει να συγκλίνει στα επιθυμητά είδη μουσικής το συντομότερο δυνατό, πάντα λαβαίνοντας ως δεδομένα τα αρνητικά δείγματα (skips) του χρήστη.
4. Το σύστημα συστάσεων μας δεν θα λειτουργεί με βάρη. Οπότε, αν ο χρήστης επιλέξει να ακούσει ενα κομμάτι που προκύπτει από τη τυχαία λίστα αναπαραγωγής, ο recommender δεν θα προσθέτει βάρος σε αυτό και το είδος του. Στόχος της έρευνας είναι να αξιολογηθεί η επίδοση του συστήματος με αποκλειστική τροφοδότηση μέσω αρνητικών δειγμάτων, δυσκολεύοντας ουσιαστικά το έργο του.

5. Καθόλη τη διάρκεια της διαδικασίας, η τυχαία αναπαραγωγή θα πρέπει να συνεχίζεται ακάθεκτη.
6. Το θεωρητικό υπόβαθρο θέλει τον χρήστη να έχει επιλέξει ως λίστα αναπαραγωγής μια που έχει δημιουργήσει εκείνος και που αποτελείται από πολλά & διάφορα μουσικά είδη (το κύριο use case και λόγος ύπαρξης της ιδέας θα αναλυθεί παρακάτω). Οπότε, η λίστα αναπαραγωγής δεν θα περιέχει μόνο συγκεκριμένες μουσικές κατηγορίες, αλλά αρκετές περισσότερες. Έτσι, το έργο που καλείται να φέρει σε πέρας ο recommender είναι και πάλι πιο δύσκολο.

Διατηρώντας στο νου όλες τις παραπάνω ιδιαιτερότητες, θα προσεγγίσουμε το δημιουργηθέν πρόβλημα, θα προτείνουμε μια λύση, θα την αξιολογήσουμε και τελικά θα εξάγουμε χρήσιμα συμπεράσματα. Πόσο εύκολο είναι να δημιουργηθεί ένα σύστημα συστάσεων μουσικών κομματιών ταχείας εύρεσης προτιμήσεων σε πραγματικό χρόνο, και αυτό υπό τη λειτουργία τυχαίας αναπαραγωγής των κομματιών; Μέλει να δούμε παρακάτω.

## 1.2 Οργάνωση του τόμου

Πέραν της παρούσας εισαγωγής, η διπλωματική εργασία χωρίζεται σε δυο κεντρικούς πυλώνες, το θεωρητικό & το πρακτικό μέρος, καθώς και τον επίλογο. Στο θεωρητικό μέρος θα αναλυθουν οι τεχνολογίες και οι έννοιες που χρησιμοποιούνται στην επίλυση του προβλήματος μας. Επίσης, θα αναλυθεί επαρκώς η ιδέα της εργασίας, στην οποία αναφερθήκαμε μερικώς παραπάνω. Τέλος, θα συμπεριληφθούν υπάρχουσες υλοποιήσεις που πραγματεύονται τα συστήματα συστάσεων στον χώρο της μουσικής. Στο πρακτικό μέρος θα υπάρξει πλήρης περιγραφή της αρχιτεκτονικής και των λειτουργιών της λύσης που προτείνεται. Προχωρώντας παρακάτω, θα παρουσιαστεί η υλοποίηση της λύσεις, με τη ταυτόχρονη προβολή του σχετικού κώδικα, των λειτουργιών που επιτελεί κάθε κομμάτι αυτού και των αποτελεσμάτων. Επειτα, θα παρατεθούν τα αποτελέσματα των δοκιμών της υλοποίησης. Ακολουθεί η εξαγωγή συμπερασμάτων βασισμένων στα αποτελέσματα. Η εργασία ολοκληρώνεται με τον επίλογο, στον οποίο παρατίθενται συνολικά συμπεράσματα και μελλοντικές επεκτάσεις της υλοποίησης.



## **Μέρος Ι**

### **Θεωρητικό Μέρος**

---



## Κεφάλαιο 2

# Shuffle Play, Συστήματα Συστάσεων & Μηχανική Μάθηση

---

Σε αυτό το κεφάλαιο παρουσιάζονται οι τεχνολογίες και οι έννοιες που θα χρησιμοποιηθούν για την επίλυση του προβλήματος μας. Θα αναλύσουμε το θεωρητικό και εννοιολογικό τους υπόβαθρο, ώστε να είμαστε σε θέση να κατανοήσουμε πλήρως τον τρόπο λειτουργίας τους ως μέρος της γενικής ιδέας που προτείνει η παρούσα διπλωματική εργασία.

### 2.1 Shuffle Play

#### 2.1.1 Γενικός Ορισμός

Είναι γεγονός πως δεν υπάρχει ξεκάθαρος ορισμός για τη διαδικασία της τυχαίας αναπαραγωγής, γνωστής και ως shuffle play ή shuffling. Μπορούμε όμως να αποδώσουμε έναν γενικό ορισμό. Οπότε, η τυχαία αναπαραγωγή μουσικής (shuffle play), μπορεί να οριστεί ως μια μέθοδος μουσικής αναπαραγωγής κατά την οποία τα κομμάτια της λίστας αναπαραγωγής επιλέγονται με τυχαία σειρά. Συνήθως, η τυχαία αυτή σειρά δημιουργείται με την επιλογή της δυνατότητας τυχαίας αναπαραγωγής. Η σειρά διαφοροποιείται αν η τυχαία αναπαραγωγή απενεργοποιηθεί και ενεργοποιηθεί εκ νέου. Το shuffle play υπάρχει σχεδόν σε όλα τα ψηφιακά μέσα αναπαραγωγής μουσικής, είτε πρόκειται για H/Y, είτε για έξυπνες φορητές συσκευές (smartphones, tablets), είτε για απλούστερα audio players (MP3, MP4 players κλπ). Επιπρόσθετα, η δυνατότητα τυχαίας αναπαραγωγής εμφανίζεται και σε μερικά CD players [21] (όπως έχει ήδη αναφερθεί), με τη δυνατότητα τυχαίας επιλογής κομματιού από έναν ή περισσότερους δίσκους CD (ανάλογα με το είδος του μοντέλου και το αν μπορεί να δεχεθεί παραπάνω από έναν δίσκους).

Η διαδικασία της τυχαίας αναπαραγωγής μουσικής (αλλά και της τυχαίας αναπαραγωγής γενικότερα) δεν προϋποθέτει την ύπαρξη ενός πολύπλοκου αλγόριθμου. Μοναδικός σκοπός είναι το “ανακάτεμα” των μουσικών κομματιών μιας λίστας α-

να παραγωγής, έτσι ώστε ο χρήστης να μην ακούει τα κομμάτια με τη σειρά που εκείνα είναι καταχωρημένα. Πολλές λίστες αναπαραγωγής έχουν ομαδοποιήσει τα κομμάτια ανάλογα με το είδος τους. Άλλες πάλι κάνουν το ίδιο, αυτή τη φορά ανάλογα με τον καλλιτέχνη. Η επιλογή τυχαίας αναπαραγωγής δίνει τη δυνατότητα στο χρήστη να "περιηγηθεί" ελεύθερα (και προφανώς τυχαία) στα μουσικά κομμάτια της λίστας. Συνήθως δεν υπάρχει κάποια συνθήκη κατά την τυχαία αναπαραγωγή. Η λίστα ανασχηματίζεται και παραθέτει τα μουσικά κομμάτια σε σειρά που δεν φέρει κάποια "λογική" εξήγηση. Όπως αναφέρθηκε και κατά την εισαγωγή, η ψηφιακή εποχή της μουσικής βιομηχανίας κατέστησε την ενσωμάτωση του shuffle play στα νέα audio players ιδιαίτερα απλή. Αξίζει να αναφερθεί για άλλη μια φορά πως πολλές εταιρείες χρησιμοποίησαν τη δυνατότητα τυχαίας αναπαραγωγής ως ισχυρό στοιχείο marketing για τις φορητές συσκευές τους (σσ. iPod Shuffle). Πλέον, το shuffling είναι μια βασική δυνατότητα (ίσως και η πιο βασική) σε όλες τις εφαρμογές αναπαραγωγής πολυμέσων.



**Σχήμα 2.1:** Το καθολικά "αναγνωρισμένο" σύμβολο της τυχαίας αναπαραγωγής. Η πλειοψηφία των αναπαραγωγών πολυμέσων χρησιμοποιούν το συγκεκριμένο σύμβολο των διασταυρωμένων βελών για να υποδείξουν στο χρήστη τη δυνατότητα του shuffle play.

Οι λόγοι για τους οποίους οι άνθρωποι επιλέγουν να ψυχαγωγηθούν και να απολαύσουν τα μουσικά κομμάτια μιας λίστας τυχαία, ποικίλουν. Θα έλεγε κανείς πως ο κυριότερος λόγος αφορά τη θέληση ενός ατόμου να απολαύσει ένα ευρύ φάσμα μουσικών ειδών, χωρίς να παραμένει σε ενα συγκεκριμένο είδος ή καλλιτέχνη για μεγάλο χρονικό διάστημα. Αντιλαμβανόμαστε πως η συναισθηματική τάση για μουσική ακρόαση δεν περιορίζεται μόνο σε συγκεκριμένα είδη ή τραγουδιστές, αλλά σε συνδυασμούς μελωδιών (ενδεχομένως και φωνητικών). Μέσω της μουσικής, τα συναισθήματα γίνονται πιο έντονα, παλαιότερες αναμνήσεις επαναφέρονται στο "προσκήνιο" του νευρωνικού συστήματος, οι σκέψεις εξελίσσονται με υψηλότερη πολυπλοκότητα, η φαντασία ταξιδεύει με ακόμα μεγαλύτερη ταχύτητα και ολόκληρη η προσωπικότητα ενός ατόμου μοιάζει να "συμφωνεί" με τις νότες, τις μελωδίες και τις φωνές που εμπεριέχονται στο εκάστοτε μουσικό στοιχείο. Πολλοί άνθρωποι φτάνουν σε αυτή τη συναισθηματική έκσταση μέσω διαφόρων μουσικών ειδών. Η τυχαία αναπαραγωγή συμβάλλει σε αυτό κατά το μέγιστο. Πρόκειται για εκείνους τους ανθρώπους που, όταν ρωτούνται τι είδος μουσικής ακούν, εκείνοι απαντούν "Τα πάντα". Βέβαια, μερικά άτομα νιώθουν αυτό το είδος αδρεναλίνης μόνο με συγκεκρι-

μένα μουσικά είδη ή καλλιτέχνες. Για εκείνους, η τυχαία αναπαραγωγή αφορά την εναλλαγή κομματιών σε ένα πιο περιορισμένο πλαίσιο (πχ. shuffling μόνο ρε Rock κομμάτια ή μόνο σε τραγούδια του John Lennon), όμως η σημασία της παραμένει υψηλή. Αυτό γιατί η δυνατότητα τυχαίας αναπαραγωγής αναζωωγονεί συνεχώς το ενδιαφέρον και διατηρεί σε υψηλά επίπεδα τη θέληση για μουσική. Όσο τα μουσικά κομμάτια, ακόμα και ενός συγκεκριμένου είδους, εναλάσσονται τυχαία, ο ακροατής δεν βαριέται τόσο εύκολα. Ο κορεσμός για τη μουσική ακρόαση παρατείνεται.

Όπως είναι λογικό, τα μοτίβα τυχαίας αναπαραγωγής που υποσυνείδητα δημιουργούν οι άνθρωποι, μπορούν να προσφέρουν πολλά σε διάφορους επιστημονικούς τομείς. Έχουν διεξαχθεί έρευνες που αφορούν την εξέταση των λόγων για τους οποίους οι άνθρωποι προβαίνουν σε shuffle play των κομματιών τους στη λίστα που έχουν δημιουργήσει, ή σε λίστες που είναι διαθέσιμες μέσω υπηρεσιών ροής μουσικής [22]. Η ανάλυση μουσικών μοτίβων που δημιουργούνται από τις τάσεις αναπαραγωγής των χρηστών είναι μια ιδιαίτερα ενδιαφέρουσα διαδικασία, την οποία όμως δεν πραγματεύεται η παρούσα διπλωματική εργασία. Σε κάθε περίπτωση όμως, τέτοια είδη ανάλυσης πραγματοποιούνται από τις ίδιες τις υπηρεσίες streaming για τους χρήστες τους. Παραδείγματος χάρη, οι προσωποποιημένες προτάσεις μουσικών κομματιών που δημιουργούνται σε έναν χρήστη του Spotify, είναι αποτέλεσμα ανάλυσης των προτιμήσεων του και των επιλογών που έχει κάνει κατά την ακρόαση διαφόρων κομματιών, σε πολλά από τα οποία έχει μεταπρδήσει μέσω της τυχαίας αναπαραγωγής. Όμως, ο τρόπος με τον οποίο διαμορφώνεται η τελική λίστα των κομματιών προς εκείνον, γίνεται μέσω μιας άλλης διαδικασίας, εκείνης των συστημάτων συστάσεων.

## 2.2 Συστήματα Συστάσεων

### 2.2.1 Ορισμός

Τι είναι ένα σύστημα συστάσεων, γνωστό ευρέως ως Recommender ή Recommender / Recommendation System [23]; Πρόκειται για εργαλεία λογισμικού και τεχνικές που παρέχουν προτάσεις ή υποδείξεις για προϊόντα / υπηρεσίες / ιδέες / αγαθά που είναι πιθανότερο να ενδιαφέρουν έναν συγκεκριμένο χρήστη. Οι προτάσεις - υποδείξεις που παράγουν τα συστήματα συστάσεων είναι αποτελέσματα διάφορων διεργασιών λήψης αποφάσεων που βασίζονται σε ερωτήματα όπως τα "τι προϊόν να αγοράσω;", "τι είδος μουσικής να ακούσω;", "τι κατηγορία ειδήσεων να διαβάσω;" κ.α.. Τα συστήματα συστάσεων αποτελούν υποκλάσεις των Συστημάτων Φιλτραρίσματος Πληροφοριών (Information Filtering Systems) [24], μοντέλων που αφαιρούν ανεπιθύμητες ή μη χρήσιμες πληροφορίες από μια ροή ή ένα πακέτο δεδομένων, πριν τα δεδομένα αυτά φτάσουν στον τελικό χρήστη. Τα συστήματα Information Filtering αποτελούν τη βασική οικογένεια των recommenders, περιλαμβανομένων και

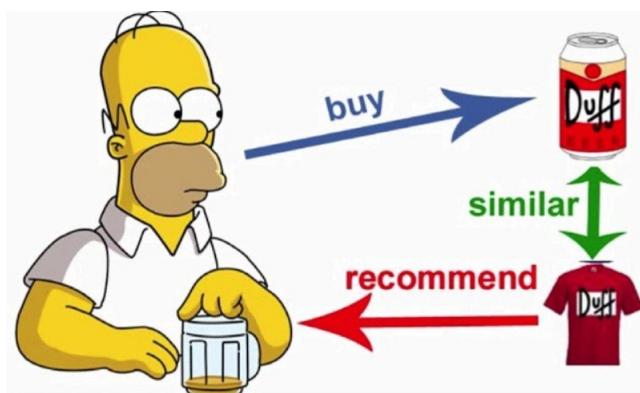
των πλατφορμών ανακάλυψης περιεχομένου (Content Discovery Platforms).

Συνήθως, ένα σύστημα συστάσεων εστιάζει στη δημιουργία προτάσεων για ένα συγκεκριμένο είδος αντικειμένου (πχ. μόνο τραγούδια ή μόνο ταινίες). Θα μπορούσε να θεωρηθεί πως ένας recommender αντικαθιστά τις προσπάθειες ενός ανθρώπου να εξετάσει και να αξιολογήσει πληροφορίες, ώστε να προτείνει εκείνος παρόμοια στοιχεία στον τελικό χρήστη. Για παράδειγμα, η λειτουργία ενός recommender σε ένα ηλεκτρονικό σύστημα δανεισμού βιβλίων από τη βιβλιοθήκη ενός πανεπιστημίου, θα αντικαθιστούσε τον ρόλο ενός βιβλιοθηκάριου που θα πρότεινε βιβλία στους φοιτητές, με βάση εκείνα που έχουν ήδη δανειστεί. Όπως είναι λογικό, τα συστήματα συστάσεων είναι δημιουργημένα ώστε (συνήθως) να παρέχουν προσωποποιημένες προτιμήσεις / προτάσεις. Αυτό σημαίνει πως ένας recommender λειτουργεί και παράγει προτάσεις που απευθύνονται κυρίως σε ενα συγκεκριμένο άτομο. Τα δεδομένα που έχουν τροφοδοτηθεί στον recommender θα προέρχονται και από το άτομο στο οποίο στοχεύουν οι προτιμήσεις. Στην απλούστερη μορφή του, ένα σύστημα συστάσεων παρέχει ως έξοδο μια λίστα με τις προτάσεις / προτιμήσεις σε σειρά συσχέτισης με τα τροφοδοτούμενα δεδομένα. Για παράδειγμα, το πρώτο στοιχείο της λίστας θα είναι και η επικρατέστερη πρόταση, το δεύτερη στοιχείο θα είναι η αμέσως επόμενη πιθανότερη κ.ο.κ..

Η ανάπτυξη των συστημάτων συστάσεων προήλθαν από την εξής απλή κοινωνική παρατήρηση: Οι άνρθρωποι τείνουν να βασίζονται στις επιλογές και τις απόψεις / αξιολογήσεις που κάνουν συνάνθρωποι τους για να αποφασίσουν στην προτίμηση ενός αντικειμένου / προϊόντος / υπηρεσίας. Αυτό σημαίνει πως φαίνεται να υπάρχει μια "διαδραστική" σχέση μεταξύ των καταναλωτών, όταν εκείνοι καλούνται να αποφασίσουν για την αγορά ή προτίμηση ενός αγαθού. Βέβαια, οι recommenders μπορούν να λειτουργήσουν με πληροφορίες που αφορούν αποκλειστικά έναν άνθρωπο (οπότε και να επιστρέψουν προτάσεις που τον αφορούν), κάτι που τονίστηκε και προηγουμένως. Τα είδη των συστημάτων που υπάρχουν αναφέρονται λίγο παρακάτω. Με βάση λοιπόν αυτή την βασική ιδέα - παρατήρηση, τα συστήματα συστάσεων εξελίχθηκαν σε ολόκληρο αντικείμενο μελέτης από επιστημονικούς κλάδους, αλλά και μεγάλο πλεονέκτημα στα χέρια των κολοσσών της παγκόσμιας βιομηχανίας & αγοράς. Από τα μέσα της δεκαετίας του 1990, τα συστήματα συστάσεων αποτελούν ένα αυτόνομο και ανεξάρτητο πεδίο έρευνας.

Ιδιαίτερα τα τελευταία χρόνια, το ενδιαφέρον γύρω από τους recommenders αυξάνεται συνεχώς. Κυριότεροι λόγοι για αυτό θα μπορούσαν να θεωρηθούν οι εξής: Πρώτον, η συντριπτική πλειοψηφία των μεγάλων παγκόσμιων ιστοτόπων χρησιμοποιεί κάποιο είδος recommender για τη καλύτερη δυνατή παροχή υπηρεσίων προς τον τελικό χρήστη. Τέτοιοι ιστότοποι είναι, μεταξύ άλλων, τα Amazon, YouTube, Google, Netflix, Spotify, Linkedin, Facebook, IMDb και Tripadvisor. Τα αποτελέσματα που μπορεί να έχει η υιοθέτηση ενός συστήματος συστάσεων στους σημερινούς ιστοτόπους είναι τεράστια, με απότερο σκοπό την αύξηση του κέρδους μέσω

καλύτερα επιλεγμένων διαφημίσεων ως προς τους χρήστες / επισκέπτες. Δεύτερον, εχουν αναπτυχθεί και ιδρυθεί συνεργασίες, συνέδρια και workshops που αφορούν αποκλειστικά τα συστήματα συστάσεων, όπως είναι η διεθνής κοινότητα Association of Computing Machinery (ACM), η οποία διεξάγει συνέδρια για τους recommenders (Conference Series on Recommender Systems (RecSys)). Πέραν αυτού, οι recommenders έχουν ενταχθεί ως θέματα σε άλλα επιστημονικά συνέδρια που προηγουμένως δεν είχαν "εκπροσώπηση". Ο τρίτος λόγος είναι πως τα συστήματα συστάσεων αποτελούν πλέον αυτόνομα μαθήματα σε πολλά ακαδημαϊκά ιδρύματα ανά τον κόσμο, τόσο σε προπτυχιακό, όσο και σε μεταπτυχιακό επίπεδο. Αυτό φυσικά έχει οδηγήσει και σε μια αυξημένη κινητικότητα του αντίστοιχου ερευνητικού σκέλους, ενώ ινστιτούτα όπως το IEEE έχουν αφιερώσει αρκετά issues επιστημονικών δημοσιεύσεων στους recommenders.



**Σχήμα 2.2:** Οι "commercial recommenders" παίζουν πλέον βασικό ρόλο στη προσπάθεια των εταιρειών να προσελκύσουν πελάτες, παρέχοντας τους όσο το δυνατόν καλύτερες προτάσεις προϊόντων που να αντιστοιχούν στις προτιμήσεις τους [3].

Όσον αφορά τις υπηρεσίες και εταιρείες που υιοθετούν τη χρήση των συστημάτων συστάσεων (συγκεκριμένα των commercial recommenders, καθώς έχουμε ήδη αναφέρει πως η χρήση τους δεν περιορίζεται μόνο εκεί), είναι προφανές πως μπορούν να αυξήσουν το κέρδος τους, εφόσον προωθούν τις προτιμήσεις / προτάσεις που φαίνεται να ταιριάζουν στον εκάστοτε χρήστη. Το κυριότερχο στοιχείο είναι πως στοχεύουν στην αύξηση των συνολικών πωλήσεων των αγαθών τους, αλλά δεν είναι μόνο αυτό. Η χρήση recommenders μπορεί να συμβάλλει στη πώληση προϊόντων / αγαθών που δεν είναι τόσο δημοφιλή (που σημαίνει ότι δεν αποφέρουν ιδιαίτερα κέρδη στον κατασκευαστή), αλλά θα ταιριάζουν στις προτιμήσεις του τελικού χρήστη. Επιπρόσθετα, η προώθηση των αγαθών που ανταποκρίνονται στα "θέλω" ενός χρήστη, θα αυξήσει ταυτόχρονα και το ενδεχόμενο ευχαρίστησης / ικανοποίησης εκείνου, οπότε και συνολικά τα ποσοστά ικανοποίησης των πελατών. Ακόμα, μέσω των recommenders, οι ιστότοποι είναι σε θέση να αναγνωρίσουν έναν πελάτη που τους έχει επισκεφθεί στο παρελθόν, αναγνωρίζοντας του το με τη παροχή προσωποποιημένων προσφορών ή εκπτώσεων. Τέλος, το σύστημα συστάσεων θα είναι τελικά σε θέση να κατανοήσει τι

πραγματικά θέλει ένας χρήστης καθώς επισκέπτεται τον ιστότοπο μιας υπηρεσίας / εταιρείας. Εν ολίγοις, ο recommender επιτυγχάνει τη "ψυχολόγηση" ενός πελάτη, κάτι που θα έκανε ένας υπάλληλος μέσω διαπροσωπικής επαφής, αλλά σε καλύτερο βαθμό και με υψηλότερα ποσοστά επιτυχίας.

Φυσικά, όφελος δεν υπάρχει μόνο από τις εταιρείες και τις υπηρεσίες που χρησιμοποιούν συστήματα συστάσεων. Τα πλεονεκτήματα πρέπει να γίνονται αντιληπτά και από τους ίδιους τους χρήστες. Η ύπαρξη των recommenders δεν αποσκοπεί μόνο στο όφελος εκείνων που το χρησιμοποιούν, αλλά φυσικά και εκείνων για τους οποίους χρησιμοποιείται. Το γενικό σύνολο των λειτουργιών των συστημάτων συστάσεων, μαζί με τους ουσιαστικούς λόγους για την υιοθέτηση αυτών, έχει οριστεί [25] και αποτελείται από έντεκα περιπτώσεις. Πρόκειται για τις ακόλουθες:

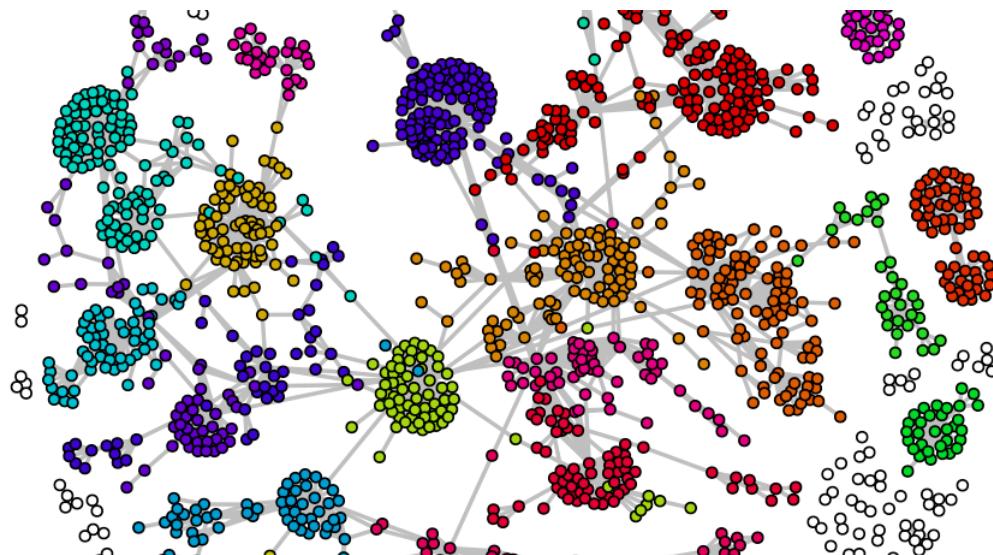
- **Εύρεση μερικών καλών αγαθών / προϊόντων / υπηρεσιών:** Παρουσίαση μερικών προτάσεων στο χρήστη, οι οποίες θεωρείται πως θα αρέσουν περισσότερο σε εκείνον. Πρόκειται για τη κυριότερη και πλέον συνηθισμένη λειτουργία των recommenders. Σε μερικές περιπτώσεις, οι προτάσεις που θα παρουσιαστούν, μπορούν να συνοδεύονται από το συντελεστή πιθανότητας επιλογής (πχ. μαζί με τα αστέρια που έχουν από το 1 έως το 5, ή μαζί με το ποσοστό άλλων που προτίμησαν το ίδιο).
- **Εύρεση όλων των καλών αγαθών / προϊόντων / υπηρεσιών:** Παρουσίαση όλων των προτάσεων στο χρήστη. Η βασική διαφορά με τη πρώτη περίπτωση είναι πως, αρχικά, ο recommender προτείνει μερικά αντικείμενα μέσα από μια λίστα πολλών. Οπότε, καλείται να κρατήσει μόνο εκείνα που κρίνει οτι του ταιριάζουν περισσότερο. Τώρα όμως, το σύστημα πρέπει να παρουσιάσει όλες τις επιλογές στο χρήστη, ακόμα και αν κάποιες δεν συγκεντρώνουν πολλές πιθανότητες προτίμησης. Αυτό είναι ουσιώδες στις περιπτώσεις που η λίστα διαθέσιμων αγαθών είναι μικρή, οπότε δεν υπάρχει περιθώριο αποκλεισμού κάποιων. Επίσης, το ίδιο ισχύει για ιδιαίτερα σημαντικές περιπτώσεις, όπου δεν πρέπει να αποκλειστεί κανένα ενδεχόμενο.
- **Επισήμανση προτιμήσεων σε γενικό πλαίσιο:** Δοθέντος ενός συγκεκριμένου γενικού πλαισίου (πχ. λίστα με προϊόντα από το super market), ο recommender καλείται να επισημάνει εκείνα που θεωρεί πως είναι πιο σημαντικά / πιθανό να προτιμηθούν, με βάση το ιστορικό επιλογών του χρήστη στο συγκεκριμένο πλαίσιο. Στο παράδειγμα των προϊόντων του super market, το σύστημα συστάσεων θα λάβει υπόψιν του το μακρύ ιστορικό αγορών του χρήστη και θα επισημάνει ποια προϊόντα είναι πιθανό να επιλεχθούν από τη δοθείσα τυχαία λίστα.
- **Πρόταση μιας ακολουθίας:** Ο recommender καλείται να προτείνει μια ακολουθία προτιμήσεων που κρίνει πως αντιστοιχούν στο χρήστη. Αυτό σημαίνει πως

δεν θα προτείνει αριθώς ανάλογα αγαθά / προϊόντα, αλλά εκείνα που ενδέχεται να επιλέξει μελλοντικά ο χρήστης, ως κομμάτι μιας ακολουθίας / μοτίβου επιλογών. Για παράδειγμα, αν ένας αναγνώστης επιλέξει ένα βιβλίο σχετικό με την ανάλυση δεδομένων, ο recommender θα είναι σε θέση να του προτείνει βιβλία που αφορούν τα... συστήματα συστάσεων, τη μηχανική μάθηση, τη τεχνητή νοημοσύνη κλπ.

- **Πρόταση μιας δέσμης:** Το σύστημα συστάσεων καλείται να προτείνει μια δέσμη αγαθών / προϊόντων / υπηρεσιών που φέρουν συσχετίσεις μεταξύ τους και “ταιριάζουν”. Για παράδειγμα, ένα πλάνο ταξιδιού θα περιλαμβάνει προορισμούς, αξιοθέατα, εστιατόρια και υπηρεσίες αναψυχής που βρίσκονται στο ίδιο γεωγραφικό σημείο. Επίσης, ένα κινητό τηλέφωνο μπορεί να συνοδεύεται από θήκη, μεμβράνη προστασίας και ακουστικά κεφαλής.
- **Έξυπνη πλοήγηση:** Κατά τη παραμονή του σε έναν ιστότοπο, μια υπηρεσία ψυχαγωγίας ή κάτι παρόμοιο, ο χρήστης θα λαμβάνει τη βοήθεια του recommender ώστε τα αντικείμενα που “πέφτουν” στο οπτικό του πεδίο όσο πλοηγείται, να είναι όσο το δυνατόν πιο κοντά στις προτιμήσεις του εκείνη τη χρονική περίοδο. Η συγκεκριμένη λειτουργία του recommender αφορά μόνο τη διάρκεια της πλοήγησης εκείνη τη στιγμή.
- **Διασφάλιση αξιοπιστίας προς το χρήστη:** Όπως είναι λογικό, πολλοί χρήστες δεν εμπιστεύονται τις προτάσεις που παρέχουν τα συστήματα συστάσεων. Άλλωστε, πολλοί ιστότοποι χρησιμοποιούν recommenders που προτείνουν προϊόντα παρόμοια με εκείνα που οι χρήστες έχουν ήδη αγοράσει, καθιστώντας την όλη διαδικασία ανούσια, εφόσον η αγορά έχει ήδη γίνει. Άλλοι χρήστες “παίζουν” με τα συστήματα συστάσεων, προσπαθώντας να δουν αν όντως εκείνα ανταποκριθούν στις πραγματικές προτιμήσεις τους. Ένας recommender μπορεί να παρέχει μικρές διεργασίες στον χρήστη, μέσω των οποίων εκείνος θα δοκιμάζει τις δυνατότητες πρόβλεψης προτιμήσεων. Για παράδειγμα, ένας ιστότοπος θα μπορεί να χρησιμοποιεί recommender που παροτρύνει τον χρήστη να κάνει μια δοκιμή (tutorial) των δυνατοτήτων του, διαλέγοντας ένα τυχαίο αγαθό / προϊόν και παρουσιάζοντας τις προβλέψεις.
- **Βελτίωση του προσωπικού παράγοντα:** Ένα σύστημα συστάσεων θα πρέπει να είναι σε θέση να βελτιώνει διαρκώς τη γνώση του γύρω από τον χρήστη, ώστε να κάνει τις μελλοντικές προτάσεις ακόμα πιο προσωπικές. Βέβαια, αυτό δεν είναι μόνο “στο χέρι” του recommender, εξαρτάται και από τη τροφοδότηση δεδομένων που θα λάβει. Σε κάθε περίπτωση, ο recommender πρέπει να είναι έτσι δομημένος, ώστε να καθιστά τις προβλέψεις του διαρκώς πιο σωστά τροποποιημένες ως προς στα “θέλω” του χρήστη.

- Έκφραση απόψεων από το χρήστη: Μερικοί χρήστες αδιαφορούν για τις προτάσεις που τους εμφανίζονται από τα συστήματα συστάσεων. Είτε πλοηγούνται στο διαδίκτυο ώστε να αγοράσουν ένα προϊόν, είτε ψάχνουν μουσική να ακούσουν, είτε εναλάσσονται τις διαθέσιμες ταινίες μιας υπηρεσίας για να παρακολουθήσουν μια, πολλοί χρήστες δεν δίνουν σημασία στα πεδία του τύπου "Μπορεί να σας αρέσουν:". Όμως, αρκετοί είναι εκείνοι που θέλουν να εκφράζουν τις απόψεις τους για το αγαθό / προϊόν / υπηρεσία, δίνοντας του μια βαθμολογία (καλό - κακό, ωραίο - άσχημο κλπ). Ένας recommender πρέπει να αξιοποιεί την παραχθείσα αυτή πληροφορία, ώστε να ενισχύσει τις προτάσεις που θα δώσει σε άλλους χρήστες, ή ακόμα και στον ίδιο.
- Βοήθεια άλλων: Παρόμοια με την παραπάνω περίπτωση, πολλοί χρήστες αφήνουν κριτικές, απόψεις, βαθμολογίες και εμπειρίες για διάφορα μέρη / προϊόντα / υπηρεσίες κλπ. Το σύστημα συστάσεων μπορεί να μην έχει άμεση εφαρμογή σε εκείνους, αλλά οι πληροφορίες που παράγουν, θα βοηθήσουν μελλοντικά στη δημιουργία προβλέψεων για άλλους χρήστες. Ο recommender συλλέγει όλη τη παραχθείσα γνώση και την αξιοποιεί. Η διαφορά με την παραπάνω περίπτωση είναι πως, τώρα, οι χρήστες εν γνώσῃ τους βαθμολογούν με σκοπό να βοηθήσουν άλλους (πχ. ενας αγοραστής ενός I.X. γράφει μια ειλικρινή αξιολόγηση για να προτρέψει - αποτρέψει άλλους από το να προβούν στην αγορά του).
- Επιρροή άλλων: Σε αντίθεση με τις προηγούμενες δυο περιπτώσεις (κυρίως την ακριβώς πρηγούμενη), πολλοί χρήστες έχουν ως σκοπό να παρασύρουν / επηρεάσουν άλλους μέσω κριτικών, αξιολογήσεων και βαθμολογιών για αγαθά και υπηρεσίες που δεν ανταποκρίνονται στη πραγματικότητα. Το κίνητρο τους μοιάζει να έχει κακή χροιά. Ισως πάλι να έχουν ως σκοπό να ωθήσουν τους καταναλωτές προς μια συγκεκριμένη κατεύθυνση, ενισχύοντας τις πωλήσεις / προβολές ενός συγκεκριμένου αγαθού. Ένας recommender θα πρέπει να έχει στοιχεία αντίληψης, ώστε κομμάτια πληροφορίας & γνώσης που συλλέγει να απορρίπονται.

Μέσα από αυτές τις έντεκα περιπτώσεις / λειτουργίες, θεσμοθετείται η ύπαρξη και το βασικό λειτουργικό πλαίσιο των συστημάτων συστάσεων. Φυσικά, οι recommenders εξελίσσονται συνεχώς, κάτι που σημαίνει πως οι λειτουργίες τους αυξάνονται. Πλέον ισχύουν πολλά περισσότερα από τα προαναφερθέντα. Μέσα από τις παραπάνω αναφορές όμως, αντιλαμβάνεται κανείς πως τα recommendation systems έχουν τεράστια εφαρμογή στο σήμερα. Η διαρκής, συνεχής συλλογή πληροφοριών και γνώσης με σκοπό τη δημιουργία προτάσεων & προβλέψεων για έναν ή περισσότερους ανθρώπους, είναι πραγματικά αξιοσημείωτη διαδικασία. Αν συλλογιστεί κανείς τη ποικιλία των εισερχόμενων πληροφοριών, αντιλαμβάνεται το "δύσκολο" έργο που έχουν να φέρουν σε πέρας τα σημερινά συστήματα συστάσεων.



**Σχήμα 2.3:** Οι εσερχόμενες πληροφίες στα συστήματα συστάσεων, όπως και οι κατηγορίες αυτών, είναι αχανείς. Θα μπορούσαν να παρουσιαστούν σε έναν γράφο εισερχόμενων στοιχείων γνώσης διαφορετικών ειδών. Μέσα από όλα αυτά, οι recommenders καλούνται να εξάγουν χρήσιμα συμπεράσματα και ασφαλείς προβλέψεις για χρήστες [4].

Οι χρήσεις των συστημάτων συστάσεων είναι αμέτρητες. Αναλογιζόμενοι τις παραπάνω λειτουργίες και καταστάσεις οι οποίες "διέπουν" τους recommenders, αντιλαμβανόμαστε πως ισχύει η αρχική συνθήκη, που δεν είναι άλλη από το γεγονός ότι οι χρήση τέτοιων συστημάτων ευνοεί τόσο εκείνους που την υιοθετούν, όσο και εκείνους για τους οποίους υιοθετείται. Τα recommendation systems δεν αναπτύχθηκαν ώστε οι εταιρείες να μπορέσουν να αυξήσουν το κέρδος τους. Όπως έχει ήδη αναφερθεί, οι recommenders μπορούν να συμβάλουν στη διευκόλυνση διαφόρων καταστάσεων και διεργασιών. Μπορεί όντως να χαίρουν αναγνώρισης από τη βιομηχανία και τη παγκόσμια αγορά, αλλά σε καμία περίπτωση αυτός δεν είναι ο αυτοσκοπός της ύπαρξης τους. Άλλωστε, η ιδέα της παρούσας διπλωματικής εργασίας χρησιμοποιεί συστήματα συστάσεων για να βοηθήσει τους χρήστες να απολαμβάνουν ακόμα πιο ποιοτικά τις περιόδους αναπαραγωγής & ακρόασης μουσικής. Σε κάθε περίπτωση, οι recommenders χωρίζονται σε διαφορετικά είδη, ανάλογα με τη χρήση και το σκοπό που έχει ο εκάστοτε υλοποιητής τους...

## 2.2.2 Είδη & Κατηγορίες

Οι κατηγορίες των recommenders διαφέρουν ανάλογα με τη τεχνική και το είδος των δεδομένων με τα οποία τροφοδοτούνται. Οπότε, ένα σύστημα συστάσεων καθορίζεται από τον τρόπο λειτουργίας του αλγορίθμου του και από τα δεδομένα με τα οποία τροφοδοτείται. Ανάλογα με τον σκοπό και τον λόγο χρήσης, γίνεται επιλογή και του αντίστοιχου συστήματος. Οι κύριες κατηγορίες των recommenders είναι οι εξής [23][26]:

- **Content-Based Recommendation:** Σε αυτή τη κατηγορία, το σύστημα μαθαίνει να κάνει προτάσεις αντικειμένων που αντιστοιχούν / ταιριάζουν / είναι παρόμοια με αντικείμενα που ο χρήστης επέλεξε (ή του άρεσαν) στο παρελθόν. Κάθε αντικείμενο συνοδεύεται από χαρακτηριστικά, των οποίων συγκρίνεται η ομοιότητα. Τα αντικείμενα που φέρουν τους υψηλότερους δείκτες ομοιότητας μεταξύ τους, επιστρέφονται από τον recommender ως έξοδο. Παραδείγματος χάρη, αν ένας χρήστης έχει παρακολουθήσει κωμωδίες αρκετές φορές στο παρελθόν, ο recommender θα τείνει να προτείνει ταινίες που ανήκουν σε αυτή τη κατηγορία. Αυτό το πέτυχε συγκρίνοντας την ομοιότητα χαρακτηριστικών διαφόρων ταινών με εκείνες που έχει ήδη δει ο χρήστης (και ανήκουν στη κατηγορία της κωμωδίας), επιστρέφοντας μια (ή περισσότερες) ταινία της οποίας τα σκορ ομοιότητας ήταν υψηλά. Σε γενικές γραμμές, αυτή είναι και η κύρια μέθοδος λειτουργίας των content-based recommenders. Συγκρίνουν την ομοιότητα χαρακτηριστικών κάποιων αντικειμένων και προχωρούν σε προβλέψεις. Πρόκειται για μια ευρέως διαδεδομένη μέθοδο. Όμως, αυτό το είδος συστημάτων συστάσεων είναι χρήσιμο μόνο σε συγκεκριμένες περιπτώσεις. Επίσης, μοιάζει να μην λαμβάνει υπόψιν του δεδομένα από άλλους χρήστες. Οι content-based recommenders είναι χρήσιμοι, αλλά δεν εφαρμόζονται σε όλα.
- **Collaborative Filtering:** Πρόκειται για τη δημοφιλέστερη και πλέον χρησιμοποιούμενη μέθοδο συστημάτων συστάσεων. Το collaborative filtering είναι αποδεδειγμένα μια ασφαλής επιλογή για πολλές περιπτώσεις στις οποίες χρειάζεται η υιοθέτηση ενός recommender. Η κυριότερη υλοποίηση θέλει το σύστημα να κάνει προτάσεις αντικειμένων στον χρήστη με βάση τις προτιμήσεις άλλων χρηστών που είχαν τα ίδια "γούστα" στο παρελθόν. Η ομοιότητα μεταξύ δυο χρηστών υπολογίζεται με βάση την ομοιότητα του ιστορικού βαθμολογήσεων / αξιολογήσεων τους σε αντικείμενα. Αυτός είναι και ο λόγος που το collaborative filtering έχει χαρακτηριστεί και ως people-to-people correlation. Η διαδικασία μπορεί να χωριστεί σε δυο διαφορετικές προσεγγίσεις. Η πρώτη εστιάζει στις συσχετίσεις μεταξύ των αντικειμένων και η δεύτερη στις συσχετίσεις μεταξύ των χρηστών. Οπότε, έχουμε μια item-based και μια user-based προσέγγιση. Στην πρώτη, η προτίμηση ενός χρήστη σε ένα προϊόν προβλέπεται μέσω των βαθμολογήσεων / αξιολογήσεων που έχει δώσει εκείνος σε παρόμοια προϊόντα. Προτιμάται σε "commercial" περιπτώσεις (ο recommender λειτουργεί με σκοπό τη προώθηση / πώληση προϊόντων) όταν οι χρήστες αριθμητικά είναι περισσότεροι από το πλήθος των διαθέσιμων προϊόντων προς πώληση, παρέχοντας αξιόπιστες προβλέψεις χωρίς την απαίτηση υψηλών υπολογιστικών πόρων. Στη δεύτερη προσέγγιση, οι προτιμήσεις ενός χρήστη υπολογίζονται μέσα από τα παρόμοια μοτίβα βαθμολογήσεων / αξιολογήσεων άλλων χρηστών. Εν ολίγοις, επιλύνει το πρόβλημα όντας "πιο πίσω" στην ιεραρχία

δεδομένων, αφού επεξεργάζεται μοτίβα βαθμολογήσεων των χρηστών και όχι τα συγκεκριμένα αντικείμενα που εκείνοι επέλεξαν. Προτιμάται κυρίως λόγω της "αυθεντικότητας" & πρωτοτυπίας των προτάσεων που παράγει, δηλαδή του γεγονότος πως προχωρεί σε προβλέψεις που ενδεχομένως να μην ακολουθούν ένα ίδιο προβλέψιμο μοτίβο, εφόσον αναφερόμαστε σε ομοιότητες χρηστών και όχι των αντικειμένων αυτών. Πλέον, το collaborative filtering κατηγοριοποιείται και σε δυο άλλους τύπους, ανάλογα με τον τρόπο λειτουργίας και όχι μόνο με το είδος των δεδομένων που τροφοδοτείται. Πρόκειται για τις memory based και model based περιπτώσεις [27]. Στη πρώτη περίπτωση, χρησιμοποιούνται ουσιαστικά οι item & user based προσεγγίσεις που μόλις αναλύθηκαν. Οι προσεγγίσεις αυτές είναι γνωστές και ως λύσεις "γειτονιάς", αφού εξετάζονται οι συσχετίσεις γειτονικών αντικειμένων ή χρηστών (με βάση την ομοιότητα των χαρακτηριστικών τους). Οπότε, πρόκειται για την ήδη διαδεδομένη περίπτωση που γνωρίζουν όσοι ασχολούνται με τα συστήματα συστάσεων. Στη δεύτερη περίπτωση όμως, αναπτύσσονται μοντέλα υπολογισμού προβλέψεων με τη χρήση τεχνικών ανάλυσης δεδομένων & μηχανικής μάθησης. Πρόκειται για διαδικασία που ενσωματώνει τις τελευταίες τάσεις των ερευνητικών και επιστημονικών κλάδων στους H/Y. Υλοποιώντας ένα model based collaborative filtering recommender, επιτυγχάνεται η μείωση διάστασης των δεδομένων αλλά και η απόδοση όσο το δυνατόν καλύτερων προβλέψεων.

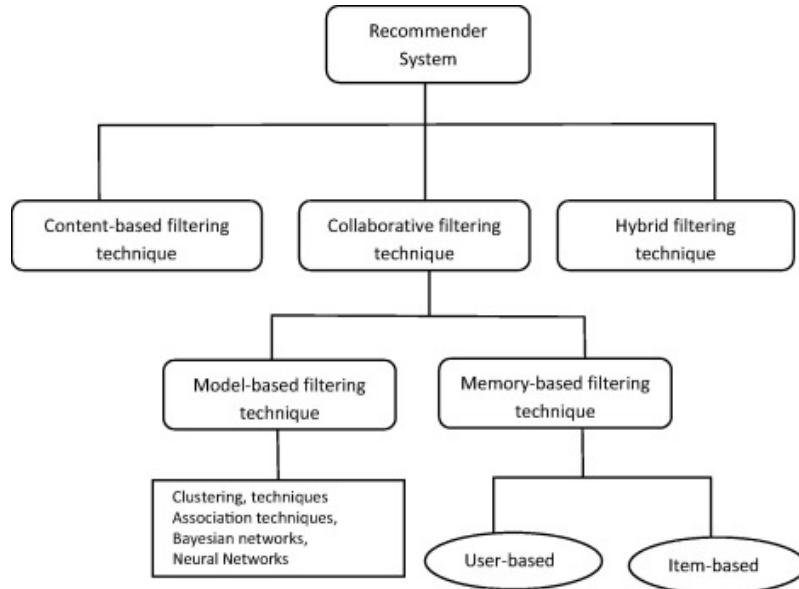
- **Demographic Recommendation:** Αυτή η κατηγορία συστημάτων συστάσεων προτείνει αντικείμενα στο χρήστη με βάση τα δημογραφικά στοιχεία του προφίλ του. Η κεντρική ιδέα είναι πως, διαφορετικές συστάσεις θα πρέπει να γίνονται για διαφορετικές δημογραφικές πληροφορίες. Για παράδειγμα, ένας ιστότοπος που χρησιμοποιεί αυτής της κατηγορίας τον recommender, θα παρέχει διαφορετικές προτάσεις στους χρήστες, ανάλογα με τη τοπική περιοχή, τη χώρα προέλευσης ή τη γλώσσα που εκείνοι ομιλούν. Επίσης, οι προτάσεις μπορεί να διαφέρουν και ανάλογα με την ηλικία του επισκεπτόμενου. Αξίζει να σημειωθεί πως, μέχρι σήμερα, η συγκεκριμένη κατηγορία δεν έχει την ίδια απήχηση με τις δυο πρώτες. Πρόκειται για μια κατηγορία που χρησιμοποιεί μόνο συγκεκριμένα είδη δεδομένων και αυτό ενδεχομένως να αποκλείει πολλές υλοποιήσεις στις οποίες θα μπορούσε να ενταχθεί. Βέβαια, αυτός είναι και ο λόγος ύπαρξης των demographic recommenders, να παρέχουν αξιόπιστες προτάσεις / προβλέψεις στους χρήστες μέσα από δημογραφικά στοιχεία και όχι μέσα από αντικείμενα ή βαθμολογήσεις.
- **Knowledge-Based Recommendation:** Η συγκεκριμένη κατηγορία δίνει μια διαφορετική προσέγγιση στον τρόπο λύσης του προβλήματος των συστάσεων. Αυτοί οι recommenders προτείνουν αντικείμενα, βασιζόμενοι σε συγκεκριμένη γνώση που αφορά τα χαρακτηριστικά των αντικειμένων και πως αυτά ταιριάζουν

στις ανάγκες, προτιμήσεις & τα "θέλω" τους. Αντιλαμβανόμαστε πως σε αυτή τη κατηγορία, οι λεπτομέρειες και επιπρόσθετες πληροφορίες (γύρω από συγκεκριμένα αντικείμενα) στη γνώση που τροφοδοτεί τους recommenders, παιζουν βασικό ρόλο στη καλύτερη δυνατή λειτουργία του. Τύποι των knowledge-based συστημάτων είναι τα i) case-based, όπου τα σκορ ομοιότητας των χαρακτηριστικών των αντικειμένων αποτελούν και το βασικό δείκτη πρότασης ή μη, ii) constraint-based, τα οποία δεν χρησιμοποιούν σκορ ομοιότητας για να προχωρήσουν σε προτάσεις, αλλά με βάση κανόνων που αφορούν τη συσχέτιση των αναγκών του χρήστη με τα χαρακτηριστικά των αντικειμένων, iii) community-based, όπου οι υπολογισμοί των προτάσεων γίνονται με βάση πληροφορίες για τις προτιμήσεις / συνήθειες και το ιστορικό των φίλων & γνωστών του χρήστη. Όπως και η κατηγορία δημογραφικών recommenders, έτσι και η τωρινή δεν χαίρει ιδιαίτερης απήχησης συγκριτικά με τα content-based & collaborative filtering μοντέλα.

- **Hybrid Recommenders:** Δεν πρόκειται για μια αυτόνομη κατηγορία συστημάτων συστάσεων, αλλά για έναν συνδυασμό κάποιων εκ των παραπάνω. Συγεκριμένα, ένας recommender θεωρείται υβριδικός, όταν συνδυάζει λειτουργίες και υλοποιήσεις διαφορετικών συστημάτων (πχ. content-based με collaborative filtering), αξιοποιεί τα πλεονεκτήματα αυτών και επιλύει μέρος των μειονεκτημάτων τους. Όπως είναι λογικό, έχουν προταθεί και χρησιμοποιηθεί διάφοροι συνδυασμοί recommenders. Μάλιστα, στη συγκεκριμένη κατηγορία επάγονται και οι συνδυασμοί εννοιών των συστημάτων συστάσεων με άλλες τεχνολογίες, όπως η μηχανική μάθηση, η ανάλυση δεδομένων και η τεχνητή νοημοσύνη. Ουσιαστικά, υβριδικός recommender θα μπορούσε να χαρακτηριστεί και το model-based collaborative filtering, αλλά ενδέχεται πχ. να αξιοποιεί μόνο γνώσεις μηχανικής μάθησης, χωρίς κάποια "παραδοσιακή" τεχνική συστάσεων. Όπως είναι λογικό, υλοποιήσεις υβριδικών συστημάτων θα κάνουν όλο και συχνότερα την εμφάνιση τους, ενώ ήδη φαίνεται να συγκεντρώνουν πάνω τους το ενδιαφέρον του ερευνητικού & επιστημονικού χώρου. Άλλωστε, ο συνδυασμός recommenders με ανάλυση δεδομένων και μηχανική μάθηση μπορεί να αποφέρει ιδιαίτερα ποιοτικά αποτελέσματα.

Σήμερα, τρεις είναι οι κατηγορίες που χρησιμοποιούνται κυρίως. Πρόκειται για τους content-based, collaborative filtering και hybrid recommenders. Αν συλλογιστεί κανείς πως ο συνδυασμός των τριών αυτών μοντέλων μπορεί να δημιουργήσει λύσεις για όλες τις περιπτώσεις, τότε θα του φανεί και λογική η σταδιακή εγκατάλειψη των demographic & knowledge-based κατηγοριών. Φυσικά, οποιοσδήποτε το επιθυμεί, μπορεί να δοκιμάσει μια "αυστηρή" υλοποίηση ενός από τους δύο, αλλά πλέον αυτό δεν συνηθίζεται τόσο. Θα μπορούσαμε να χαρακτηρίσουμε αυτές τις δύο περιπτώσεις ως πιο ειδικές, αφού πραγματεύονται συγκεκριμένα είδη πληροφοριών.

Ίσως αυτός να είναι και ο λόγος που χάνουν έδαφος συγκριτικά με τις πρώτες τρεις κατηγορίες.



**Σχήμα 2.4:** Σήμερα, τρεις εκ των προαναφερθεισών κατηγοριών *recommenders* χαίρουν υιοθέτησης και ενδιαφέροντος. Τα *content-based* συστήματα, η δημοφιλής *collaborative filtering* τεχνική και ο φουτουριστικός *hybrid* συνδυασμός υλοποιήσεων είναι οι κατηγορίες που έχουν ξεχωρίσει [5].

Επιπρόσθετα, αξίζει να σημειωθεί πως οι κατηγορίες των *recommenders* μπορούν να μπουν σε γενιές (συγκεκριμένα τρεις), ανάλογα με χρονικό στάδιο στο οποίο βρίσκονταν κάθε φορά η έρευνα γύρω από αυτούς [28]. Με τη πάροδο των τελευταίων ετών, όπου και αναπτύχθηκε ιδιαίτερα το ενδιαφέρον γύρω από τα συστήματα συστάσεων, έχουν λάβει χώρα σημαντικές βελτιώσεις & αλλαγές στη κατανόηση και τη λειτουργία των μοντέλων αυτών. Θα μπορούσαμε λοιπόν να ορίσουμε τις πρώτες βασικές κατηγορίες που αναλύσαμε παραπάνω (*Content-based*, *collaborative filtering*, *knowledge-based* και *hybrid recommendation*) στη πρώτη γενιά συστημάτων συστάσεων, όπως ορίστηκαν αρχικά και ξεκίνησε η έρευνα γύρω από εκείνα. Καθώς όμως ανακαλύφθηκαν και εξετάστηκαν νέες τεχνικές συστάσεων, όπως και συνδυασμοί με άλλες υπάρχουσες τεχνολογίες & έννοιες, περάσαμε στη δεύτερη γενιά *recommenders*, εντός της οποίας βρίσκονται οι εξελιγμένες προσεγγίσεις των πρώτων περιπτώσεων (πχ. *model-based collaborative filtering* και *εξερεύνηση* των δυνατοτήτων των υβριδικών μοντέλων). Τέλος, η τρίτη γενιά φέρνει τους συνδυασμούς των συστημάτων με τεχνικές ανάλυσης δεδομένων, μηχανικής μάθησης & τεχνητής νοημοσύνης. Εδώ συναντάμε μοντέλα *collaborative filtering* με χρήση αλγορίθμων βαθειάς μάθησης (*deep learning*), καθώς και υβριδικούς *recommenders* νέας γενιάς, με συνδυασμό τροφοδοτούμενων δεδομένων από χρήστες & αντικείμενα. Σε κάθε περίπτωση, είναι εμφανές πως η έννοια των συστημάτων συστάσεων εξελίσσεται συνεχώς. Καθώς ο χρόνος κυλάει και η μηχανική μάθηση διεισδύει σε όλο και

περισσότερους τομείς, μοιάζει θέμα χρόνου μέχρι να υιοθετηθεί καθολικά από όλες τις σύγχρονες υλοποιήσεις recommenders..!

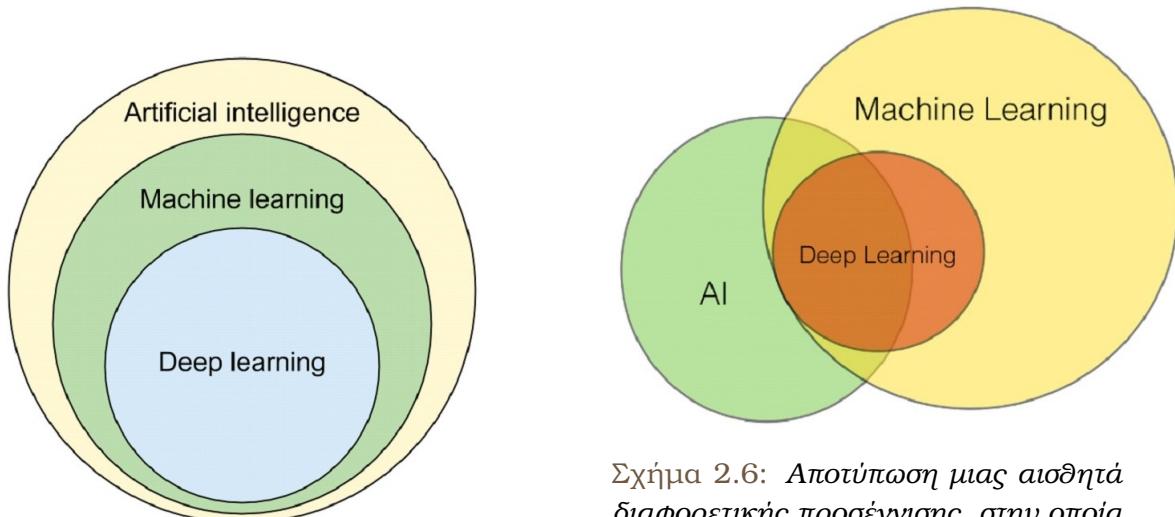
## 2.3 Μηχανική Μάθηση

### 2.3.1 Ορισμός

Η Μηχανική Μάθηση (Machine Learning) αποτελεί αδιαμφισβήτητα έναν από τους πλέον αναπτυσσόμενους κλάδους της επιστήμης των υπολογιστών. Μάζι με την ανάλυση δεδομένων, η μηχανική μάθηση (και η "αδελφική" ειδίκευση της, τεχνητή νοημοσύνη) μονοπωλούν το παγκόσμιο ερευνητικό ενδιαφέρον εδώ και αρκετά χρόνια, αρχίζοντας μάλιστα από τα μέσα του προηγούμενου αιώνα. Λόγω του γεγονότος πως οι τεχνικές της έχουν απήχηση σε αμέτρητες εφαρμογές και υλοποιήσεις στη παγκόσμια βιομηχανία, ιατρική, έρευνα & ακαδημαϊκή κοινότητα, η μηχανική μάθηση είναι ξεκάθαρα μια από τις επιστημονικές εμβαθύνσεις του μέλλοντος. Τι εστί Μηχανική Μάθηση [29]; Πρόκειται για τη μελέτη αλγορίθμων υπολογιστών που βελτιώνονται αυτόματα μέσω αποκόμισης εμπειρίας. Θεωρείται ως κομμάτι της τεχνητής νοημοσύνης ή πρόδρομος αυτής, αφού μέσω της μηχανικής μάθησης ένα σύστημα μπορεί να αποκτήσει ευφυΐα. Στην απλούστερη μορφή τους, τα μοντέλα μηχανικής μάθησης τροφοδοτούνται με συγκεκριμένα σύνολα δεδομένων, γνωστά ως "δεδομένα εκπαίδευσης" (training data), μέσω των οποίων εκπαιδεύουν τα αλγορίθμικά συστήματα τους και τελικά προβαίνουν στη παραγωγή προβλέψεων ή αποφάσεων. Το αξιοσημείωτο με τα μοντέλα αυτά είναι πως η εξαγωγή προβλέψεων δεν αποτελεί ξεκάθαρη προγραμματιστική εντολή από τον μηχανικό λογισμικό. Το αποτέλεσμα που παράγεται είναι αποκλειστική δημιουργία του μοντέλου. Λόγω της σημασίας που έχουν τα σύνολα δεδομένων στη μηχανική μάθηση, ο κλάδος της συχνά εξετάζεται μαζί με εκείνον της επιστήμης αυτών (των δεδομένων). Τα παραγόμενα αποτελέσματα των μοντέλων συνήθως αναλύονται και εξερευνώνται, με σκοπό τη καλύτερη δυνατή εξαγωγή συμπερασμάτων.

Η πρώτη αναφορά στον όρο "Machine Learning" έγινε το 1959 από τον Αμερικανό Arthur Lee Samuel, ο οποίος θεωρείται και εκ των πρωτοπόρων στον χώρο της τεχνητής νοημοσύνης, αλλά και στην εξέλιξη των ηλεκτρονικών παιχνιδιών H/Y [30]. Εκτότε, ο όρος "αιωρούνταν" στην επιστημονική & ερευνητική κοινότητα του AI (Artificial Intelligence), ξεκινώντας με βασικές μελέτες πάνω σε απλούς αλγορίθμους, μέχρι που τις επόμενες δεκαετίες εντάθηκε η έρευνα γύρω από πιο πολύπλοκα ζητήματα, όπως τα εξελιγμένα τεχνητά νευρωνικά δίκτυα, η αναγνώριση μοτίβων κλπ. Καθώς η ανθρωπότητα εισήλθε στη νέα χιλιετία, η μηχανική μάθηση έχαιρε άνθισης και παγκόσμιου ενδιαφέροντος, κάτι που εξακολουθεί να ισχύει (και σε ακόμα μεγαλύτερο βαθμό) μέχρι σήμερα. Όπως προαναφέρθηκε, οι δυνατότητες που φέρνει αυτός ο επιστημονικός κλάδος και το εύρος των τομέων στους οποίους μπορούν να

εφαρμοστούν οι τεχνικές του, τον καθιστούν έναν από τους πλέον αναπτυσσόμενους [31]. Η μηχανική μάθηση έχει ήδη εισχωρήσει στη καθημερινότητα. Όπουδήποτε και να κοιτάξουμε γύρω μας, η συντριπτική πλειοψηφία των ηλεκτρικών / ηλεκτρονικών συστήματων έχουν τώρα πια αποκτήσει στοιχειώδη ευφυΐα, συμβάλλοντας στη διευκόλυνση της ζωής μας, αλλά και στη καλύτερη δυνατή επιτέλεση οργανικών λειτουργιών στους χώρους των τηλεπικοινωνιών, της ιατρικής, της αγοράς, του διαδικτύου, της βιομηχανίας, των μεταφορών κ.α.. Δεν τίθεται αμφιθολία πως οι επεκτάσεις της μηχανικής μάθησης θα γίνονται ολοένα και πιο αισθητές. Για παράδειγμα, έχει ήδη σημειωθεί τεράστια πρόοδος στην έρευνα γύρω από τα αυτόνομα οχήματα, δηλαδή αυτοκίνητα που είναι σε θέση να πλοηγηθούν μόνα τους, χωρίς την ανάγκη παρέμβασης ανθρώπινου παράγοντα. Αντιλαμβάνεται λοιπόν κανείς πόσο ουσιαστικό είναι το βήμα της μηχανικής μάθησης, πριν το επόμενο βήμα της τεχνητής νοημοσύνης.



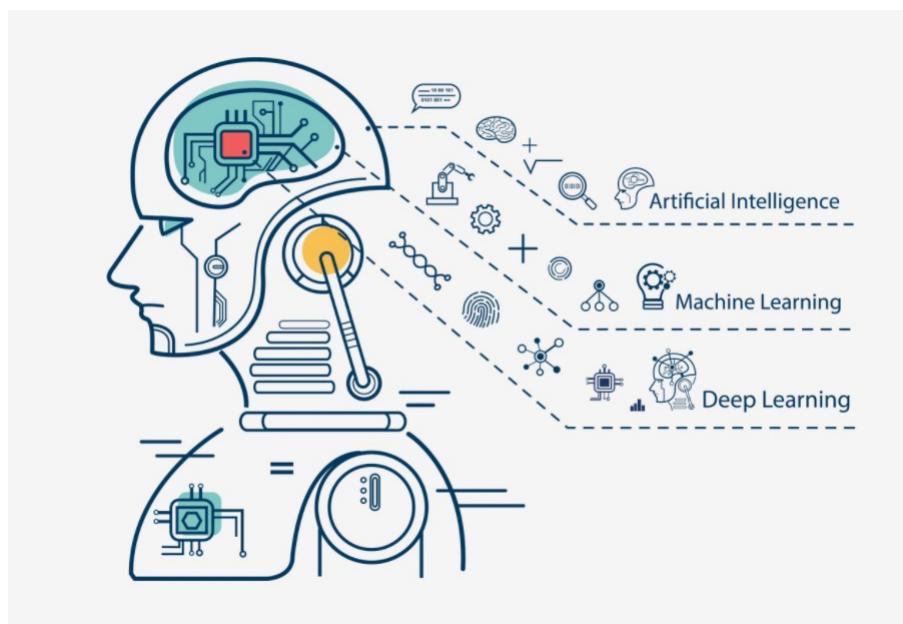
**Σχήμα 2.5:** Αποτύπωση της δημοφιλούς άποψης πως ο κλάδος της τεχνητής νοημοσύνης εμπεριέχει εκείνους της μηχανικής μάθησης και της νεότερης τάσης στο χώρο, το deep learning [6].

**Σχήμα 2.6:** Αποτύπωση μιας αισθητά διαφορετικής προσέγγισης, στην οποία η μηχανική μάθηση μοιράζεται αρκετά κοινά στοιχεία με τον τομέα της τεχνητής νοημοσύνης, αλλά η μία δεν εμπεριέχεται στην άλλη. Σε κάθε περίπτωση, το deep learning αποτελεί στοιχείο εμβάθυνσης της μηχανικής μάθησης [7].

Όμως, αξίζει να σημειωθεί πως, ιδιαίτερα από τις αρχές της δεκαετίας του 2010, το ενδιαφέρον έχει στραφεί σε μια υπο-κατηγορία του Machine Learning. Πρόκειται για τη Βαθεία Μάθηση, γνωστή ευρέως ως Deep Learning. Η επινόηση αυτής της έννοιας και του γεγονότος πως η μάθηση μοτίβων σε έναν υπολογιστή μπορεί να λάβει νέες εκτάσεις, είχε αναλυθεί από τις αρχές του 2000, όμως ήταν στα τέλη αυτής της δεκαετίας που ο επιστημονικός χώρος μπόρεσε να αξιοποιήσει πλήρως τη θεωρητική γνώση [32]. Με τον όρο "deep learning" ορίζουμε την οικογένεια των μεθόδων μηχανικής μάθησης που βασίζονται καθαρά σε προηγμένα συστήματα τεχνητών νευ-

ρωικών δικτύων και άλλων νεότερων τεχνικών (πχ. representation learning). Όπως είναι λογικό, η λειτουργία των μοντέλων βαθείας μάθησης απαιτούν υψηλότερη υπολογιστική ισχύ συγκριτικά με άλλα παραδοσιακά μοντέλα ML, ενώ οι περιπτώσεις στις οποίες χρησιμοποιούνται είναι συνήθως και πιο εξειδικευμένες.

Γίνεται λοιπόν γρήγορα αντιληπτό πως οι έννοιες της μηχανικής / βαθείας μάθησης και τεχνητής νοημοσύνης συνδέονται άρρητα μεταξύ τους. Βέβαια, κάθε ένας από αυτούς τους κλάδους αποτελείται από πολλά παρακλάδια, τα οποία είτε αποτελούν προσεγγίσεις επίλυσης, είτε άλλες θεωρητικές διατυπώσεις. Σε κάθε περίπτωση όμως, απαιτείται συθαρή αποκόμιση γνώσης γύρω από τα προαναφερθέντα πεδία, ώστε να είναι κανείς σε θέση να κατανοήσει πλήρως τις δυνατότητες, τον τρόπο λειτουργίας και τα αποτελέσματα αυτών. Παρακάτω, εμείς θα εξετάσουμε τις κυριότερες προσεγγίσεις / τεχνικές μοντέλων μηχανικής μάθησης, πάνω στις οποίες έχουν βασιστεί οι αλγόριθμοι που σήμερα αποτελούν κινητήρια γρανάζια σε ευφυή υπολογιστικά συστήματα. Να τονίσουμε πως, πάνω στις ίδιες προσεγγίσεις της μηχανικής μάθησης είναι δομημένη και η βαθεία μάθηση. Η παρούσα διπλωματική εργασία θα δοκιμάσει τη χρήση μοντέλων από ένα είδος τεχνικής Machine Learning, κάτι το οποίο θα παρατεθεί στο επόμενο κεφάλαιο. Επιπρόσθετα, να σημειώσουμε πως η επιστημονική ειδικότητα της μηχανικής μάθησης απαιτεί πολλές σελίδες ανάλυσης, απλώς και μόνο για να καλυφθούν οι βασικές αρχές, γνώσεις & έννοιες. Η παρούσα διπλωματική εργασία παρέχει μια απλή εισαγωγή στον κόσμο των έξυπνων υπολογιστικών συστημάτων, που επιτυγχάνουν την διαρκή τους εξέλιξη μέσω πολύπλοκων μοντέλων μηχανικής μάθησης.



**Σχήμα 2.7:** Απεικόνιση της πολυπλοκότητας που χαρακτηρίζει τα νέα συστήματα εκπαίδευσης υπολογιστών για απόκτηση ευφυΐας, αλλά και τη διασύνδεση μεταξύ τους. Η μηχανική μάθηση, η τεχνητή νοημοσύνη και το "νεότερο" deep learning έχουν άμεση σχέση στην επιτυχημένη λειτουργία ενός τεχνητού νευρώνα [8].

### 2.3.2 Προσεγγίσεις

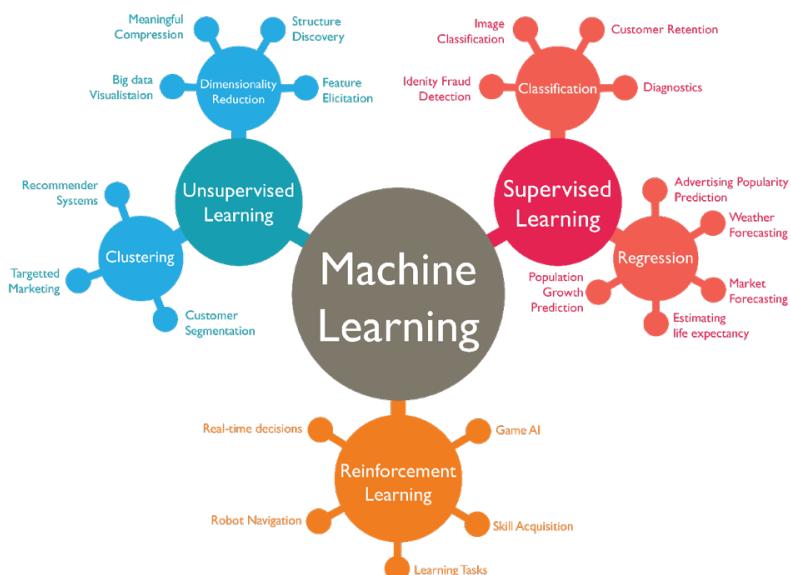
Όπως προαναφέρθηκε, το πεδίο της μηχανικής μάθησης χωρίζεται σε μια σειρά από είδη & προσεγγίσεις. Ανάλογα με το σύνολο των δεδομένων, τον επιθυμητό στόχο και τις υποδομές, γίνεται επιλογή του κατάλληλου τύπου / είδους. Οι κυριότερες προσεγγίσεις στις οποίες χωρίζεται η μηχανική μάθηση έχουν ως εξής [33]:

- **Supervised Learning:** Η ελληνική ονομασία αποδίδεται ως Επιβλεπόμενη ή Επιτηρούμενη Μάθηση. Πρόκειται ίσως για τον δημοφιλέστερο τύπο αλγορίθμων μηχανικής μάθησης. Στόχος της επιβλεπόμενης μάθησης είναι ο χαρακτηρισμός δεδομένων με βάση κάποια δεδομένα εκπαίδευσης. Ο αλγόριθμος “αντιστοιχεί” στοιχεία εισόδου με στοιχεία εξόδου, βασιζόμενο σε ενα σύνολο ζευγαριών εισόδου - εξόδου που του έχει ήδη τροφοδοτηθεί. Το σύνολο δεδομένων που περιέχει τους συνδυασμούς εισόδου - εξόδου, πάνω στους οποίους θα εκπαιδευτεί το μοντέλο του αλγορίθμου supervised learning, ονομάζεται σύνολο δεδομένων εκπαίδευσης (training data). Συνήθως, κάθε στοιχείο αποτελείται από ένα σύνολο τιμών εισόδου (μπορεί να είναι μια ή και περισσότερες) και μια επιθυμητή τιμή εξόδου (γνωστή και ως target ή label). Μετά την τροφοδότηση από τα δεδομένα, ο αλγόριθμος παράγει ένα εκπαιδευμένο μοντέλο που είναι σε θέση να κάνει προβλέψεις τιμών εξόδου για νέα δοθέντα στοιχεία εισόδου. Τα βήματα που ακολουθούνται συνήθως, η πλειοψηφία των οποίων είναι ίδια για όλες τις προσεγγίσεις, έχουν ως εξής: α) καθορισμός του είδους των δεδομένων εκπαίδευσης, β) συλλογή δεδομένων, γ) καθορισμός του τρόπου αναπαράστασης των εισερχόμενων στοιχείων από το μοντέλο εκμάθησης, δ) καθορισμός της δομής του μοντέλου & του αλγορίθμου εκμάθησης, ε) ολοκλήρωση του σχεδιασμού του συστήματος, στ) αξιολόγηση ακρίβειας του μοντέλου εκμάθησης. Γνωστοί αλγόριθμοι επιβλεπόμενης μάθησης είναι οι Support Vector Machines (SVMs), Linear Regression (Γραμμική Παλιδρόμηση), Logistic Regression (Λογιστική Παλινδρόμηση), Naive Bayes, Linear Discriminant Analysis (Γραμμική Διακριτική Ανάλυση), Decision Trees (Δέντρα Αποφάσεων), k-Nearest Neighbor, είδη Νευρωνικών Δικτύων (Multilayer Perceptron) και Similarity Learning (Μάθηση Ομοιότητας). Ουσιαστικά, ο σκοπός των αλγορίθμων επιβλεπόμενης μάθησης είναι να “προβλέψουν την επόμενη τιμή”, αφού έχουν εκπαιδευτεί από ενα σχετικό σύνολο δεδομένων.
- **Unsupervised Learning:** Η ελληνική απόδοση είναι Μη-επιβλεπόμενη ή Μη-επιτηρούμενη Μάθηση. Στόχος της μη-επιβλεπόμενης μάθησης είναι η ανακάλυψη πιθανής δομής που ενδεχομένως να “κρύβεται” πίσω από δεδομένα που δεν έχουν χαρακτηριστεί, δηλαδή δεν έχουν targets ή labels. Το ότι τα εισερχόμενα δεδομένα δεν είναι χαρακτηρισμένα, σημαίνει πως δεν έχουν ενα αντίστοιχο ζευγάρι εξόδου, όπως συμβαίνει στη περίπτωση του supervised

learning. Όλα τα στοιχεία του συνόλου, αποτελούν μόνο μονάδες εισόδου. Οπότε, ο αλγόριθμος μη-επιβλεπόμενης μάθησης καλείται να ανακαλύψει καποια πιθανή δομή μεταξύ τους, κάποια ομοιότητα, εφόσον δεν είναι σε θέση να γνωρίζει κάτι παραπάνω. Μάλιστα, εφόσον τα δεδομένα δεν είναι χαρακτηρισμένα, δεν υπάρχει σφάλμα ή σήμα ανταμοιβής ώστε να αξιολογηθούν οι πιθανές λύσεις. Εν ολίγοις, δεν μπορούμε να κάνουμε αξιολόγηση της ακρίβειας, όπως είναι το τελευταίο βήμα των supervised τεχνικών. Αυτό είναι που ξεχωρίζει τη μη-επιβλεπόμενη μάθηση από άλλες προσεγγίσεις. Οι δυο βασικές ομάδες αλγορίθμων μη-επιβλεπόμενης μάθησης είναι οι Πιθανοτικές Μέθοδοι και είδη Νευρωνικών Δικτύων. Στις πιθανοτικές μεθόδους συγκαταλέγονται τα Μοντέλα Εντοπισμού Ανωμαλιών (Local Outlier Factor, Isolation Forest), οι μέθοδοι Συσταδοποίησης (Ιεραρχική Συσταδοποίηση, αλγόριθμοι K-Means, DBSCAN & OPTICS), μοντέλα Λανθάνουσας Μεταβλητής (Expectation-Maximization - EM), αλλά και άλλοι αλγόριθμοι όπως οι Principal Component Analysis (PCA), Independent Component Analysis κλπ. Στα είδη Νευρωνικών Δικτύων συναντάμε τα Hopfield, Boltzmann, Restricted Boltzmann Machines (RBM), Helmholtz, Autoencoders Variational Autoencoders (VAE). Με λίγα λόγια, στόχος των unsupervised αλγορίθμων είναι να βγάλουν συμπεράσματα από τα δοθέντα δεδομένα, χωρίς να έχουν έναν "χάρτη" από labels, οπότε και ανακαλύπτουν μόνα τους τη δομή μεταξύ αυτών (των δεδομένων). Χρησιμοποιούνται συνήθως για διεργασίες συσταδοποίησης, αλλά και μείωσης διάστασης δεδομένων.

- **Semi-supervised Learning:** Πρόκειται για μια προσέγγιση μηχανικής μάθησης η οποία συνδυάζει ένα μικρό σύνολο χαρακτηρισμένων δεδομένων (με labels / targets), με ένα μεγάλο σύνολο μη-χαρακτηρισμένων. Ουσιαστικά η Ημι-επιβλεπόμενη Μάθηση αποτελεί ένα ενδιάμεσο "στάδιο" των επιβλεπόμενων & μη-επιβλεπόμενων προσεγγίσεων. Έχει αποδειχθεί πως η ανάμιξη μικρών συνόλων χαρακτηρισμένων δεδομένων με μεγάλα σύνολα unlabeled datasets μπορεί να παράξει βελτιωμένα μοντέλα υψηλότερης ακρίβειας. Η ημι-επιβλεπόμενη μάθηση αποτελεί χαρακτηριστικό παράδειγμα της έννοιας του weak supervision (ασθενής επίβλεψη) στη μηχανική μάθηση.
- **Reinforcement Learning:** Στα ελληνικά είναι γνωστή ως Ενισχυτική Μάθηση. Πρόκειται για το σύνολο αλγορίθμων και τεχνικών στις οποίες το σύστημα μάθησης προσπαθεί να μάθει μέσα από την άμεση αλληλεπίδραση με το περιβάλλον. Ουσιαστικά, ο αλγόριθμος ενισχυτικής μάθησης αλληλεπιδρά με ένα δυναμικό περιβάλλον στο οποίο πρέπει να επιτευχθεί ένας συγκεκριμένος στόχος. Χρησιμοποιείται στον έλεγχο κίνησης των ρομπότ, στη βελτιστοποίηση εργασιών σε εργοστάσια, στην επίλυση επιτραπέζιων παιχνιδιών κλπ. Παραδείγματα χρήσης μοντέλων reinforcement learning είναι η αυτόνομη οδήγηση ή

οι πράκτορες (AI robots) επιτραπέζιων παιχνιδιών. Η έννοια της ενισχυτικής μάθησης είναι εμπνευσμένη από τη μάθηση με επιβράβευση και τιμωρία, όπως συναντάται στα ερευνητικά / επιστημονικά μοντέλα μάθησης των έμβιων όντων. Σε γενικές γραμμές, ο αλγόριθμος καλείται να ανακαλύψει μόνος του τις όποιες ενέργειες πρέπει να κάνει, χωρίς να καθοδηγείται από κάποιον εξωτερικό επιβλέποντα. Φυσικά, ο τρόπος λειτουργίας του κάθε αλγορίθμου θεσπίζεται από την αρχιτεκτονική του. Δημοφιλείς υλοποιήσεις της ενισχυτικής μάθησης είναι τα μοντέλα Monte Carlo, Q-Learning, State-Action-Reward-State-Asction (SARSA) & Deep Q Network. To reinforcement learning είναι μια προσέγγιση που χρησιμοποιείται αρκετά στις μέρες μας, αφού είναι ιδιαίτερα αποδοτική σε πολύπλοκες διαδικασίες (λόγω της επικοινωνίας των αλγορίθμων με το περιβάλλον). Γι' αυτό, γίνεται έρευνα για την ανάπτυξη εξειδικευμένων μοντέλων βαθειάς ενισχυτικής μάθησης (Deep Reinforcement Learning).



Σχήμα 2.8: Οι κυριότερες προσεγγίσεις της μηχανικής μάθησης (εξαιρώντας τη semi-supervised περίπτωση), μαζί με τα είδη χρήσεων και τις εφαρμογές τους [9].

Όπως προαναφέρθηκε, οι τέσσερις αυτές προσεγγίσεις αποτελούν και τις κυριότερες. Η πλειοψηφία των αλγορίθμων μηχανικής μάθησης που χρησιμοποιείται σήμερα, αποτελούν σκέλη κάποιων εκ των supervised, unsupervised, semi-supervised & reinforcement learning. Όμως, υπάρχουν και άλλες προσεγγίσεις, πιο εξειδικευμένες και λιγότερο γνωστές. Μερικές εξ' αυτών υιοθετούνται σε αλγορίθμους που υπάγονται ήδη σε μια από τις τέσσερις βασικές κατηγορίες. Το Self Learning είναι μια περίπτωση που τείνει να μοιάζει με την ενισχυτική μάθηση, αλλά φέρει θεμελιώδεις διαφορές. Η διαδικασία μάθησης δεν φέρει καμία συμβουλή "δασκάλου", ούτε έχει κάποιο σύστημα ανταμοιβής [34]. Ο αλγόριθμος (ο οποίος συνήθως είναι ένα δίκτυο ονόματι Crossbar Adaptive Array - CAA) υπολογίζει μόνος του τις

αποφάσεις για δράσεις & αποτελέσματα των συνεπειών κάθε κατάστασης δράσης. Το Feature learning (γνωστό και ως Representation Learning) είναι μια προσέγγιση που μοιράζεται αρκετά κοινά με εκείνη της μη-επιβλεπόμενης μάθησης. Οι τεχνικές της προέρχονται σε ένα σύστημα μηχανικής μάθησης να ανακαλύψει αυτόματα αναπαραστάσεις από τα δεδομένα εισόδου, ώστε να επιτευχθεί ο εντοπισμός & εξαγωγή χαρακτηριστικών (feature detection). Εν ολίγοις, σκοπός των αλγορίθμων feature learning είναι να ανακαλύψουν καλύτερες αναπαραστάσεις των δεδομένων εκπαίδευσης (εισόδου) [35]. Παράδειγμα αυτής της κατηγορίας είναι η τεχνική PCA, που είδαμε προηγουμένως και στο unsupervised σκέλος. Άλλα είδη μηχανικής μάθησης αποτελούν τα Sparse dictionary learning, ο Εντοπισμός Ανωμαλιών (Anomaly Detection - έχει συμπεριληφθεί και παραπάνω ως μέρος του unsupervised learning), η Ρομποτική Μάθηση (Robot Learning - συγκενικές διεργασίες με εκείνες της ενισχυτικής μάθησης), αλλά και το σύνολο μεδόδων μάθησης με κανόνες, γνωστό ως Rule-based Machine Learning.

### 2.3.3 Αρνητικά Δείγματα ως Δεδομένα Δειγματοληψίας

Παρά το γεγονός πως δεν αποτελεί άμεσα κομμάτι της Μηχανικής Μάθησης, θεωρήθηκε καλό να ενταχθεί σε αυτό το σημείο ένας σύντομος ορισμός σχετικά με τα Αρνητικά Δείγματα. Άλλωστε, ένα σύνολο δειγμάτων μπορεί να αποτελεί ένα σύνολο δεδομένων εισόδου σε machine learning αλγόριθμο...! Εεκάθαρος ορισμός για την αρνητική δειγματοληψία δεν υπάρχει. Το Negative Sampling μπορεί να οριστεί ως η τεχνική που χρησιμοποιείται σε συστήματα & αλγορίθμους επεξεργασίας δεδομένων (μηχανικής μάθησης ή μη), όταν οι αρνητικές παρατηρήσεις των στοιχείων στα δεδομένα χαίρουν μεγαλύτερης σημασίας. Στη περίπτωση της μουσικής αναπαραγωγής, ως αρνητικά δείγματα μπορούν να χαρακτηριστούν τα κομμάτια τα οποία ο χρήστης "προσπερνάει" (κάνει skip) κατά τη διάρκεια ακρόασης μουσικής μέσα από μια λίστα αναπαραγωγής. Όσα τραγούδια επιλέγει να ακούει, αποτελούν τα θετικά δείγματα του συνόλου δεδομένων. Ο ορισμός του negative sampling μπορεί να διαμορφωθεί ανάλογα το είδος των δεδομένων και το σκοπό του προβλήματος. Θα μπορούσαμε να πούμε με ασφάλεια όμως πως, κατά βάση, τα στοιχεία που κατατάσσονται ως αρνητικά, φέρουν τους χαρακτηριστικούς του "ανεπιθύμητου" και "κακού" στο σύνολο δεδομένων που ανήκουν. Αρνητικά δείγματα σε ένα σύνολο φωτογραφιών θα μπορούσαν να είναι όσες είναι παραμορφωμένες, σε μια γκαρνταρόμπα ρούχων όσα είναι μικρότερου μεγέθους ή ξεθωριασμένα. Ακόμα, αρνητικά δείγματα θα μπορούσαν να είναι και όσα πιάτα δεν διαλέξουμε από ένα μενού εστιατορίου. Ανάλογα με τη περίπτωση, αλλάζουν και κάποια κομμάτια του ορισμού. Η κύρια έννοια όμως παραμένει αναλλοίωτη.

## Κεφάλαιο 3

# Περιγραφή θέματος

---

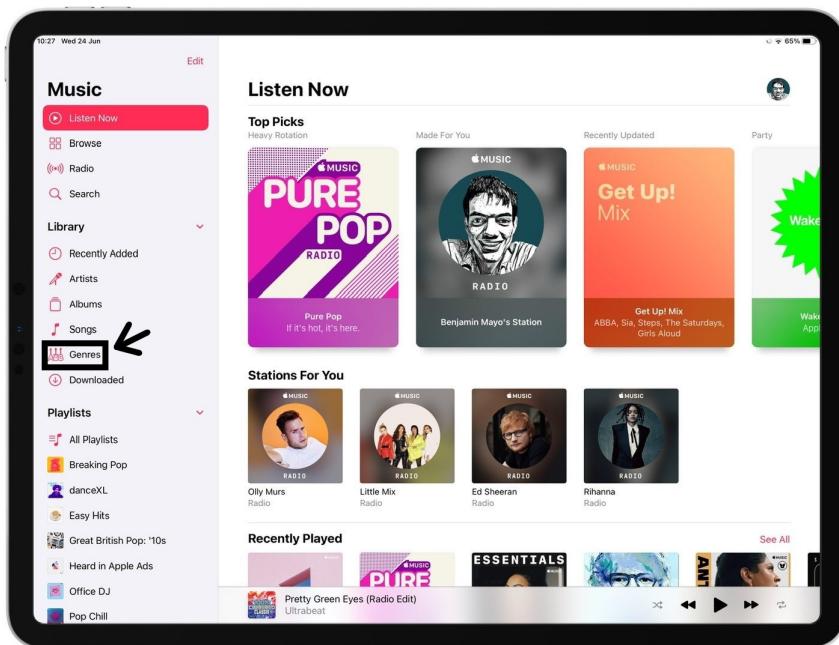
Στο παρόν κεφάλαιο θα προχωρήσουμε σε μια βασική ανάλυση του θέματος και της κεντρικής ιδέας που πραγματέυεται η διπλωματική μας εργασία. Αρχικά, καταγράφεται το "γιατί" να δοκιμάσουμε την υλοποίηση ενός τέτοιου συστήματος, καθώς και τους τρόπους που θα δοκιμαστούν κατά τη διάρκεια της επίλυσης αυτού. Έπειτα, παρουσιάζονται σχετικές εργασίες (υλοποιήσεις) που υπάρχουν ήδη στο διαδίκτυο και αφορούν την διεκπαιρέωση κομματιών που παρόμοια τους υπάρχουν και στη παρούσα διπλωματική. Ουσιαστικά, το τρίτο κεφάλαιο αποτελεί τη καλύτερη δυνατή εισαγωγή για το σάδιο της υλοποίησης που βρίσκεται παρακάτω.

### 3.1 Η Κεντρική Ιδέα

Η κεντρική ιδέα πίσω από την ανάπτυξη του θέματος της παρούσας διπλωματικής είναι η εξής: Οι περισσότεροι από τους ανθρώπους (αν όχι όλοι) που ακούν καθημερινά μουσική στις συσκευές τους, έχουν δημιουργήσει μια προσωπική λίστα αναπαραγωγής, η οποία συνήθως φιλοξενεί εκατοντάδες (ή και χιλιάδες) κομμάτια από διαφορετικά μουσικά είδη. Πρόκειται λοιπόν για μια συλλογή τραγουδιών που δεν μπορεί να χαρακτηριστεί από συγκεκριμένα είδη, ακριβώς λόγω της μεγάλης ποικιλίας αυτών εντός της. Είναι αυτή η "μεγάλη playlist" που έχουμε και στην οποία αποθηκεύουμε τα μουσικά κομμάτια που μας κινούν το ενδιαφέρον, αντανακλώντας τα συναισθήματα και τις σκέψεις μας. Συχνά, όταν θέλουμε να απολαύσουμε τα κομμάτια της προσωπικής μας λίστας, απλώς προχωρούμε σε τυχαία αναπαραγωγή αυτής. Όσα τραγούδια δεν τυχαίνει να "συμφωνούν" με τη διάθεση της στιγμής, απλώς προσπερνώνται, δηλαδή τους κάνουμε "skip". Όμως, υπάρχουν και περιπτώσεις που θέλουμε να ακούσουμε μόνο ένα συγκεκριμένο μουσικό είδος εντός της λίστας αναπαραγωγής μας. Δυστυχώς, δεν είναι εφικτό να απομονώσουμε εκείνη τη στιγμή όσα κομμάτια συμφωνούν στο είδος που επιθυμούμε. Το μόνο που επιτρέπει η λίστα, είναι να ακούσουμε κομμάτια ίδιου καλλιτέχνη στη σειρά (εφόσον τα τραγούδια είναι κατηγοριοποιημένα με αλφαριθμητική σειρά). Δεν μπορούμε όμως να ακούσουμε (για παράδειγμα) όλα τα rock κομμάτια. Έτσι, συνήθως καταφεύγουμε

στη κλασική λύση της τυχαίας αναπαραγωγής, κάνοντας συχνά και πολλαπλά skips σε όσα μουσικά κομμάτια δεν υπάγονται στο είδος που θέλουμε εκείνη τη στιγμή.

Σε αυτό το σημείο αξίζει να τονιστεί πως, πολλές εφαρμογές media / audio players έχουν ενσωματωμένους απλούς αλγόριθμους που προχωρούν σε κατηγοριοποίηση / κατανομή των αποθηκευμένων μουσικών κομματιών ανά το είδος. Παράδειγμα μιας τέτοιας εφαρμογής είναι το "Music" των συσκευών του οικοσυστήματος της Apple (προηγουμένως "iTunes" για προϊόντα με λογισμικό macOS, όμως πλέον φέρει την ίδια ονομασία σε όλες τις πλατφόρμες). Όμως, αυτό δεν συμβαίνει στα πλαίσια μιας συγκεκριμένης λίστας αναπαραγωγής. Εν ολίγοις, κάποιες εφαρμογές κατηγοριοποιούν όλα τα μουσικά κομμάτια που είναι αποθηκευμένα (ανεξαρτήτως λίστας), αλλά δεν κάνουν το ίδιο και για τα κομμάτια εντός μιας συγκεκριμένης λίστας αναπαραγωγής. Άλλωστε, μπορεί στη συσκευή μας να εχουμε δημιουργήσει παραπάνω από μια λίστες, ή να έχουμε τραγούδια "ελέυθερα", που δεν ανήκουν σε κάποια playlist. Οπότε, το αρχικό πρόβλημα παραμένει. Επιθυμούμε να πλοηγούμαστε σε ενα συγκεκριμένο είδος, εντός μιας συγκεκριμένης λίστας αναπαραγωγής.



**Σχήμα 3.1:** Η επιλογή κατηγοριοποίησης των μουσικών κομματιών ανά είδος, όπως φαίνεται στα αριστερά της εφαρμογής Apple Music, σε μια συσκευή Apple iPad. Όμως, δεν υπάρχει η ίδια δυνατότητα και για τα κομμάτια εντός των λιστών αναπαραγωγής [10].

Για ποιον λόγο όμως να θέλουμε κάτι τέτοιο; Ποιο είναι το σενάριο χρήσης του συγκεκριμένου προβλήματος; Το κυριότερο, πως προέκυψε η συγκεκριμένη ιδέα και εξελίχθηκε στη παρούσα διπλωματική εργασία; Συγκεκριμένα, δυο παραδείγματα-σενάρια χρήσης ήταν εκείνα που οδήγησαν στην ιδέα που πραγματεύεται η διπλωματική μας. Και τα δυο σενάρια έχουν αρκετά κοινά στοιχεία, ως κομμάτια της

καθημερινής ρουτίνας πολλών ανθρώπων. Στη πραγματικότητα, πρόκειται για περιπτώσεις που οι περισσότεροι φίλοι της μουσικής θα έχουν βιώσει, ιδιαίτερα άτομα των οποίων η μουσική αποτελεί αναπόσπαστο κομμάτι της καθημερινότητας. Άλλωστε, σε εκείνους απευθύνεται περισσότερο τούτη η εργασία. Στόχος της είναι να βελτιώσει μια λειτουργία των εφαρμογών αναπαραγωγής μουσικής, ώστε να καταστήσει πιο ποιοτική τη ψυχαγωγία μας μέσω αυτής. Ας δούμε πιο αναλυτικά τα δυο σενάρια-χρήσης που πυροδότησαν το αντικείμενο της διπλωματικής.

**Το πρώτο σενάριο** εχει ως εξής: Έστω ότι ενα άτομο επισκέπτεται το γυμναστήριο της γειτονιάς του, θέλοντας να κάνει το καθιερωμένο πρόγραμμα εκγύμνασης. Πάντα κατά τη διάρκεια της άσκησης έχει και τη συντροφιά της μουσικής, που τον/την βοηθάει δίνοντας ψυχολογική ώθηση. Το άτομο αυτό τυχαίνει να είναι συνδρομητής σε υπηρεσία ροής μουσικής (music streaming, όπως έχουμε ήδη αναλύσει παραπάνω), οπότε τις περισσότερες φορές επιλέγει τις λίστες αναπαραγωγής που η υπηρεσία του προτείνει πως ταιριάζουν καλύτερα κατά τη γυμναστική. Όμως, δεν είναι λίγες οι περιπτώσεις που το άτομο επιλέγει να ακούσει κομμάτια από τη δικιά του λίστα αναπαραγωγής, εκείνη που έχει δημιουργήσει μόνο του. Βέβαια, η λίστα αυτή περιέχει εκατοντάδες (ή και χιλιάδες) μουσικά κομμάτια, από πολλά και διάφορα είδη. Το άτομο δεν θέλει να ακούσει τραγούδια από όλα τα είδη της λίστας, μόνο εκείνα που θα τον βοηθήσουν / ωθήσουν ψυχολογικά κατά τη διάρκεια της γυμναστικής του. Η μόνη επιλογή που έχει, είναι να ξεκινήσει τη διαδικασία τυχαίας αναπαραγωγής και να κάνει συνεχόμενα skips μέχρι να “πέφτει πάνω” σε κομμάτια του είδους που θέλει να ακούσει και τα οποία είναι κατάλληλα για “παρέα” στην ώρα της εκγύμνασης. Κάτι τέτοιο δυσκολεύει τη κατάσταση για το άτομο, αφού ενδέχεται να χρειάζονται πολλά skips μεταξύ δυο κομματιών του είδους που θέλει. Όμως, δεν έχει άλλη επιλογή, από το να συνεχίσει έτσι..!

**Το δεύτερο σενάριο** είναι το ακόλουθο: Ένα άτομο εργάζεται σε εκπαιδευτικό ίδρυμα της Αθήνας ως ερευνητής, εκπονώντας ταυτόχρονα και τη διδακτορική του διατριβή. Λόγω του γνωστού κυκλοφοριακού προβλήματος της πρωτεύουσας, ιδιαίτερα στις περιοχές που βρίσκονται πιο κοντά στο κέντρο της, το άτομο καλείται καθημερινά να περνάει αρκετά λεπτά εντός του αυτοκινήτου του, καθώς κινείται με βραδύ ρυθμό λόγω της κίνησης το δρόμο. Η καλύτερη δυνατή συντροφιά που μπορεί να έχει εκείνη τη στιγμή, είναι λίγη μουσική. Παρά την αρχική του επιλογή να ακούσει ραδιόφωνο, αποφασίζει να συνδέσει τη φορητή συσκευή του με το ηχοσύστημα του αυτοκινήτου, θέλοντας να απολαύσει ένα συγκεκριμένο είδος από τα κομμάτια της δικιάς του λίστας αναπαραγωγής, το οποίο και θα τον τονώσει για τη δύσκολη ημέρα εν όψει. Όπως στο προηγούμενο σενάριο, έτσι και τώρα, το άτομο είναι χρήστης υπηρεσίας streaming, οπότε η λίστα αναπαραγωγής του είναι διαμορφωμένη εντός της εφαρμογής της υπηρεσίας. Καθώς επιλέγει τη λίστα του και ξεκινάει την ακρόαση, συνειδητοποιεί πως, για να ακούσει τα τραγούδια της συγκεκριμένης κατηγορίας ειδών που επιθυμεί, θα πρέπει να περνά και πολλά που δεν θέλει να ακούσει εκείνη

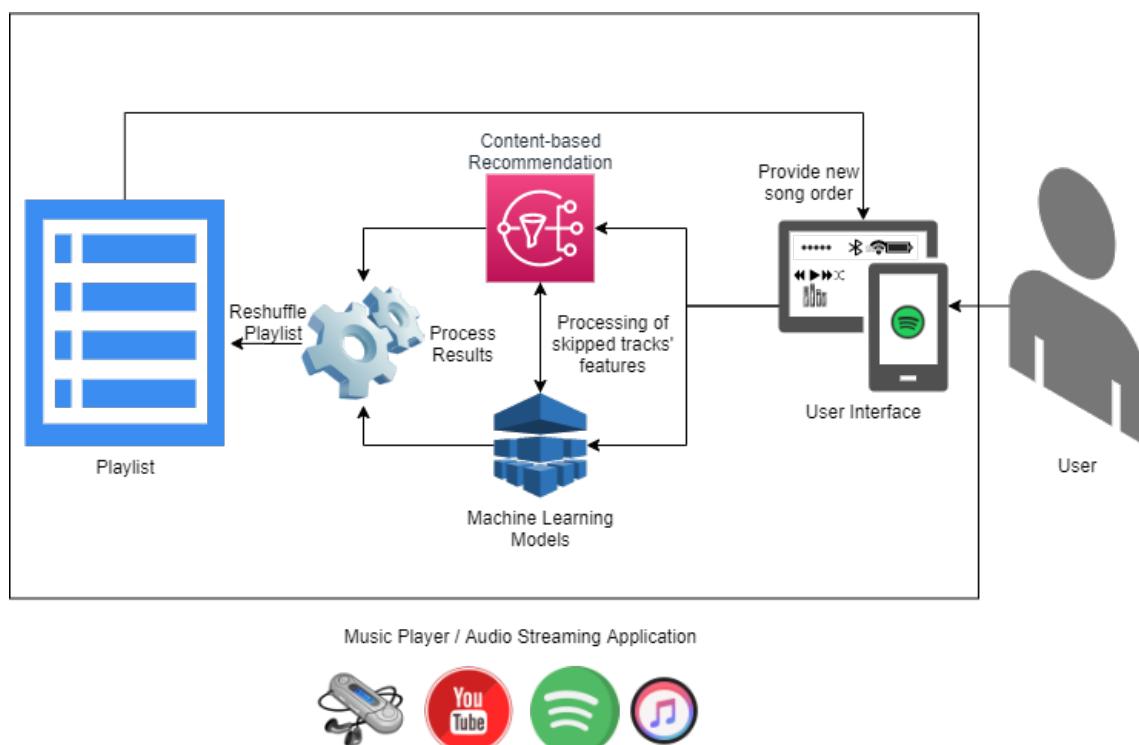
τη στιγμή. Εν ολίγοις, είναι “εγκλωβισμένος” σε μια διαδικασία πολλαπλών skips μεταξύ των τραγουδιών του είδους που αποζητά. Ουσιαστικά, πρόκειται για το ίδιο ακριβώς πρόβλημα που εντοπίζουμε και στο πρώτο σενάριο χρήστης...

Μέσα από τα δυο προαναφερθέντα σενάρια, εξάγουμε ενα βασικό συμπέρασμα. Μέχρι τώρα, οι εφαρμογές μουσικής ακρόασης δεν δίνουν τη δυνατότητα επιλογής συγκεκριμένου είδους μουσικής εντός των λιστών αναπαραγωγής. Για να απολαύσει κάποιος μουσικά κομμάτια συγκεκριμένης κατηγορίας ειδών εντός της λίστας, θα πρέπει να μεταπηδά από τραγούδι σε τραγούδι, προσπερνώντας όσα δεν θέλει. Κάπως έτσι, μέσα από τα “βιώματα” των παραπάνω σεναρίων, επήλθε η ιδέα δημιουργίας ενός νέου και διαφορετικού συστήματος συστάσεων, το οποίο θα μπορεί να εντοπίζει μια κατηγορία μουσικών ειδών που ο χρήστης επιθυμεί να ακούσει τη συγκεκριμένη χρονική στιγμή, ανάλογα με τη συμπεριφορά του, δηλαδή ανάλογα με τα skips που κάνει! Με λίγα λόγια, τι θα λέγαμε, αν η εφαρμογή μπορούσε να αντιληφθεί τις μουσικές προτιμήσεις του χρήστη εκείνη τη στιγμή, μέσω των τραγουδιών που θα προσπεράσει; Έτσι, μετά από μερικά skips, ο αλγόριθμος της εφαρμογής θα προσαρμόσει τα ακόλουθα μουσικά κομμάτια, “ελπίζοντας” πως αυτά θα ανταποκρίνονται στα “θέλω” του ατόμου εκείνη τη στιγμή. Αυτή είναι και η κεντρική ιδέα της παρούσας διπλωματικής εργασίας. Ας εξετάσουμε τα επί μέρους στοιχεία της, καθώς και τη προσέγγιση επίλυσης του προβλήματος.

Αρχικά, ας ξεκαθαρίσουμε πως, το πρόβλημα θα μπορούσε να λυθεί αν οι εφαρμογές ακρόασης μουσικής είχαν τη δυνατότητα κατηγοριοποίησης των κομματιών μιας λίστας αναπαραγωγής ανά είδος. Όμως, κάτι τέτοιο δεν υπάρχει, τουλάχιστον στις δημοφιλέστερες πλατφόρμες / υπηρεσίες ροής μουσικής. Συγκεκριμένα, η διπλωματική εργασία θα “πατήσει” πάνω σε δεδομένα και δυνατότητες που παρέχονται από το Spotify, που θεωρείται η δημοφιλέστερη streaming υπηρεσία σήμερα. Οπότε, επιστρέφοντας στο ζητούμενο, καλούμαστε να δημιουργήσουμε εμείς εναν τρόπο κατηγοριοποίησης των κομματιών εντός μιας λίστας. Στη πραγματικότητα όμως, η πρόταση κομματιών στον χρήστη (και στη προκειμένη εκείνων που συμφωνούν με τις μουσικές προτιμήσεις του τη συγκεκριμένη χρονική στιγμή) αποτελεί δουλειά των συστημάτων συστάσεων. Άρα, μοιάζει επιτακτική η ανάγκη δημιουργίας ενός τέτοιου συστήματος. Γνωρίζουμε όμως πως οι recommenders χρειάζονται σωστή τροφοδότηση δεδομένων για να λειτουργήσουν σωστά. Τι είδους δεδομένα θα τροφοδοτούμε το σύστημα μας; Πρόκειται για τα μουσικά κομμάτια τα οποία ο χρήστης θα προσπερνά, δηλαδή εκείνα στα οποία θα κάνει skip. Άρα, εφόσον δεν αναφερόμαστε στα θετικά δείγματα της λίστας (τα τραγούδια που θέλει να ακούσει), ορίζουμε το συγκεκριμένο σύνολο ως αρνητικό, πράγμα που σημαίνει πως ο recommender μας θα λειτουργεί με αρνητικά δείγματα.

Έχοντας εξακριβώσει το κεντρικό πλάνο της ιδέας, πλέον καλούμαστε να ερευνήσουμε τους τρόπους με τους οποίους θα δημιουργήσουμε το σύστημα συστάσεων μας. Επαναλαμβάνοντας άλλη μια φορά, στόχος είναι η κατανόηση των μουσικών

προτιμήσεων του χρήστη (ποια κατηγορία ειδών θέλει να ακούσει) μια συγκεκριμένη χρονική στιγμή, μέσα από μια λίστα αναπαραγωγής και με βάση τα αρνητικά του δείγματα (δηλαδή τα κομμάτα που προσπερνάει). Φυσικά, θα πρέπει να μελετηθεί και η ταχύτητα λειτουργίας του recommender. Το ιδανικό σενάριο θα ήταν να έχουμε αποτελέσματα σε πραγματικό χρόνο, εφόσον καλούμαστε να αντιληφθούμε τη προτίμηση του χρήστη εκείνη τη στιγμή. Αυτό θα είναι και ένα από τα ερωτήματα που θα απαντήσει η παρούσα διπλωματική εργασία. Αν δηλαδή είναι εφικτός ο εντοπισμός των προτιμήσεων σε πραγματικό χρόνο, ώστε να επιλυθεί το πρόβλημα που προκύπτει από σενάρια όπως τα δυο προαναφερθέντα. Οπότε, η αρχιτεκτονική του συστήματος συστάσεων είναι και εκείνη που πρέπει να οριστεί πρώτη, ώστε να προχωρήσουμε στο στάδιο της έρευνας & ανάπτυξης. Φυσικά, μαζί με την αρχιτεκτονική θα πρέπει να αναλύσουμε και τα δεδομένα μας (δηλαδή τα μουσικά κομμάτια), ώστε να επιλέξουμε τη καλύτερη δυνατή αξιοποίηση των πληροφοριών που λαβαίνουμε από αυτά. Οι βασικές προσεγγίσεις που θα ακολουθήσουμε για το σύστημα συστάσεων μας περιλαμβάνουν έναν content-based αλγόριθμο, καθώς και μοντέλα μηχανικής μάθησης. Στόχος μας είναι να εξετάσουμε αν οι δυο αυτές κατηγορίες μπορούν να συνδυαστούν, δημιουργώντας ετσι ένα υβριδικό σύστημα συστάσεων. Αν όχι, θα απαντήσουμε στο ποια εκ των δυο μπορεί να επιλύσει το πρόβλημα μας.



Σχήμα 3.2: Μια αρχιτεκτονική απεικόνιση του αρχικού θεωρητικού συστήματος που πραγματεύεται και ερευνά η διπλωματική εργασία. Δοκιμή μοντέλων μηχανικής μάθησης και content-based συστημάτων συστάσεων για εντοπισμό μουσικών προτιμήσεων χρήστη. Θα εξεταστούν μεμονομένα και αν τα αποτελέσματα είναι αξιόλογα, θα γίνει δοκιμή συνδυασμού αυτών.

Παρακάτω, στο πρακτικό μέρος της εργασίας, θα αναλυθούν οι προσεγγίσεις των content-based και machine learning κομματιών και θα οριστικοποιηθεί στο σύστημα συστάσεων. Καθώς κανείς κατανοεί την ιδέα της παρούσας διπλωματικής, ενδέχεται να υποθέτει πως ουσιαστικός της στόχος είναι η αναγνώριση, κατηγοριοποίηση & πρόβλεψη των μουσικών ειδών (genre recognition). Κάτι τέτοιο όμως δεν είναι αληθές. Στη πραγματικότητα, το genre recognition αποτελεί τον λειτουργικό στόχο των εργαλείων λογισμικού, με απότερο σκοπό τη σύσταση κομματιών που βρίσκουν σύμφωνες τις προτιμήσεις του χρήστη. Πολύ απλά, η αναγνώριση μουσικών ειδών είναι ένα υποσύστημα του κεντρικού μας συστήματος. Πρόκειται για βασικό διεκπαιρεωτικό εργαλείο, αλλά όχι την ουσία. Στόχος μας είναι να προτείνουμε τραγούδια, με βάση άλλα μουσικά κομμάτια που ο χρήστης δεν θέλει. Ο πιο εύκολος τρόπος να προσθέσουμε ή αφαιρέσουμε κοινά μουσικά κομμάτια από μια λίστα, είναι με βάση το είδος τους. Γι' αυτόν ακριβώς τον λόγο, επιλέγουμε να πειραματιστούμε με την αναγνώριση, τη πρόβλεψη & τη κατηγοριοποίηση των μουσικών genres. Αν μοναδικός μας στόχος ήταν αυτός, ενδεχομένως το έργο μας να ήταν ευκολότερο. Η σύσταση μουσικών κομματιών μπορεί να θεωρηθεί ως επέκταση της κατηγοριοποίησης των ειδών. Εξού και η θεώρηση πως το genre recognition αποτελεί υποσύστημα του γενικού μοντέλου.

Επιπρόσθετα, ένα άλλο ερώτημα που θα μπορούσε να έρθει στο μυαλό ενός αναγνώστη, είναι το γιατί δεν λαμβάνονται υπόψη και τα θετικά δείγματα (δηλαδή τα τραγούδια που ο χρήστης ακούει και δεν προσπερνάει), πέραν των αρνητικών, κάτι που θα μπορούσε να διευκολύνει τον τελικό στόχο. Είναι αλήθεια πως η δημιουργία recommenders βασίζεται, κατά κύριο λόγο, στα θετικά δείγματα των χρηστών. Για παράδειγμα, πολλά συστήματα συστάσεων προτείνουν στους καταναλωτές προϊόντα παρόμοια με εκείνα που αγόρασαν. Κάτι τέτοιο μοιάζει πιο εύκολο, κοιτώντας το υπό το πρίσμα της υλοποίησης. Όμως, εκείνο που έχει πραγματικά ενδιαφέρον, είναι να δούμε αν είναι εφικτό το να μπορέσουν να προταθούν στον χρήστη προϊόντα που ενδέχεται να του αρέσουν, τροφοδοτώντας όμως τους recommenders όλα εκείνα τα αγαθά που οι χρήστες προσπέρασαν κατά την αναζήτηση τους. Το πρόβλημα τώρα μοιάζει πιο δύσκολο, αλλά και πιο ενδιαφέρον ερευνητικά. Αυτός είναι και ο λόγος που η παρούσα διπλωματική εργασία δεν θα καταπιαστεί με τα μουσικά κομμάτια που ο χρήστης επιλέξει να ακούσει, δηλαδή με τα θετικά δείγματα. Αν μπορούσαμε να χρησιμοποιήσουμε μια λαϊκή ρήση, το σύστημα που αναπτύχθηκε προσεγγίζει το αρχικό πρόβλημα "μέσω Λαμίας", θέλοντας να διερευνηθεί κατά πόσο κάτι τέτοιο είναι ουσιώδες, χρήσιμο και εφικτό. Άλλωστε, αν λαμβάνονταν υπόψιν τα θετικά δείγματα, το σύστημα θα μπορούσε να εστιάσει στα μουσικά είδη των κομματιών που ο χρήστης άκουσε, να "θάψει" όλα τα υπόλοιπα και να συνεχίσει την αναπαραγωγή όσων "συμφωνούν" μελωδικά με τα πρώτα. Εν ολίγοις, το πρόβλημα θα μπορούσε να επιλυθεί αρκετά γρήγορα. Κατά πάσα πιθανότητα, η πιο λειτουργική λύση είναι ενας συνδυασμός των δυο, δηλαδή αξιοποίηση τόσο των θετικών, όσο και των αρνητι-

κών δειγμάτων Όμως, η παρούσα διπλωματική θα επιχειρήσει να πετύχει ένα καλό αποτέλεσμα, αξιοποιώντας μόνο τα αρνητικά δεδομένα μας.

## 3.2 Σχετικές Υλοποιήσεις

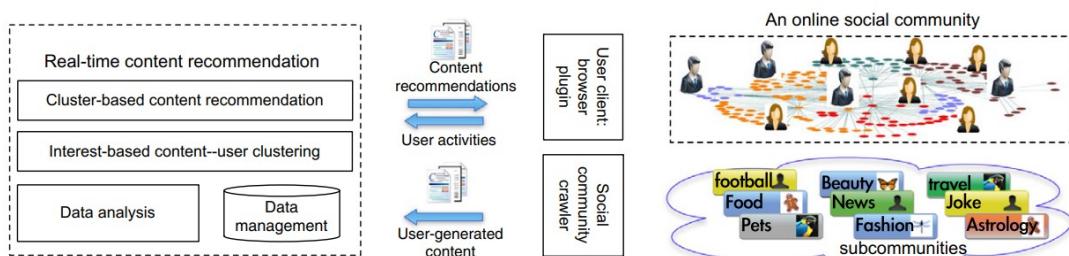
Θα χωρίσουμε το κομμάτι των σχετικών υλοποιήσεων σε δυο σκέλη. Το πρώτο σκέλος θα αφορά τις σχετικές ιδέες, δημοσιεύσεις και υλοποιήσεις που αφορούν το content-based recommendation. Το δεύτερο σκέλος αφορά σχετική δουλειά που έχει γίνει πάνω στην αναγνώριση μουσικών ειδών, με βάση τα χαρακτηριστικά τραγουδιών και με χρήση μοντέλων μηχανικής μάθησης. Ας εξετάσουμε κάθε σκέλος ξεχωριστά και ας δούμε τι έχει ήδη γίνει από άλλους ερευνητές, προγραμματιστές και επαγγελματίες. Όπως είναι λογικό, θα καλύψουμε τα κυριότερα θέματα και εκείνα που βρίσκονται πιο κοντά στην ιδέα της διπλωματικής. Άλλωστε, γενικές δημοσιεύσεις που αφορούν τα συστήματα συστάσεων και τα μοντέλα machine learning είναι εκατοντάδες χιλιάδες. Ας αναλύσουμε τα δυο σκέλη...

### 3.2.1 Υλοποιήσεις Συστημάτων Συστάσεων

Ξεκινώντας με το κομμάτι του **content-based recommendation**, αξιζει να τονιστεί πως η ιδέα που αναλυέται στη παρούσα διπλωματική φέρει κάποιες ιδιαιτερότητες που την ξεχωρίζουν από τις υπόλοιπες υλοποιήσεις. Συγκεκριμένα, τα χαρακτηριστικά της προσπάθειας δημιουργίας συστάσεων σε πραγματικό χρόνο (real-time), αλλά και του γεγονότος πως τα τροφοδοτούμενα δεδομένα θα είναι αρνητικά δείγματα (τα μουσικά κομμάτια που θα προσπερνώνται), καθιστούν το σύστημα της εργασίας μας ξεχωριστό. Σε κάθε περίπτωση, ας δούμε σχετικές ιδέες που συναντώνται σε επιστημονικές δημοσιεύσεις. Η έννοια / προσέγγιση του content-based recommendation περιλαμβάνει γενικούς αλγορίθμους, οι οποίοι μπορούν να εφαρμοστούν σε διάφορα είδη δεδομένων. Εν ολίγοις, δεν είναι ανάγκη να εντοπίσουμε ένα σύστημα συστάσεων που να ειδικεύεται στα μουσικά κομμάτια. Άλλωστε, ο τρόπος λειτουργίας των content-based συστημάτων είναι συγκεκριμένος. Οπότε, θα χωρίσουμε τις σχετικές προσεγγίσεις ως εξής. Πρώτα θα παραθέσουμε συστήματα συστάσεων που έχουν ήδη δημοσιευτεί και λειτουργούν σε πραγματικό χρόνο και έπειτα θα γνωρίσουμε recommenders που αφορούν συγκεκριμένα τη μουσική. Μοιάζει ως η πλέον λογική λύση για το σωστό διαχωρισμό των σχετικών υλοποιήσεων στα πλαίσια των συστημάτων συστάσεων.

Όσον αφορά τις δημοσιεύσεις που αφορούν recommenders πραγματικού χρόνου, αξιζει να σημειωθεί πως δεν εντοπίστηκε κάποιο μοντέλο που να επιχειρεί real-time συστάσεις μουσικών κομματιών. Όμως, υπάρχουν πολλές δημοσιεύσεις για recommendation systems που πληρούν τη προϋπόθεση της ταχείας δημιουργίας προτιμήσεων. Ιδιαίτερα ενδιαφέρουσα είναι μια υλοποίηση που δημοσιεύτηκε πριν από

μερικά χρόνια και γεννήθηκε μέσα από συνεργασία ερευνητών του Πανεπιστημίου Fudan της Κίνας & εκείνου του Κολοράντο των ΗΠΑ [11]. Πρόκειται για έναν 1-διαίτερα εξελιγμένο recommender που παράγει προτιμήσεις αγαθών / υπηρεσιών για χρήστες που ανήκουν σε κοινότητες εντός διαφόρων κοινωνικών δικτύων. Το σύστημα εκμεταλλεύεται πληροφορίες που υπάρχουν εντός των ψηφιακών κοινοτήτων, συλλέγοντας τες μέσω ενος crawler. Έπειτα, λαβαίνει χώρα η ανάλυση των συλλεχθέντων δεδομένων, η διαχείριση αυτών και τελικά η δημιουργία συστάσεων. Το αξιοσημείωτο είναι πως το σύστημα καταφέρνει να παράξει τις προτιμήσεις άμεσα, σε πραγματικό χρόνο. Έτσι, ο χρήστης “βλέπει” προτάσεις προϊόντων που ενδέχεται να τον ενδιαφέρουν, με βάση τις κοινότητες των κοινωνικών δικτύων στις οποίες είναι μέλος. Πρόκειται πραγματικά για μια πολύ ενδιαφέρουσα υλοποίηση (όπως προαναφέρθηκε). Η κύρια διαφορά με την ιδέα που παρουσιάζει η παρούσα διπλωματική είναι πως στην υλοποίηση αυτή γίνεται αξιοποίηση των θετικών δειγμάτων, δηλαδή πληροφοριών από τις κοινότητες του χρήστη. Σε κάθε περίπτωση, πρόκειται για ένα μοντέλο που θα μπορούσε να αξιοποιηθεί από την αγορά, κάτι που πιθανόν να έχει συμβεί (άλλωστε, η δημοσίευση μετράει ήδη 9 χρόνια).



**Σχήμα 3.3:** Η αρχιτεκτονική της προαναφερθείσας δημοσίευσης, που παρουσιάζει ένα προηγμένο σύστημα συστάσεων, το οποίο μάλιστα λειτουργεί σε πραγματικό χρόνο [11]. Αξιοποιεί πληροφορίες από ψηφιακές κοινότητες (κοινωνικών δικτύων) χρηστών ώστε εν συνεχείᾳ να τους προτείνει σχετικά αγαθά & προϊόντα.

Μια άλλη δημοσίευση που πραγματεύεται τα συστήματα συστάσεων πραγματικού χρόνου έρχεται από ερευνητές των πανεπιστημίων Ζεϊζάνγκ της Κίνας & Τμήματος Υπολογιστών του Εθνικού Πανεπιστημίου της Σιγκαπούρης [36]. Ο συγκεκριμένος recommender λειτουργεί εντός κοινωνικών δικτύων όπως το Twitter (γνωστά και ως microblogging systems) και έχει ως στόχο τη δημιουργία προτιμήσεων και προτάσεων με βάση τα “tags” που έχει “γράψει” ή χρησιμοποιήσει ένας χρήστης. Τα tags είναι στοιχεία που χρησιμοποιούνται ευρέως σε δίκτυα όπως το Twitter. Πρόκειται για μικρά κομμάτια κειμένου, συνήθως λέξεις ή φράσεις, που χαρακτηρίζουν ολόκληρα θέματα και κατηγορίες συζητήσεων. Συνήθως, το Twitter παρέχει στους χρήστες σχετικά tags ανάλογα με τη περιοχή και τη χώρα στην οποία βρίσκονται. Όμως, ο recommender του επιστημονικού αυτού άρθρου καταφέρνει να παράξει νέες προτάσεις από τα tags που συνδέονται άμεσα με τους λογαριασμούς των χρηστών, αφού έχει ήδη δημιουργήσει προσωπικούς γράφους πληροφοριών (βασισμένους στα

tags) για τους χρήστες, σε πραγματικό χρόνο. Πρόκειται για μια ενδιαφέρουσα προσέγγιση υπό το πρίσμα του content-based recommendation. Για άλλη μια φορά, γίνεται αποκλειστική χρήση των θετικών δειγμάτων, δηλαδή των tags που έχει χρησιμοποιήσει ο χρήστης.

Τέλος, μια ακόμα αξιοσημείωτη υλοποίηση real-time recommender έγινε από ερευνητές του πανεπιστημίου MIT (ΗΠΑ) και ερευνητικών ινστιτούτων από τη Γαλλία & το Κατάρ [37]. Συγκεκριμένα, πρόκειται για ένα μοντέλο που τροφοδοτείται με σχόλια χρηστών σε άρθρα του διαδικτύου και εν συνεχεία προχωράει στην ταχεία δημιουργία προτάσεων άρθρων που ενδέχεται να ενδιαφέρουν έναν συγκεκριμένο χρήστη, μια συγκεκριμένη χρονική στιγμή. Στόχος του μοντέλου είναι η εύρεση των πιο σχετικών άρθρων για έναν χρήστη που εισέρχεται σε έναν συγκεκριμένο ιστότοπο, σε πραγματικό χρόνο. Αυτό που χρειάζεται να κάνει ο χρήστης, είναι να διαβάσει τουλάχιστον ένα άρθρο του ιστοτόπου, ώστε ο recommender να τροφοδοτηθεί με στοιχεία & λέξεις-κλειδιά από τα σχόλια σε αυτό και να παράξει νέες προτάσεις άρθρων. Είναι μια έξυπνη ιδέα, που συνδυάζει το content-based recommendation με τη δυνατότητα παραγωγής προτιμήσεων σε πραγματικό χρόνο. Όπως ισχύει και με τις προγηούμενες δυο, το σύστημα αυτό δεν χρησιμοποιεί αρνητικά δείγματα, αλλά μόνο θετικά (σχόλια των άρθρων που διαβάζει ο χρήστης). Είναι μια ιδέα που αποδεικνύει πως όλων των ειδών τα δεδομένα μπορεί να είναι χρήσιμα (όχι μόνο το κείμενο ενός άρθρου, αλλά και τα σχόλια αυτού).

Προχωρώντας σε σχετικές υλοποιήσεις που αφορούν συστήματα συστάσεων εξειδικευμένα σε μουσικά στοιχεία, για άλλη μια φορά έχουμε αρκετές σχετικές περιπτώσεις. Ας ξεκινήσουμε με μια προσέγγιση υλοποίησης συστήματος συστάσεων που εκμεταλλεύεται και τα αρνητικά δείγματα. Πρόκειται για μια πρόσφατη δημοσίευση που παρουσιάζει έναν recommender που είναι σε θέση να παράξει μουσικές προτάσεις στον χρήστη, εκμεταλλεύμενος και κομμάτια που εκείνος δεν έχει επιλέξει να ακούσει [38]. Η παραδοχή των συγγραφέων αναφορικά με τα αρνητικά δείγματα είναι η εξής: Αν ο χρήστης δεν επιλέξει να ακούσει κάποια μουσικά κομμάτια, εκείνα αυτομάτως χαρακτηρίζονται ως αρνητικά δείγματα. Επιπρόσθετα, στα θετικά δείγματα συγκαταλέγονται, όχι μόνο τα τραγούδια που έχει ακούσει, αλλά και τα δημοφιλή μιας συγκεκριμένης χρονικής περιόδου, εφόσον πρόκειται για κομμάτια που ακούν πολλοί χρήστες, οπότε και είναι θετικά δείγματα εκείνων. Η αξολόγηση φανέρωσε πως ο recommender ήταν αρκετά αποδοτικός στα αποτελέσματα δημιουργίας ορθών προτιμήσεων στους χρήστες, όμως δεν προσπαθεί να παράξει τις συστάσεις σε πραγματικό χρόνο. Σε κάθε περίπτωση, η χρήση αρνητικών δειγμάτων είναι ενα κοινό χαρακτηριστικό με την ιδέα που παρουσιάζει η παρούσα διπλωματική εργασία.

Άλλη μια δημοσίευση παρουσιάζει τη δημιουργία διαφορετικών playlists από συλλογές τραγουδιών, με στόχο κάποια playlist να αποτελεί και πρόταση σε συγκεκριμένους χρήστες [39]. Με βάση τις μουσικές προτιμήσεις και τα χαρακτηριστικά ενός χρήστη, γίνεται δημιουργία λίστας αναπαραγωγής που είναι πολύ πιθανό να

ταιριάζει στα "ακουστικά θέλω" του. Το αυτοματοποιημένο playlist generation / recommendation υπάρχει εδώ και αρκετά χρόνια, τόσο στην αγορά (πχ. σε υπηρεσίες ροής μουσικής), όσο και σε ερευνητικές προσπάθειες για τη βελτίωση του. Η δημοσίευση πραγματεύεται αυτό το ζήτημα και αναλύει τρόπους για τη ποιοτικότερη δημιουργία τέτοιων συστημάτων συστάσεων. Επιπρόσθετα, η ανάγκη για δημιουργία μουσικών recommenders που θα είναι δυναμικοί και θα προσαρμόζουν το περιεχόμενο των προτάσεων συνεχώς, έχει τονιστεί και σχετικά πρόσφατα [40]. Η έρευνα γύρω από τα συστήματα συστάσεων με ειδίκευση στη μουσική πρέπει να συνεχιστεί. Όπως τονίζεται και μέσω της παρούσας διπλωματικής εργασίας, οι μελλοντικοί recommenders πρέπει να είναι σε θέση να προσαρμόζονται συνεχώς, ανάλογα με τη διάθεση και την όρεξη των χρηστών. Η συγκεκριμένη θεωρητική δημοσίευση ώθησε ακόμα περισσότερο τη θέληση συγγραφής της παρούσας εργασίας, εφόσον προσπαθεί να κάνει ακριβώς αυτό: Να δοκιμάσει τις δυνατότητες των μουσικών συστημάτων συστάσεων ώστε να προσαρμόζονται κάθε φορά διαφορετικά, και τη μελέτη αυτής της κατάστασης σε πραγματικό χρόνο.



Σχήμα 3.4: *Tι μπορεί να βρίσκεται στο επίκεντρο των μουσικών συναισθημάτων & προτιμήσεων ενός χρήστη κάθε φορά;* Αυτό είναι μια πραγματική πρόκληση που καλούνται να αντιμετωπίσουν τα μελλοντικά συστήματα συστάσεων [12]. Η ιδέα της παρούσας διπλωματικής εργασίας πειραματίζεται με το προαναφερθέν ερώτημα.

Αναφορικά με τα συναισθήματα και τη διάθεση των χρηστών, γνωρίζουμε πολύ καλά πως εκείνα συνδέονται άρρηκτα με τα μουσικά / μελωδικά ερεθίσματα. Όπως έχει ήδη αναφερθεί, το σύστημα που παρουσιάζεται μέσω τούτης της διπλωματικής εργασίας καλείται να προσαρμόζει την ακολουθία των κομματιών μιας λίστας αναπαραγωγής με βάση την κατηγορία μουσικών ειδών που ταιριάζουν στη διάθεση

του χρήστη μια συγκεκριμένη χρονική στιγμή. Όμως, έχουν γίνει αρκετές έρευνες σχετικά με τη συσχέτιση μουσικής και συναισθημάτων / διάθεσης. Μια τέτοια δημοσίευση εξετάζει τη σύνδεση στοιχείων του χαρακτήρα ατόμων με τις μουσικές τους προτιμήσεις [41], κάτι που θα μπορούσε να οδηγήσει στη δημιουργία νέων, πιο αποδοτικών recommenders. Επίσης, μια άλλη έρευνα επιχείρησε να συνδυάσει τις συσχετίσεις συναισθημάτων - μουσικής με συγκεκριμένες τοποθεσίες, δημιουργώντας έτσι ακόμα μεγαλύτερες ομάδες συνδυασμών [42]. Ένα άτομο θα μπορούσε να διακατέχεται από συγκεκριμένα συναισθήματα σε κάποια τοποθεσία, στην οποία θα ταιριάζει μουσική που "αντιστοιχεί" στα συναισθήματά του. Η έρευνα αυτή θα μπορούσε να οφελήσει τα συστήματα συστάσεων που λαμβάνουν υπόψιν τους και τη τοποθεσία των χρηστών, υλοποιώντας έτσι έναν μουσικό recommender με... location services.

Μια επίσης ενδιαφέρουσα δημοσίευση αφορά την έρευνα γύρω από το είδος των δεδομένων που χρειάζονται για ορθή και αποδοτική υλοποίηση μουσικών recommenders [43]. Συγκεκριμένα, το άρθρο ερευνά μετα-δεδομένα μουσικών κομματιών και την ενδεχόμενη επιρροή που θα είχαν εκείνα σε συστήματα συστάσεων. Μάλιστα, ένα από τα βασικά μετα-δεδομένα των τραγουδιών που μπορούν να συμβάλλουν στη βελτίωση των recommenders είναι το είδος τους. Το συμπέρασμα αυτό του άρθρου ενισχύει ακόμα περισσότερο την κεντρική ιδέα της παρούσας διπλωματικής εργασίας, η οποία βασίζεται ιδιαίτερα στα μουσικά είδη. Όπως αναφέρθηκε και προηγουμένως, είναι ευκολότερο να γίνεται διαχείριση κομματιών με βάση το είδος τους, εφόσον είναι πιο κοντά στη "ρίζα" του δέντρου της μουσικής μελωδίας. Θα μπορούσαμε να φανταστούμε ενα μουσικό είδος ως κλαδί και σχετικά του κομμάτια ως αμέτρητα παρακλάδια ή φύλλα. Αν θέλουμε να απομονώσουμε ένα είδος, είναι ευκολότερο να κόψουμε ολόκληρο το κλαδί, παρά τα παρακλάδια / φύλλα ένα-ένα. Βασιζόμενοι (και) στη παραπάνω δημοσίευση, αποφασίσαμε να αναλύσουμε τα μετα-δεδομένα συνόλων μουσικών κομματιών και να εξετάσουμε πως μπορούν αυτά να συμβάλλουν στη κατασκευή ενός συστήματος συστάσεων. Παρακάτω θα αναλυθεί επαρκώς το σύνολο δεδομένων στο οποίο καταλήξαμε, μαζί με τα χαρακτηριστικά του.

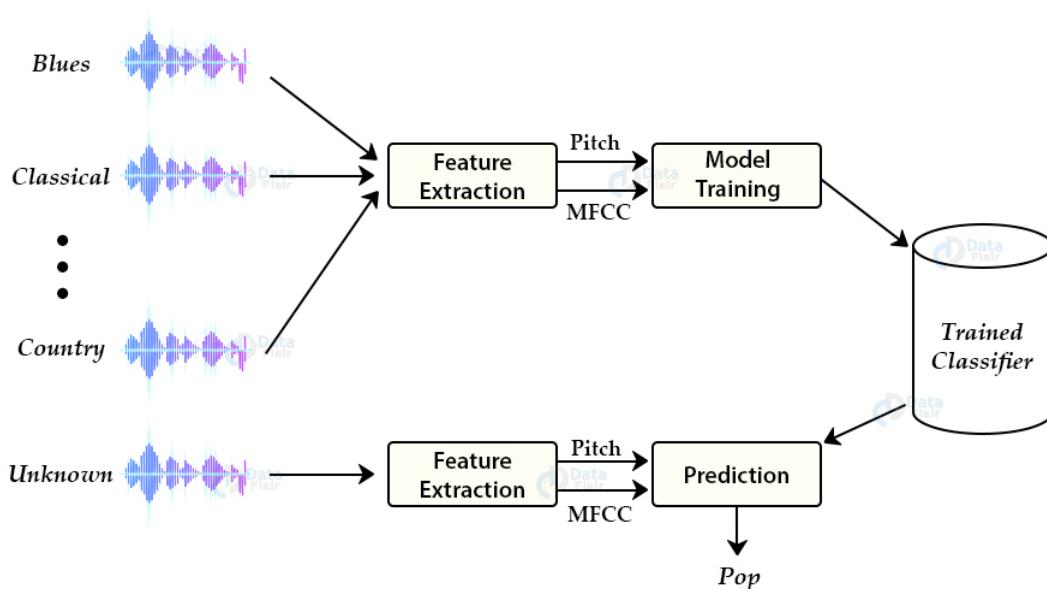
Αντιλαμβανόμαστε λοιπόν πως υπάρχουν πολλές παρόμοιες υλοποιήσεις που αφορούν recommenders πραγματικού χρόνου, αξιοποίησης αρνητικών δειγμάτων και ειδικευόμενους σε μουσικά σύνολα δεδομένων. Η κυριότερη προσέγγιση που χρησιμοποίησαν οι δημοσιεύσεις στις οποίες γίνεται αναφορά, είναι εκείνη του content-based recommendation με ενδεχόμενες παραλλαγές. Ας δούμε τώρα σχετικές ιδέες που έχουν δημοσιευτεί και αφορύν τη χρήση μοντέλων μηχανικής μάθησης. Η αναζήτηση αφορούσε κυρίως τους τρόπους με τους οποίους μπορούν να αξιοποιηθούν τα μουσικά δεδομένα από ML αλγορίθμους και τελικά τι αποτελέσματα μπορούν εκείνοι να παράξουν.

### 3.2.2 Υλοποιήσεις Μηχανικής Μάθησης

Ας δούμε το σκέλος των υλοποιήσεων με χρήση **τεχνικών μηχανικής μάθησης**. Ο λόγος μελέτης τους στη διπλωματική μας εργασία είναι το genre recognition και το κατά πόσο μπορεί να οφελήσει η δημιουργία ενός υβριδικού recommender, του οποίου ένα σκέλος θα επιχειρεί ακριβώς αυτό μέσω μοντέλων machine learning. Μπορούμε να ξεκινήσουμε εξετάζοντας ιδέες & δημοσιεύσεις που αφορούν τον συνδυασμό μοντέλων Machine Learning με εκείνο των αρνητικών δειγμάτων. Παρόλο που δεν βρέθηκαν έρευνες σχετικές με μηχανική μάθηση, μουσικά δεδομένα και αρνητικά δείγματα αυτών, υπάρχουν όμως υλοποιήσεις με άλλο είδος δεδομένων. Συγκεκριμένα, μια δημοσίευση πραγματεύεται την εκπαίδευση αλγορίθμων machine learning και πρόβλεψη πρωτεϊνικών αλληλεπιδράσεων, με χρήση αρνητικών δεδομένων [44]. Στη περίπτωση αυτή, τα αρνητικά δεδομένα ορίζονται ως τα δείγματα πρωτεϊνικών στοιχείων που δεν αλληλεπιδρούν μεταξύ τους, δηλαδή το αντίθετο από το επιθυμητό αποτέλεσμα πρόβλεψης των υλοποιηθέντων ταξινομητών. Πρόκειται για μια έρευνα που πειραματίζεται τόσο με τη μηχανική μάθηση, όσο και με τα αρνητικά δείγματα. Ένα άλλο επιστημονικό άρθρο παρουσιάζει τη δυνατότητα υλοποίησης ενός Παραγωγικού Αντιπαραθετικού δικτύου (Generative Adversarial Network, γνωστό με τη συντομογραφία του, GAN), το οποίο θα λειτουργεί με τα αρνητικά δείγματα ενός συνόλου δεδομένων [45], στα πλαίσια της Αναπαράστασης Γνώσης. Η δημιουργία αρνητικών (“αντίθετων”) δεδομένων και η ενσωμάτωση τους στο σύστημα έχει ως σκοπό τη βελτίωση της ποιότητας λειτουργίας του GAN δικτύου. Είναι μια εξειδικευμένη έρευνα και απευθύνεται σε άτομα που καταπιάνονται σχεδόν αποκλειστικά με τη μηχανική μάθηση.

Μια ακόμα δημοσίευση αφορά την εκπαίδευση ταξινομητών μηχανικής μάθησης με σκοπό την ορθή κατηγοριοποίηση προδρόμων microRNA [46]. Τα αρνητικά δείγματα στη προκειμένη περίπτωση αποτελούν όσα στοιχεία δεν είναι πρόδρομοι microRNA. Αξίζει να σημειωθεί πως τα machine learning μοντέλα του άρθρου δεν εκπαιδεύτηκαν αποκλειστικά και μόνο με αρνητικά δείγματα, αλλά με συνδυασμό. Αυτό ενισχύει ακόμα περισσότερο την παραδοχή που κάναμε παραπάνω, πως η αξιοποίηση τόσο των θετικών, όσο και των αρνητικών δειγμάτων μπορεί να αποφέρει τα καλύτερα δυνατά αποτελέσματα σε ένα σύστημα. Σε κάθε περίπτωση, έχοντας στο νου μας τις τρεις παραπάνω δημοσιεύσεις, αντιλαμβανόμαστε πως είναι όντως εφικτή η αξιοποίηση όλων των ειδών ενός συνόλου δεδομένων για έναν classifier μηχανικής μάθησης. Πέραν όμως αυτού, υπάρχουν διαθέσιμες πολλές επιστημονικές έρευνες που πραγματεύονται την αξιοποίηση του machine learning σε μουσικά σύνολα δεδομένων. Κύριο μοτίβο των περισσότερων περιπτώσεων είναι η ανάλυση των μουσικών χαρακτηριστικών για τα κομμάτια ενός συνόλου, η αξιοποίηση των πληροφοριών αυτών μέσω μοντέλων μηχανικής μάθησης και η τελική πρόβλεψη επιθυμητών στοιχείων. Στις περισσότερες περιπτώσεις, ο ιδανικός στόχος είναι η πρόβλεψη του

μουσικού είδους για τα κομμάτια ενός συνόλου δεδομένων. Επιστρέφουμε δηλαδή στο genre recognition, ένα θέμα γύρω από το οποίο έχουν γίνει αμέτρητες δοκιμές με χρήση machine learning ταξινομητών. Πάνω σε αυτό το σκέλος, απαιτείται σχολαστική έρευνα και ανάλυση των διθέντων μουσικών δεδομένων, ώστε να αποφασιστεί ποια χαρακτηριστικά οδηγούν καλύτερα στη πρόβλεψη του είδους. Άλλωστε, τα διαθέσιμα μουσικά σύνολα δεδομένων διαφέρουν ως προς τις πληροφορίες που παρέχουν για κάθε κομμάτι τους. Ποιες από τις παρεχόμενες πληροφορίες μπορούν να αξιοποιηθούν από μοντέλα μηχανικής μάθησης και να οδηγήσουν σε πρόβλεψη του μουσικού είδους; Αυτό είναι και ένα ερώτημα που αφορά άμεσα τη παρούσα διπλωματική εργασία.



**Σχήμα 3.5:** Η χρήση μοντέλων μηχανικής μάθησης με σκοπό τη πρόβλεψη μουσικών ειδών και κατηγοριοποίηση κομματιών με βάση αυτά, είναι ενα ιδιαίτερα ευδιαφέρον σκέλος έρευνας [13]. Έχει ουσιαστικό νόημα η χρήση του machine learning ή πρόκειται για ένα πρόβλημα που χρήζει άλλης λύσης;

Διερεύνηση των μουσικών χαρακτηριστικών που οδηγούν σε ταυτοποίηση του είδους ενός κομματιού έχει γίνει και παρουσιάζεται σε διάφορες δημοσιεύσεις, όπως προαναφέρθηκε. Μάλιστα, η έρευνα γύρω από αυτό το κομμάτι έχει ζεκινήσει εδώ και πολλά χρόνια, όπως αποδεικνύεται από δημοσιεύσεις των αρχών του 21ου αιώνα [47] [48]. Τα αρχικά αποτελέσματα έκαναν ανάλυση πάνω σε μουσικά σύνολα δεδομένων με ακουστικά αρχεία, ενώ τα συμπεράσματα τους μαρτυρούσαν την ανάγκη περαιτέρω ανάπτυξης των ταξινομητών που θα μπορούσαν να κατηγοριοποιήσουν ορθά τα μουσικά είδη σε ένα σύνολο. Φυσικά, η πάροδος των ετών έφερε νέες δημοσιεύσεις, με σημαντικά αξιόλογα συμπεράσματα. Μια εξ αυτών χρησιμοποιεί ακουστικό υλικό από τα κομμάτια ενός μουσικού συνόλου δεδομένων, με σκοπό την ανάλυση του, την εκπαίδευση ταξινομητών και τελικά τη κατηγοριοποίηση ανά

είδος [49]. Το συλλεχθέν σύνολο δεδομένων αποτελούνταν από μουσικά κομμάτια επτά (7) διαφορετικών ειδών, σχηματίζοντας αυτομάτως επτά "labels" ταξινόμησης. Παρά το επαρκές του συνόλου (40 χιλιάδες κομμάτια) και την προσεκτική δοκιμή & παραμετροποίηση των μοντέλων μηχανικής μάθησης, η ακρίβεια κατηγοριοποίησης κυμάνθηκε μεταξύ του 43% και 65%. Τα αποτελέσματα δεν μπορούν να θεωρηθούν ιδιαίτερα ενθαρρυντικά, αν συλλογιστεί κανείς πως, στη καλύτερη περίπτωση, θα κατηγοριοποιούνται ορθά κάτι παραπάνω από τα μισά στοιχεία. Οι λόγοι για τη μειωμένη ακρίβεια των μοντέλων μπορεί να ποικίλουν. Ένα βασικό χαρακτηριστικό που αξίζει να έχει κανείς υπόψιν είναι το μέγεθος των διαφορετικών labels. Οι επτά επιλογές αποτελούν ήδη υψηλό αριθμό, κάτι που είναι πολύ πιθανό να οδηγήσει σε χαμηλότερα ποσοστά ακρίβειας. Αν διακρατεί κανείς στο νου του και αυτό τον παράγοντα, τότε οι επιδόσεις των μοντέλων που παρουσιάστηκαν σε αυτή τη δημοσίευση δεν είναι τόσο άσχημες. Όμως, τα αποτελέσματα αυτά ενδέχεται να οδηγούν και σε ένα άλλο συμπέρασμα. Αν η κατηγοριοποίηση επτά διαφορετικών μουσικών ειδών δεν φαίνεται να στέφεται με επιτυχία λόγω του πλήθους τους, τι θα γίνει σε περιπτώσεις όπου καλούμαστε να ταξινομήσουμε περισσότερα; Άλλωστε, υπάρχουν χιλιάδες διαθέσιμα μουσικά είδη. Είναι πιθανό, η φύση του προβλήματος (δηλαδή η ποικιλία των μουδικών ειδών) να καθιστά ιδιαίτερα δύσκολο το έργο των μοντέλων μηχανικής μάθησης [50];

Άλλες σχετικά πρόσφατες υλοποιήσεις που αξιοποιούν τα ακουστικά δεδομένα από μουσικά σύνολα κομματιών, αποφέρουν καλύτερα αποτελέσματα. Μια επιστημονική έρευνα κατάφερε να πετύχει ακρίβεια κατηγοριοποίησης δέκα διαφορετικών μουσικών ειδών που ξεπερνούσε το 80 τοις εκατό [51]. Το σύνολο δεδομένων ήταν μόλις χίλια τραγούδια, ενώ κάθε ενα εκ των δέκα ειδών εκπροσωπούνταν από 100 κομμάτια. Η χρήση μοντέλου SVM απέσπασε ακρίβεια κοντά στο 85%, ενώ ο συνδυασμός SVM & LSTM κατάφερε να αγγίζει το ποσοστό του 90%. Τα ακουστικά δεδομένα επεξεργάστηκαν επιτυχώς και τελικά κατάφεραν να τροφοδοτήσουν τους ταξινομητές με ορθό και αποδοτικό τρόπο. Φαίνεται λοιπόν πως είναι εφικτή η δυνατότητα υψηλής ακρίβειας κατηγοριοποίησης μουσικών ειδών, ακόμα και όταν εκείνα είναι πυκνά στον αριθμό. Επίσης, σημαντικό ρόλο φαίνεται να παίζει και η ορθή επεξεργασία των ακουστικών δεδομένων του συνόλου. Επιπρόσθετα, μια έρευνα ερευνητών από το πανεπιστήμιο Ποντίου Φάθρα, απέδειξε πως η πλήρης αξιοποίηση όλων των πιθανών πληροφοριών για ένα μουσικό σύνολο δεδομένων μπορεί να έχει εξαιρετικά αποτελέσματα [52]. Συγκεκριμένα, η ερευνητική ομάδα συγκέντρωσε δεδομένα αλφαριθμητικά, ηχητικά και οπτικά δεδομένα από μουσικά άλμπουμ, επεξεργάζοντας τα και αξιοποιώντας τις παραχθείσες πληροφορίες. Τα τελικά αποτελέσματα ήταν ιδιαίτερα ενθαρρυντικά, αφού η ακρίβεια των εκπαιδευμένων μοντέλων μηχανικής μάθησης κυμαίνονταν στα όρια του 90 τοις εκατό. Αξίζει να έχουμε στο νου μας πως η λύση που προτάθηκε είναι αρκετά εξειδικευμένη, αλλά σε κάθε περίπτωση πλήρως αποδοτική.

Ας επιχειρήσουμε όμως να εξετάσουμε το πρόβλημα από μια διαφορετική οπτική γωνία. Ας δούμε αν είναι εφικτή η αποδοτική ταξινόμηση των ειδών σε ένα μουσικό σύνολο, αφήνοντας στην άκρη τα ακουστικά δεδομένα (ή ακόμα και τα... οπτικά όπως είδαμε στη προηγούμενη δημοσίευση) και αξιοποιώντας μόνο αλφαριθμητικές πληροφορίες για κάθε κομμάτι. Μπορούμε να έχουμε ίδια ή και παρόμοια αποτελέσματα χωρίς την ανάλυση & επεξεργασία των ακουστικών αρχείων; Συγκεκριμένα, εφόσον η παρούσα διπλωματική εργασία βασίζεται πάνω στα δεδομένα που παρέχονται από το Spotify (όπως προαναφέρθηκε), θα μπορόσαμε να εκπαιδεύσουμε ταξινομητές με σύνολα της δημοφιλούς υπηρεσίας ροής μουσικής; Αρχικά, να σημειώσουμε πως έχουν γίνει (και δημοσιευτεί) κάποιες έρευνες που πραγματεύονται την αναγνώριση μουσικών ειδών για σύνολα δεδομένων που προέρχονται από το Million Song Dataset. Πρόκειται για έναν ιστότοπο που παρέχει δωρεάν πρόσθιαση σε ένα τεράστιο σύνολο δεδομένων από μουσικά κομμάτια, με τα οποία είναι κανείς ελέυθερος να πειραματιστεί. Είναι ένα ιδιαίτερα δημοφιλές σύνολο δεδομένων, αφού παρέχει ποικιλία χαρακτηριστικών για κάθε ένα μουσικό κομμάτι. Όπως τονίστηκε προηγουμένως, στη παρούσα διπλωματική εργασία επιλέχθηκαν δεδομένα από το Spotify έναντι του Million Song Dataset. Ο λόγος αφορά μέρος του στόχου της εργασίας που αφορά υλοποίηση σε σενάρια πραγματικού κόσμου. Αν μελλοντικά το Spotify αποφασίσει να ενσωματώσει τεχνικές αναγνώρισης και κατηγοριοποίησης μουσικών ειδών εντός των λιστών αναπαραγωγής, θα το κάνει με δικά του δεδομένα και όχι με datasets άλλης πηγής. Άλλωστε, κατηγοριοποίηση θα γίνεται στα μουσικά κομμάτια λιστών αναπαραγωγής του ίδιου του Spotify. Με λίγα λόγια, δεν μπορούν τα train και test sets να είναι διαφορετικής πηγής, πόσο μάλλον όταν φέρουν κάποιες διαφορές μεταξύ τους. Εκεί βασίζεται και η απόφαση για αποκλειστική χρήση δεδομένων από τη γνωστή streaming υπηρεσία.

Γιατί όμως γίνεται μνεία στο Milion Song Dataset; Ο λόγος είναι εκείνος που αναφέραμε στη μέση της προηγούμενης παραγράφου. Υπάρχουν σχετικές υλοποιήσεις αναγνώρισης μουσικών ειδών μόνο με βάση τα αλφαριθμητικά στοιχεία τους. Εφόσον τα δεδομένα μεταξύ Million Song Dataset & Spotify μοιράζονται αρκετές ομοιότητες, θα εξετάσουμε εν τάχει δυο δημοσιεύσεις που επιχείρησαν ταξινόμηση μουσικών ειδών με χρήση συνόλου δεδομένων από τη πρώτη πηγή. Πρόκειται για μια έρευνα [53] καθώς και μια μεταπτυχιακή διπλωματική εργασία [54] που φέρουν την ίδια προσέγγιση και προβαίνουν σε ταξινόμηση ενός μουσικού συνόλου δεδομένων, ανάλογα με το είδος τους κάθε κομματιού. Στη περίπτωση της έρευνας, το σύνολο των διαφορετικών ειδών εντός του συνόλου είναι δέκα, αν και μερικά αποτελούν συνδυασμούς κατηγοριών (πχ. rock και indie μαζί). Απεναντίας, ο μεταπτυχιακός φοιτητής αποφάσισε να "απλώσει" ακόμα περισσότερο το εύρος των μουσικών ειδών, έχοντας 18 εξ αυτών στο σύνολο δεδομένων που διάλεξε. Και οι δυο περιπτώσεις εκμεταλλεύονται το περιεχόμενο των στίχων του κάθε κομματιού, κάτι που συνδυάζουν με τα μουσικά χαρακτηριστικά τα οποία το συνοδεύουν. Στη περίπτωση της

δημοσίευσης, τα τελικά αποτελέσματα ακρίβειας των classifiers κυμαίνονταν λίγο κάτω από το 40 τοις εκατό, ενώ στη περίπτωση της μεταπτυχιακής διπλωματικής εργασίας, οι επιδόσεις του machine learning μοντέλου άγγιξαν το 41%. Καμία από τις δύο υλοποιήσεις δεν κατάφεραν καν να φτάσουν κοντά στα επίπεδα ακρίβειας των προηγούμενων δημοσιεύσεων. Ενδεχομένως, η χρήση των μουσικών χαρακτηριστικών και των στίχων δεν είναι αρκετή για να εκπαιδεύσει επαρκώς τους machine learning ταξινομητές.

Το συμπέρασμα που προκύπτει μοιάζει να μην είναι “καλός οιωνός” για τις δοκιμές του συστήματος της παρούσας διπλωματικής εργασίας. Φαίνεται πως η αναγνωρισή μουσικών ειδών με χρήση ταξινομητών μηχανικής μάθησης είναι όντως μια ιδιαίτερη υπόθεση και για την επιτυχία αυτής υπάρχουν πολλοί παράγοντες που πρέπει να κυλήσουν σωστά. Μια εκτεταμένη επεξεργασία ηχητικών αρχείων μουσικών συνόλων δεδομένων και μια επιλογή σωστού ταξινομητή, μπορεί να οδηγήσουν σε επιθυμητά αποτελέσματα. Αυτό φαίνεται να ενισχύεται ακόμα περισσότερο στη περίπτωση χρήσης αλφαριθμητικών και οπτικών δεδομένων, πέραν των ηχητικών. Αν πρόκειται όμως μόνο για αλφαριθμητικά δεδομένα (κάτι στο οποίο δεν υπάρχει μεγάλη ποικιλία δημοσιεύσεων), τότε ίσως τα μοντέλα μηχανικής μάθησης να μην είναι σε θέση να παράξουν αξιόλογα αποτελέσματα. Αυτό που θα επιχειρήσουμε εμείς, είναι να δούμε πόσο εφικτή είναι η ταξινόμηση μουσικών ειδών για κομμάτια ενός συνόλου δεδομένων, με βάση τις πληροφορίες και τα χαρακτηριστικά που παίρνουμε από το Spotify. Παρόμοιες δοκιμές σε βασικό επίπεδο έχουν λάβει χώρα μόνο από ιστοτόπους όπως το medium.com, οι οποίοι περισσότερο εξηγούν στους χρήστες πως να το κάνουν μόνοι τους και να πειραματιστούν με το συγκεκριμένο πρόβλημα. Μια σχετικά πρόσφατη δημοσίευση επιχείρησε να κάνει ακριβώς αυτό, χρησιμοποιώντας τα δημοφιλέστερα τραγούδια της περιόδου στο Spotify και χρησιμοποιώντας τα δοθέντα μουσικά χαρακτηριστικά τους [55]. Με ένα εύρος δέκα διαφορετικών μουσικών ειδών, το μοντέλο μηχανικής μάθησης που ανέπτυξαν οι ερευνητές της δημοσίευσης “έπιασε” μόλις 46 τοις εκατό ακρίβεια. Όπως η επιστημονική ομάδα της δημοσίευσης, έτσι και εμείς θα εξετάσουμε τα δοθέντα χαρακτηριστικά των μουσικών κομματιών. Έπειτα, θα ενσωματώσουμε διαφορετικούς classifiers μηχανικής μάθησης και θα δούμε αν έχει νόημα η ενσωμάτωση κάποιου εξ αυτών στο τελικό σύστημα συστάσεων μας. Επαρκούν τα αλφαριθμητικά μουσικά δεδομένα του Spotify για πετυχημένη εκπαίδευση machine learning μοντέλων; Θα το μάθουμε παρακάτω...

## Μέρος ΙΙ

### Πρακτικό Μέρος

---



## Κεφάλαιο 4

# Ανάλυση Υλοποίησης

---

Στο κεφάλαιο αυτό παρουσιάζεται η τεχνική ανάλυση του συστήματος που πραγματεύεται η διπλωματική εργασία. Αρχικά, αναλύεται το σύνολο δεδομένων, ενώ εν συνεχεία παρουσιάζονται οι ταξινομητές μηχανικής μάθησης που δοκιμάστηκαν, καθώς και λεπτομέρειες για τη διαμόρφωση του content-based συστήματος συστάσεων που αναπτύχθηκε.

### 4.1 Το Σύνολο Δεδομένων

#### 4.1.1 Γνώστα Σύνολα Διαθέσιμα στο Διαδίκτυο

Όπως έχει γίνει ήδη κατανοητό, η παρούσα διπλωματική εργασία θέλησε να εξετάσει τη δημουργία ενός συστήματος συστάσεων μέσα από δεδομένα που παρέχονται από το Spotify. Πριν περάσουμε στην ανάλυση αυτών, αξίζει να γίνει μνεία σε δυο άλλα σύνολα δεδομένων, τα οποία είναι ιδιαίτερα δημοφιλή στο διαδίκτυο και χρησιμοποιούνται ευρέως για ανάλυση και εξαγωγή συμπερασμάτων. Το πρώτο (και δημοφιλέστερο) δεν είναι άλλο από το Million Song Dataset [56]. Πρόκειται για ένα από τα μεγαλύτερα μουσικά σύνολα που είναι ελέυθερα και διαθέσιμα μέσω διαδικτύου. Επίσης, πάνω σε αυτό έχουν διεξαχθεί αμέτρητες έρευνες, αντικείμενα επιστημονικών δημοσιεύσεων. Αποτελείται από εκατοντάδες χιλιάδες διαθέσιμα μουσικά κομμάτια (ή “ένα εκατομμύριο” όπως ξεκαθαρίζει ο ιστότοπος), παρέχει αρκετά στοιχεία για κάθε ένα εξ αυτών, κάτι που το καθιστά ιδιαίτερα χρήσιμο για πολλές μελέτες. Η δημουργία του dataset αποτελεί ένα project των The Echo Nest και LabROSA. Το πρώτο ήταν μια πλατφόρμα μουσικής ευφυΐας και δεδομένων για προγραμματιστές και ενδιαφερόμενες εταιρείες. Η εταιρεία απορροφήθηκε από το Spotify, το 2014. Το δεύτερο (LabROSA), γνωστό και με τη πλήρη ονομασία του “Laboratory for the Recognition and Organization of Speech and Audio”, είναι ένα εργαστηριακό / ερευνητικό τμήμα του πανεπιστημίου Columbia των ΗΠΑ. Μέχρι και σήμερα, το Million Song Dataset αποτελεί έναν από τους δημοφιλέστερους “προορισμούς” για όσους θέλουν να πειραματιστούν με την ανάλυση μουσικών συνόλων.

Στην πλήρη έκδοση του, το MSD (αρχικά του Million Song Dataset) κυμαίνεται στα 280 GB, μια πληροφορία που από μόνη της αρκεί για να εξηγήσει των όγκο των διαθέσιμων πληροφοριών.

dataset <sup>1</sup>	#clips	#artists	year	audio
RWC [12]	465	-	2001	yes
CAL500 [45]	500	500	2007	yes
Ballroom [13]	698	-	2004	yes
GTZAN [46]	1,000	~ 300	2002	yes
MusiClef [36]	1,355	218	2012	yes
Artist20 [7]	1,413	20	2007	yes
ISMIR2004	1,458	-	2004	yes
Homburg [15]	1,886	1,463	2005	yes
103-Artists [30]	2,445	103	2005	yes
Unique [41]	3,115	3,115	2010	yes
1517-Artists [40]	3,180	1,517	2008	yes
LMD [42]	3,227	-	2007	no
EBallroom [23]	4,180	-	2016	no <sup>2</sup>
USPOP [1]	8,752	400	2003	no
CAL10k [44]	10,271	4,597	2010	no
MagnaTagATune [20]	25,863 <sup>3</sup>	230	2009	yes <sup>4</sup>
Codaich [28]	26,420	1,941	2006	no
<b>FMA</b>	<b>106,574</b>	<b>16,341</b>	<b>2017</b>	<b>yes</b>
OMRAS2 [24]	152,410	6,938	2009	no
MSD [3]	1,000,000	44,745	2011	no <sup>2</sup>
AudioSet [10]	2,084,320	-	2017	no <sup>2</sup>
AcousticBrainz [32]	2,524,739 <sup>5</sup>	-	2017	no

<sup>1</sup> Names are clickable links to datasets' homepage.

<sup>2</sup> Audio not directly available, can be downloaded from [ballroomdancers.com](http://ballroomdancers.com), [7digital.com](http://7digital.com), [youtube.com](http://youtube.com).

<sup>3</sup> The 25,863 clips are cut from 5,405 songs.

<sup>4</sup> Low quality 16 kHz, 32 kbit/s, mono mp3.

<sup>5</sup> As of 2017-07-14, of which a subset has been linked to genre labels for the MediaEval 2017 genre task.

Σχήμα 4.1: Μια λίστα με τα δημοφιλέστερα μουσικά σύνολα δεδομένων, διαδέσιμα μέσω διαδικτύου [14]. Το dataset του Spotify δεν περιλαμβάνεται στη λίστα, πιθανότατα λόγω του γεγονότος πως το δημιουργεί κανείς μόνος του, δυναμικά, μέσω του παρεχόμενου API.

Ένα άλλο γνωστό σύνολο δεδομένων είναι το Free Music Archive (FMA), ένα (επίσης) ανοιχτό και ελεύθερο dataset κατάλληλο για ανάλυση και δοκιμή εφαρμογών μηχανικής μάθησης [57]. Έχει δημιουργηθεί από ερευνητές των πανεπιστημίων EPFL της Ελβετίας και NTU της Σιγκαπούρης. Αποτελείται συνολικά από 106 χιλιάδες μουσικά κομμάτια, ενώ περιέχει πάνω από 16 χιλιάδες καλλιτέχνες. Αξίζει να σημειωθεί πως το FMA ενσωματώνει και ηχητικά αρχεία, κάτι που το Million Song Dataset παρέχει μέσω τρίτων ιστοτόπων. Αυτό είναι και ένα εκ των χαρακτηριστικών που καθιστούν το FMA προτιμητέο από πολλούς, παρά την αισθητή διαφορά στον αριθμό των μουσικών κομματιών (100 χιλιάδες απέναντι σε ένα εκατομμύριο). Βέβαια, όλα έρχονται με... κόστος. Το μέγεθος του πλήρους FMA Dataset ξεπερνάει τα 900 GB. Όμως, αποτελεί μια από τις καλύτερες διαθέσιμες επιλογές. Αυτά τα

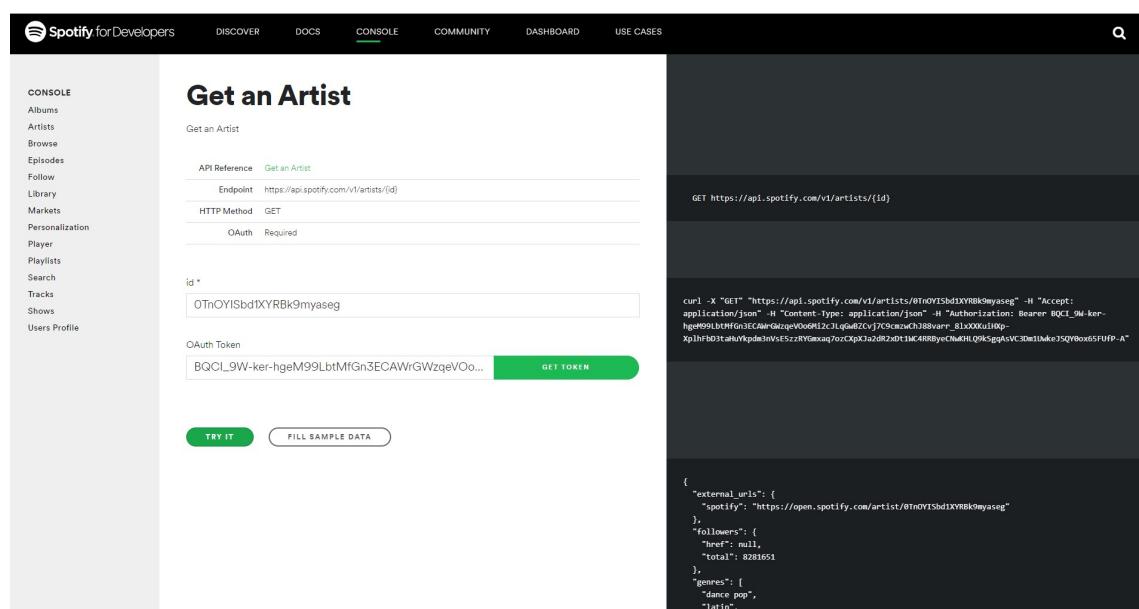
δυο σύνολα δεδομένων αποτέλεσαν κομμάτια αρχικής μελέτης της παρούσας διπλωματικής εργασίας, κυρίως για ερευνητικούς λόγους και θέληση αποκόμισης γνώσης γύρω από διάφορα μουσικά datasets που κυκλοφορούν στο διαδίκτυο. Όπως είναι λογικό, υπάρχουν και άλλα διαθέσιμα σύνολα δεδομένων, με τα οποία όμως δεν έγινε πειραματισμός εντός της διπλωματικής (φαίνονται στην εικόνα παραπάνω). Έχοντας κάνει αναφορά σε δυο εκ των δημοφιλέστερων όμως, δεν έχουμε παρά να ξεκινήσουμε την ανάλυση του συνόλου δεδομένων που παρέχεται από το Spotify σε όσους επιθυμούν. Αυτό γίνεται μέσω της διεπαφής προγραμματισμού εφαρμογών (API) που παρέχεται από την ίδια την υπηρεσία ροής μουσικής. Για να μπορεί κανείς να αποκτήσει πρόσθιαση, πρέπει να συνδεθεί ως developer μέσω του λογαριασμού που διατηρεί στη πλατφόρμα. Ας το εξετάσουμε εν βάθος.

### 4.1.2 Spotify Dataset

To Spotify Dataset. Πρόκειται για το κεντρικό αντικείμενο μελέτης της παρούσας διπλωματικής εργασίας, πάνω στο οποίο ερευνάται η εφαρμογή αλγορίθμων μηχανικής μάθησης, καθώς και προσεγγίσεων συστημάτων συστάσεων. Βέβαια, υπάρχει κάτι που πρέπει να τονιστεί. Το σύνολο δεδομένων αυτό δεν είναι διαθέσιμο ως ένα έτοιμο "πακέτο" που παρέχεται από τη γνωτή υπηρεσία ροής μουσικής. Δημιουργείται δυναμικά από τους χρήστες, μέσω του Spotify Web API. Βασισμένο σε υπηρεσίες REST, το Web API παρέχει μια σειρά από endpoints, τα οποία καλούνται από τους χρήστες ανάλογα με το είδος των δεδομένων που επιθυμούν να αποκτήσουν. Το είδος των δεδομένων που επιστρέφονται είναι σε μορφή JSON. Οι χρήστες μπορούν να αποκτήσουν πρόσθιαση σε πληροφορίες & μετα-δεδομένα που αφορούν μουσικούς καλλιτέχνες, άλμπουμ και τραγούδια, απευθείας από τους επίσημους καταλόγους του Spotify. Πιο συγκεκριμένα, εκτός του γεγονότος πως μπορεί κανείς να λάβει λεπτομερείς πληροφορίες για άλμπουμ / καλλιτέχνες / κομμάτια της επιλογής του, μπορεί επίσης να περιηγηθεί σε λίστες αναπαραγωγής που είναι δημοσιευμένες στο spotify, αλλά και σε προσωπικές του. Μπορεί να αποκτήσει πληροφορίες που αφορούν τον προσωπικό του λογαριασμό, να δημιουργήσει νέες λίστες αναπαραγωγής, να προσφαφαιρέσει μουσικά κομμάτια από αυτές, να λάβει τα κορυφαία μουσικά κομμάτια μιας περιόδου ανάλογα με την αγορά (χώρα) που τον ενδιαφέρει, να ακολουθήσει (follow) ή να διακόψει την ακολουθία (unfollow) σε συγκεκριμένους καλλιτέχνες, χρήστες ή λίστες, αλλά και πολλά άλλα πράγματα. Εν ολίγοις, οι επιλογές που έχει στη διάθεση του ένας χρήστης μέσω του Spotify Web API είναι αρκετές.

Αξίζει να σημειωθεί πως υπάρχουν και Kit Ανάπτυξης Λογισμικού (Software Development Kits - SDKs) που παρέχονται από το Spotify και δίνουν τη δυνατότητα σύνδεσης με τις δυνατότητες και τα δεδομένα αυτού, μέσω λογισμικού για Android, iOS κλπ. Σε κάθε περίπτωση, για να μπορέσει κανείς να "ζητά" δεδομένα από το σύστημα της δημοφιλούς υπηρεσίας ροής μουσικής, πρέπει να έχει λάβει από

την ίδια κάποια στοιχεία αυθεντικοποίησης. Για να λάβει ένας χρήστης τα στοιχεία αυτά, καλείται πρώτα να συνδεθεί μέσω του ήδη υπάρχοντος λογαριασμού του, να δημιουργήσει μια εικονική εφαρμογή (δηλαδή να δηλώσει πως επιθυμεί να αναπτύξει λογισμικό που θα κάνει χρήση των δεδομένων της εταιρείας), και έπειτα η νεο-δημιουργηθείσα εφαρμογή θα αποκτήσει τα απαραίτητα στοιχεία αυθεντικοποίησης, που κατόπιν χρησιμοποιεί ο χρήστης θα την επικοινωνία με το API. Το λογισμικό της παρούσας διπλωματικής εργασίας έχει αναπτυχθεί μέσω της γλώσσας προγραμματισμού Python [58], ενώ για τη διασύνδεση με το Web API έγινε χρήση της εξειδικευμένης (για αυτό το σκοπό) βιβλιοθήκης spotipy [59], η οποία διευκολύνει αισθητά τη διαδικασία και την επικοινωνία με τη πλατφόρμα διεπαφής του Spotify. Χωρίς τη χρήση του spotipy, θα έπρεπε να συνταχθούν API requests (σε γλώσσα Python) και προσεκτικά να αναγνωρίζονται οι επιστροφές αυτών σε κείμενα τύπου JSON. Αν και δεν πρόκειται για μια δύσκολη υπόθεση, η αντικατάσταση αυτού του σκέλους με τη βιβλιοθήκη spotipy απεδείχθη ιδιαίτερα βολική. Στο επόμενο κεφάλαιο θα παρουσιαστούν τα κομμάτια κώδικα που κάνουν χρήση του spotipy και δημιουργούν τα σχετικά σύνολα δεδομένων.



**Σχήμα 4.2:** Ένα παράδειγμα χρήσης του Spotify Web API μέσω της επίσημης κουσόδηλας που παρέχει για δοκιμές η υπηρεσία. Στο συγκεκριμένο λειτουργεί για μήψη πληροφοριών ενός συγκεκριμένου καλλιτέχνη. Στο πεδίο *id* γίνεται εισαγωγή του μοναδικού χαρακτηριστικού κωδικού ενός καλλιτέχνη, ενώ στο αμέσως επόμενο πεδίο γίνεται εισαγωγή ενός στοιχείου αυθεντικοποίησης για τη λειτουργία του συστήματος. Δεξιά φαίνεται το *request*, καθώς και ένα κομμάτι της απάντησης. Το *id* που χρησιμοποιείται αφορά τον καλλιτέχνη Pitbull.

Για τις ανάγκες της παρούσας διπλωματικής εργασίας, έγινε δημιουργία τριών διαφορετικών συνόλων δεδομένων. Η μοναδική διαφορά τους έγκειται στο μέγεθος αυτών. Αρχικά, κατόπιν έρευνας, έγινε διαχωρισμός 19 βασικών μουσικών ειδών,

μέσα από τα οποία έγινε η συλλογή των κομματιών. Τα επιλεχθέντα 19 είδη μπορεί να θεωρηθεί πως καλύπτουν ολόκληρο το φάσμα της μουσικής, τουλάχιστον στο μεγαλύτερο τμήμα της. Εν ολίγοις, θα μπορούσαμε να θεωρήσουμε πως πρόκειται για 19 οικογένειες μουσικών ειδών. Ποια είναι αυτά τα είδη; Πρόκειται για τα **Rock, Classical, Dance, Electronic, Folk, Hip-hop, Rap, Jazz, Blues, Metal, Pop, Punk, Soul, Reggae, Country, Indie, Alternative, Latin & R&b**. Για το πρώτο και μεγαλύτερο σύνολο δεδομένων, χρησιμοποιήθηκε το API endpoint των μουσικών λιστών αναπαραγωγής, μέσα από το οποίο ζητήθηκαν 80 τυχαίες λίστες για κάθε μουσικό είδος. Έπειτα, μέσω του endpoint των μουσικών κομματιών, έγινε αναζήτηση για κάθε κομμάτι, από κάθε μια λίστα, με βάση το χαρακτηριστικό id του για τα μουσικά χαρακτηριστικά αυτού. Αυτό είχε ως αποτέλεσμα τη δημιουργία ενός συνόλου από τελικά αποτελείτο από **84307** κομμάτια (με τα διπλότυπα να έχουν αφαιρεθεί), με σχετικά ομαλή κατανομή ανά μουσικό είδος. Συγκεκριμένα, η κατανομή των κομματιών ανά genre είναι η ακόλουθη (με φθίνουσα σειρά): **Indie: 5419, Soul: 5040, Jazz: 5025, Metal: 4973, Folk: 4932, Punk: 4821, Electronic: 4781, Pop: 4714, Alternative: 4646, Classical: 4491, Dance: 4404, Hip-hop: 4396, Blues: 4309, Country: 4039, Rap: 3844, Rock: 3778, Latin: 3669, Reggae: 3587 & R&b: 3439**. Το συγκεκριμένο σύνολο αποτελεί και το βασικό για τη δοκιμή των μοντέλων μηχανικής μάθησης. Τα άλλα δύο σύνολα είναι αισθητά μικρότερα. Για το "μεσαίο" έγινε απευθείας αναζήτηση διακοσίων μουσικών κομματιών και των μουσικών στοιχείων τους ανά είδος (δεν έγινε αναζήτηση λιστών αναπαραγωγής και έπειτα των τραγουδιών καθεμίας) μέσω του spotipy, οδηγόντας σε ένα τελικό σύνολο από **3785** κομμάτια. Υπήρξαν ελάχιστα διπλότυπα, εξού και η σχεδόν πλήρως ομαλή κατανομή ανά είδος (τα περισσότερα σε ένα είδος είναι 200 - δηλαδή όσο ζητήθηκε - και τα λιγότερα 197, μόλις τρία λιγότερα). Τέλος, το "μικρό" σύνολο αποτελείται από **948** κομμάτια, αφού επιχειρήθηκε η ίδια αναζήτηση, όμως με όριο 50 τραγουδιών ανά είδος. Μόλις δυο είδη έχουν από 49 κομμάτια, δημιουργώντας μια οριακά ιδανική κατανομή. Τα δυο μικρότερα σύνολα αξιοποιούνται από το content-based σύστημα συστάσεων. Η εντολή αίτησης μουσικών λιστών αναπαραγωγής και τραγουδιών θα παρουσιαστεί στο επόμενο κεφάλαιο.

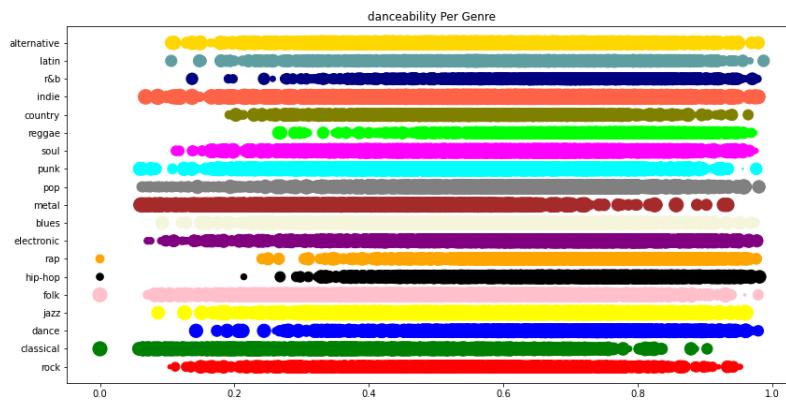
Ας θέσουμε τώρα ένα από τα πιο βασικά ερωτήματα που αφορούν το σύνολο δεδομένων μας. Ποιες πληροφορίες λαβαίνουμε για ένα μουσικό κομμάτι από το Spotify Web API, όταν προχωρούμε σε αναζήτηση με βάση το χαρακτηριστικό id του για εκείνο και τα μουσικά του στοιχεία; Σε πρώτη φάση, όταν γίνεται αναζήτηση στοιχείων για ένα κομμάτι βάσει id, ανακτώνται πληροφορίες που αφορούν το άλμπουμ στο οποίο εμφανίζεται, τον καλλιτέχνη που το κυκλοφόρησε, τη διάρκεια του, το πόσο δημοφιλές είναι στην επιλεχθείσα αγορά (χώρα), όπως φυσικά και τον τίτλο του. Από εκεί και πέρα, η δεύτερη φάση της αναζήτησης αφορά τα μουσικά του χαρακτηριστικά. Πρόκειται για εννιά (9) τιμές που αφορούν καθαρά τη μουσική ανάλυση που έχει γίνει πάνω στο κομμάτι από το Spotify. Αυτά τα μουσικά χαρακτηριστικά έχουν

ως εξής (με την αγγλική τους ονομασία):

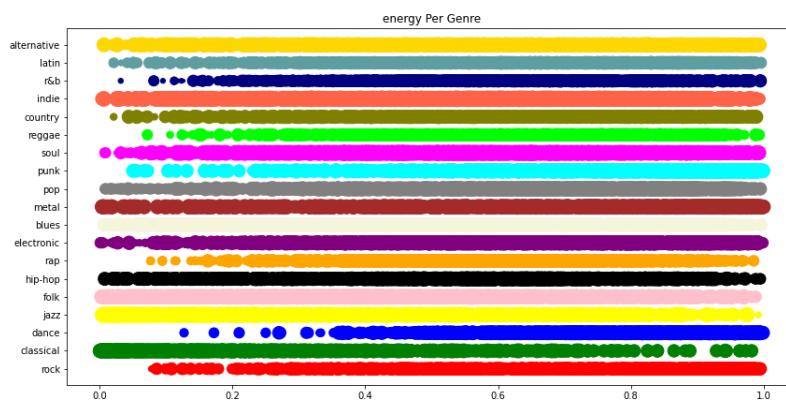
- **Danceability:** Μετρική που ορίζει το κατά πόσο ένα μουσικό κομμάτι είναι κατάλληλο για χορό. Βγαίνει με βάση έναν υπολογισμό διαφόρων παραγόντων. Κυμαίνεται στο διάστημα [0, 1].
- **Energy:** Μετρική που ουσιαστικά ορίζει το κατά πόσο ένα κομμάτι φέρει έντονα στοιχεία οργάνων όπως τα ντράμς, καθώς και το πόσο δυνατά ακούγονται κατά τη διάρκεια του τραγουδιού. Επηρεάζεται και από την ένταση της φωνής του τραγουδιστή. Κυμαίνεται στο διάστημα [0, 1].
- **Loudness:** Μετρική που ορίζει την ένταση του μουσικού κομματιού. Μετριέται σε Loudness Unit Full Scale, γνωστό ευρέως ως LUFS. Φέρει κυρίως αρνητικές τιμές ενώ σπάνια αγγίζει (ή και ξεπερνάει) το 0.
- **Speechiness:** Μετρική που ορίζει το κατά πόσο εντοπίζονται ομιλούμενες λέξεις σε ένα μουσικό κομμάτι. Κυμαίνεται στο διάστημα [0, 1].
- **Acousticness:** Μετρική που ορίζει κατά πόσο ένα μουσικό κομμάτι φέρει ακουστικά στοιχεία. Με τη φράση αυτή εννοούνται ήχοι που προέρχονται φυσικά (κλασικά όργανα, φωνές καλλιτέχνη, ήχοι φύσης), χωρίς ψηφιοποιήσεις. Κυμαίνεται στο διάστημα [0, 1].
- **Instrumentalness:** Μετρική που αφορά την ύπαρξη vocals σε ένα μουσικό κομμάτι. Όπως η πλειοψηφία των μουσικών χαρακτηριστικών, κυμαίνεται στο διάστημα [0, 1].
- **Liveness:** Μετρική που περιγράφει τη πιθανότητα να έχει γίνει καταγρασή του μουσικού κομματιού μπροστά σε live κοινό. Κυμαίνεται στο διάστημα [0, 1].
- **Valence:** Μετρική που χαρακτηρίζει τη "θετικότητα" ή μη ενός μουσικού κομματιού. Κυμαίνεται στο διάστημα [0, 1]. Κομμάτια με τιμές κοντά στο 1 τείνουν να είναι πιο χαρμόσυνα, ενώ εκείνα με τιμές κοντά στο 0 είναι πιο "στενάχωρα".
- **Tempo:** Μετρική που ορίζει τη ταχύτητα με την οποία εκτελείται ο μουσικός ρυθμός ενός κομματιού. Η πλειοψηφία των τιμών τείνει να κυμαίνεται στο διάστημα [50, 200].

Παρακάτω παρατίθενται τα διαγράμματα μουσικών ειδών ανά χαρακτηριστικό, σύμφωνα με το μεγάλο σύνολο δεδομένων που έχει δημιουργηθεί. Μέσα από τα διαγράμματα μπορούμε να εξάγουμε ασφαλή συμπεράσματα και να επιβαιθεώσουμε ήδη υπάρχουσες υποθέσεις μας. Σε κάθε περίπτωση, μπορούμε να αναμένουμε ότι (παραδείγματος χάρη) τα κλασικά κομμάτια θα έχουν χαμηλότερα επίπεδα σε μετρικές όπως το speechiness και το tempo. Αντιθέτως, τα τραγούδια του είδους hip-hop θα έχουν ακριβώς αντίθετα "αποτελέσματα" σε τέτοιες κατηγορίες. Αν και τα

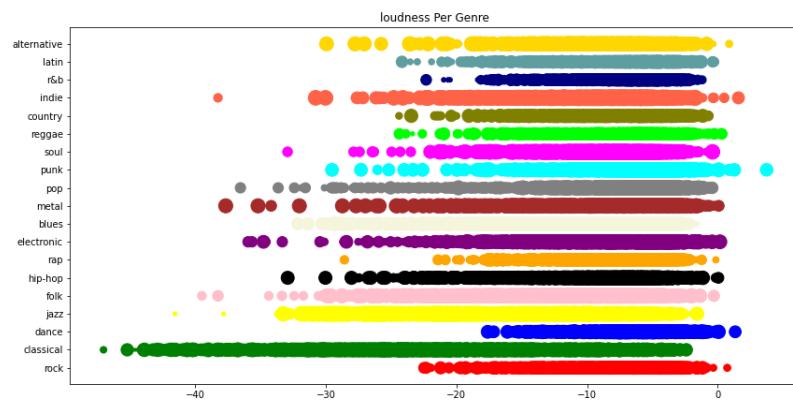
συγκεκριμένα συμπεράσματα δεν συμβάλλουν άμεσα στη βελτίωση κάποιων εκ των μοντέλων μηχανικής μάθησης ή recommender, είναι όμως χρήσιμα και ενδιαφέρονται για τη εν βάθος κατανόηση του συνόλου δεδομένων που έχει δημιουργηθεί.



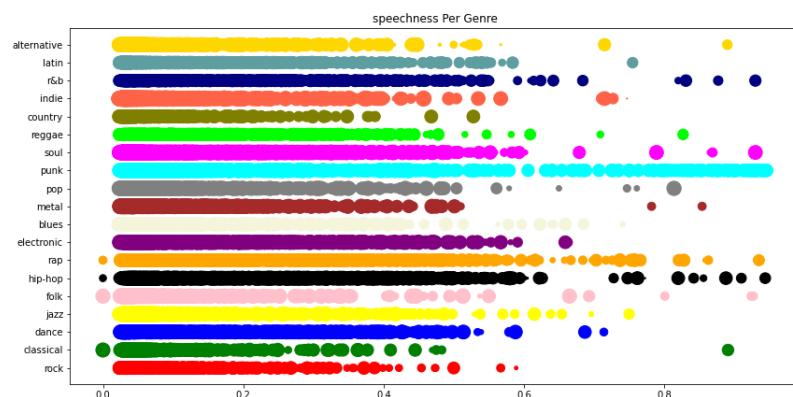
Σχήμα 4.3: *Danceability*, κατανομή ανά μουσικά είδη.



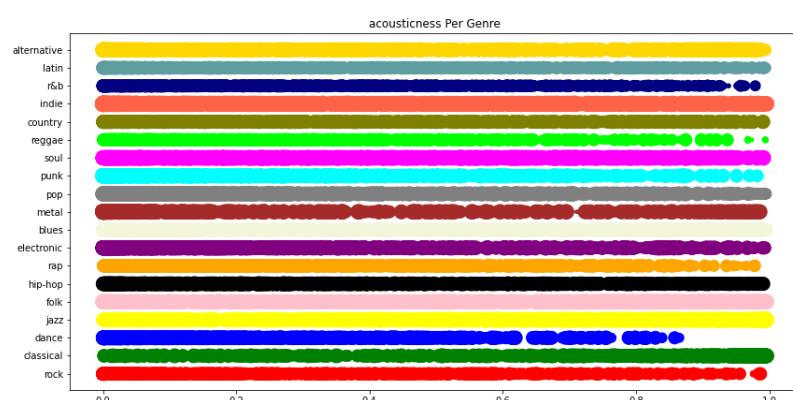
Σχήμα 4.4: *Energy*, κατανομή ανά μουσικά είδη.



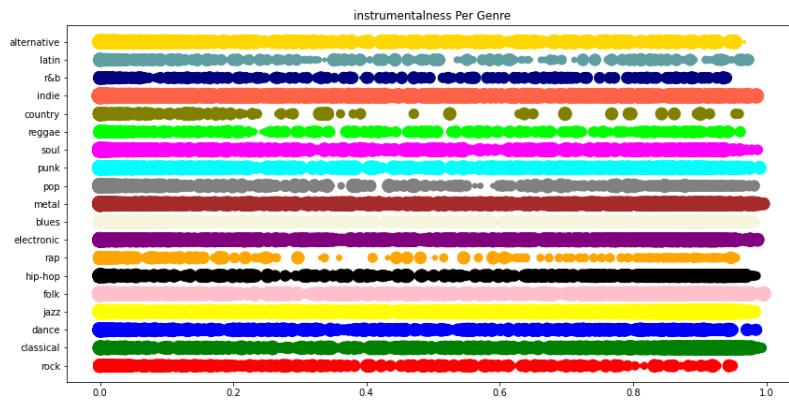
Σχήμα 4.5: *Loudness*, κατανομή ανά μουσικά είδη.



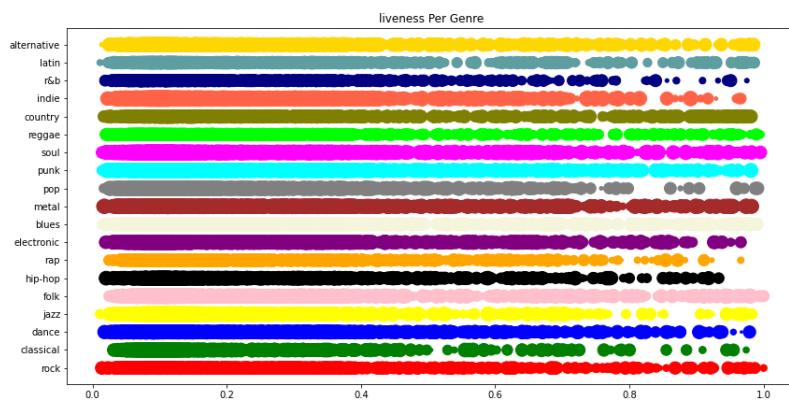
Σχήμα 4.6: *Speechiness*, κατανομή ανά μουσικά είδη.



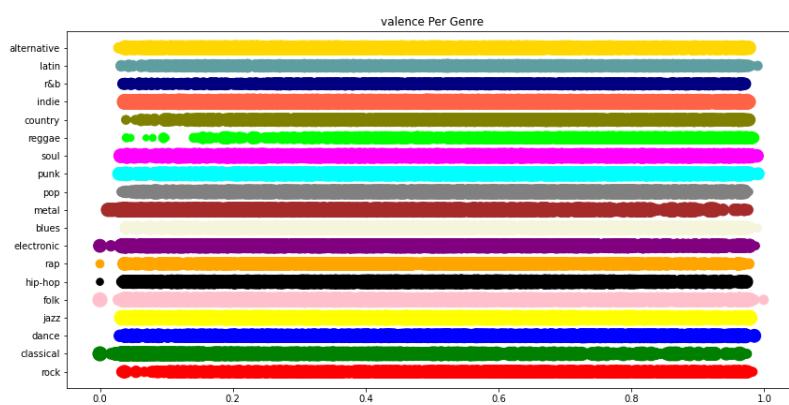
Σχήμα 4.7: *Acousticness*, κατανομή ανά μουσικά είδη.



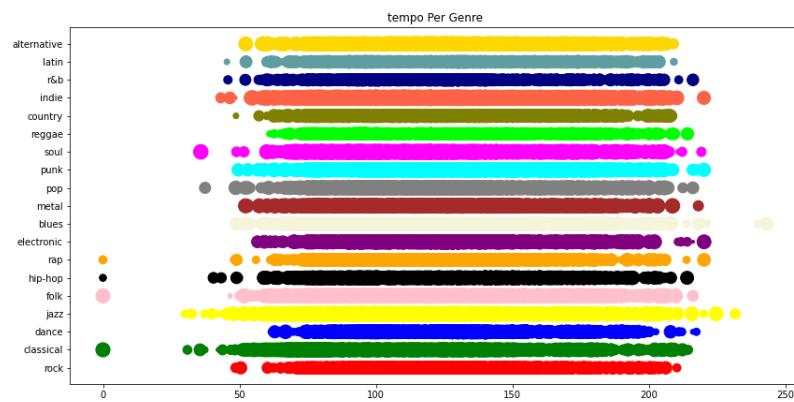
Σχήμα 4.8: *Instrumentalness*, κατανομή ανά μουσικά είδη.



Σχήμα 4.9: *Liveness*, κατανομή ανά μουσικά είδη.



Σχήμα 4.10: *Valence*, κατανομή ανά μουσικά είδη.



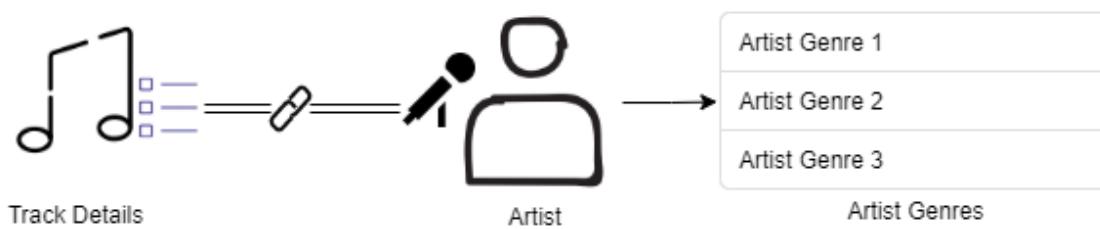
Σχήμα 4.11: *Tempo, κατανομή ανά μουσικά είδη.*

Μέσα από τα παραπάνω διαγράμματα μπορούμε να επιβεβαιώσουμε υποθέσεις μας γύρω από τα χαρακτηριστικά και ιδιότητες συγκεκριμένων μουσικών ειδών. Παραδείγματος χάρη, εξετάζοντας το πρώτο διάγραμμα (εκείνο του *danceability*), βλέπουμε πως είδη σαν το *dance* και το *latin* έχουν "ισχυρή" παρουσία στις υψηλές τιμές. Ακόμα και το *hip-hop* με το *r&b* έχουν υψηλά ποσοστά. Άλλα είδη όμως, όπως το κλασικό ή το *folk*, δεν έχουν την ίδια κατανομή. Ένα άλλο παράδειγμα μπορεί να παρατηρηθεί στο σκέλος του *speechiness*, όπου το κλασικό είδος φαίνεται να υστερεί αισθητά σε σχέση με άλλα (*rap*, *hip-hop* & *punk*). Το *rock* έχει επίσης χαμηλά ποσοστά σε αυτή τη κατηγορία, μαζί με μουσικά είδη παρόμοια σε εκείνο (*country*). Μπορούμε να προχωρήσουμε σε περαιτέρω ανάλυση των δεδομένων μας με βάση τα μουσικά τους χαρακτηριστικά, όμως αυτό δεν αποτελεί στόχο ή σκοπό της παρούσας διπλωματικής εργασίας. Σε κάθε περίπτωση, είναι ένα ενδιαφέρον ερευνητικό στοιχείο που μπορεί να αναλυθεί σε μελλοντική υλοποίηση.

Καθώς ολοκληρώνεται η κατανόηση των δεδομένων μας, αλλά και η συλλογή των χαρακτηριστικών για κάθε μουσικό κομμάτι, αντιλαμβανόμαστε πως υπάρχει ένα βασικό πρόβλημα. Το Spotify δεν παρέχει μουσικό είδος ανά τραγούδι. Αυτό είναι και ο λόγος για τον οποίο η παρούσα διπλωματική εργασία συμπεριέλαβε και το προαναφερθέν *genre recognition* στο ερευνητικό της σκέλος. Μπορεί λοιπόν να υπάρχουν πληροφορίες για τον καλλιτέχνη, το άλμπουμ και τα μουσικά χαρακτηριστικά για κάθε τραγούδι, δεν παρέχεται όμως ξεκάθαρη κατηγορία μουσικού είδους στην οποία ανήκει. Το συγκεκριμένο πρόβλημα θα ξεπεραστεί μέσω δυο διαφορετικών μεθόδων. Η πρώτη μέθοδος θέλει τα μουσικά είδη με τα οποία έγινε η αναζήτηση & συλλογή των κομματιών (τα 19 κύρια *genres*) να καθίστανται αυτομάτως και τα είδη εκείνων. Εν ολίγοις, κάθε ενα εκ των κομματιών που συλλέχθηκαν από τις 80 λίστες αναπαραγγής ανά είδος, θα λαβαίνει το είδος αυτό ως... είδος του. Οπότε, τα 5419 τραγούδια που προήλθαν από την αναζήτηση 80 playlists του Indie είδους, θα θεωρούνται Indie τραγούδια. Το ίδιο ισχύει και για τα υπόλοιπα 18 *genres*.

Αυτό εμπεριέχει τη παραδοχή πως όλα τα κομμάτια που συλλέγονται με βάση ένα συγκεκριμένο είδος, ανήκουν απόλυτα και σε εκείνο. Κάτι τέτοιο γνωρίζουμε πως δεν μπορεί να ισχύει, αλλά πρόκειται ακριβώς για μια παραδοχή την οποία είμαστε αναγκασμένοι να κάνουμε. Αυτή η πρώτη μέθοδος θα χρησιμοποιηθεί για την εκπαίδευση ταξινομητών μηχανικής μάθησης. Άρα, για τα machine learning μοντέλα θα ξεκινήσουμε με 19 labels (genres).

Η δεύτερη μέθοδος με την οποία καλούμαστε να ξεπεράσουμε το πρόβλημα της έλλειψης ειδών ανά τραγούδι, αφορά τον content-based recommender. Μέσω του Spotify Web API, για κάθε τραγούδι που συλλέγεται, γίνεται αυτομάτως αναζήτηση πληροφοριών για τον καλλιτέχνη του. Από τις πληροφορίες που έρχονται, θα κρατώνται τα μουσικά είδη του καλλιτέχνη, τα οποία και θα περνάνε στο τραγούδι. Οπότε, εδώ κάνουμε μια νέα παραδοχή, η οποία ακούει στο ότι τα μουσικά είδη ενός καλλιτέχνη, είναι πάντα μουσικά είδη και των τραγουδιών του. Για άλλη μια φορά, κάτι τέτοιο δεν ισχύει απόλυτα. Πολλοί καλλιτέχνες έχουν κυκλοφορήσει τραγούδια που είναι αισθητά έξω από τη μουσική "ζώνη ασφάλειας" τους, θέλοντας να πειραματιστούν με κάτι διαφορετικό. Επιπρόσθετα, υπάρχουν καλλιτέχνες που με τον καιρό διαφοροποιούν τη θέση τους στο μελωδικό στερέωμα. Ενδεχομένως τα μουσικά είδη που τους έχει αντιστοιχήσει το Spotify να χαίρουν αλλαγής ή ανανέωσης. Όμως, όπως και πρίν, πρόκειται για μια παραδοχή που είμαστε αναγκασμένοι να πάρουμε. Οπότε, στη δεύτερη μέθοδο θα προσδίδονται στα μουσικά κομμάτια τα είδη των καλλιτεχνών τους. Οι περισσότεροι καλλιτέχνες χαρακτηρίζονται από μια σειρά διαφορετικών ειδών (συγγενικών μεταξύ τους), άρα τα μουσικά τους κομμάτια θα "κληρονομήσουν" παραπάνω από ένα. Αυτό βοηθάει ακόμα περισσότερο το σύστημα συστάσεων μας, αν και θα ενισχυθεί περαιτέρω, όπως θα δούμε παρακάτω.



**Σχήμα 4.12:** Η διαδικασία απόκτησης των μουσικών ειδών για ένα κομμάτι. Αφού πρώτα συλλέχθουν τα μελωδικά χαρακτηριστικά του, έπειτα γίνεται αναζήτηση πληροφοριών στο Web API με βάση τον καλλιτέχνη του, και εν συνεχείᾳ συλλέγονται τα είδη που τον χαρακτηρίζουν.

Κάπως έτσι, διαμορφώνεται και το τελικό σύνολο δεδομένων μας. Όλες οι παραπάνω λεπτομέρειες σχετικά με την αντιστοιχίση μουσικών ειδών στα τραγούδια αφορά φυσικά και τα δυο μικρότερα datasets που έχουν δημιουργηθεί και στα οποία αναφερθήκαμε παραπάνω. Τελικά, το σύνολο δεδομένων διαμορφώνεται όπως φαίνεται στη παρακάτω εικόνα (οι πρώτες πέντε γραμμές του dataframe - dataset σε

περιβάλλον Python):

	trackid	trackname	artist	artistid	genres	main_genre	danceability	energy	loudness	speechness	acousticness	instrumentalness	liveness	valence	tempo
0	0hCB0YR03f6AmQaHbw/WD8	wholeotta love - 1990 remaster	led zeppelin	36QJpDe2go2KgaRleHC0Tp	albumrock classicrock hardrock	rock	0.412	0.902	-11.600	0.4050	0.04840	0.131000	0.4050	0.422	89.740
1	2ao02jRnM3A0NyLqgMN2f	all along the watchtower	jimi hendrix	776Ue845nYHJpNaStV1Ds4	acidrock albumrock bluesrock classicrock hard...	rock	0.438	0.805	-6.237	0.0624	0.00255	0.000057	0.0842	0.564	113.253
2	08mG3Y1vJYAbvDt4WdJ	back in black	ac/dc	711MCceyC8cFnzjGY4Q7Un	albumrock australianrock hardrock	rock	0.310	0.700	-5.678	0.0470	0.01100	0.009650	0.0828	0.763	188.386
3	1RJeAIwR9pZBqJA8ndZLL	paint it, black - mono	the rolling stones	22bE4uQ8BaNvSHPVcDxLCe	britishinvasion classicrock rock	rock	0.505	0.836	-7.384	0.1010	0.01030	0.077100	0.2950	0.764	159.669
4	66iOpKgTyFjOrac4S1s4g	rebel radio -1990 remaster	david bowie	0oSGxWVmOXhD2fKuz2Gy	albumrock artrock classicrock dancerock glam...	rock	0.676	0.920	-7.960	0.0435	0.00101	0.000009	0.2480	0.624	126.341

## 4.2 Ταξινομητές Μηχανικής Μάθησης

Όπως έχει ήδη ξεκαθαριστεί παραπάνω, σε αυτό το σημείο θα αξιοποιηθεί το μεγάλο σύνολο δεδομένων που δημιουργήθηκε. Στόχος των μοντέλων μηχανικής μάθησης που ενσωματώνονται στη παρούσα διπλωματική εργασία είναι να συμβάλλουν στην αναγνώριση των μουσικών ειδών για κάθε κομμάτι, κάτι που τελικά θα βοηθήσει τη σύσταση εκείνων στο υλοποιηθέν σύστημα. Βέβαια, όπως έχει ήδη ξεκαθαριστεί, δεν υπάρχουν πραγματικά και απόλυτα ποιοτικά labels - genres για κάθε τραγούδι. Στη προκειμένη, έχει γίνει η παραδοχή πως κάθε κομμάτι θα λάβει ως είδος εκείνο μέσω του οποίου συλλέχθηκε (δηλαδή το είδος από το οποίο έγινε αναζήτηση 80 λιστών αναπαραγωγής και εντός των οποίων βρίσκεται το συγκεκριμένο τραγούδι). Παρακάτω καταγράφονται οι classifiers μηχανικής μάθησης που επιλέχθηκαν για δοκιμή και χρήση. Παραπάνω, έχει ήδη γίνει εισαγωγή στη μηχανική μάθηση και τις κατηγορίες της, οπότε θα εισέλθουμε κατευθείαν στα μοντέλα που θα δοκιμαστούν. Μια αρχική ιδέα ήταν η χρήση μοντέλων μη-επιβλεπόμενης μάθησης, εφόσον μάλιστα δεν υπήρχαν και αυθεντικά labels μουσικών ειδών που να παρέχονται από το Spotify για κάθε τραγούδι. Η ιδέα ήθελε τους unsupervised ταξινομητές να δημιουργούν συστάδες ανάλογα με τις ομάδες ειδών μέσω των οποίων είχαν συλλεχθεί τα κομμάτια (δηλαδή 19 clusters, όσα και τα μουσικά genres). Όμως, οι δοκιμές απέφεραν εξαιρετικά απογοητευτικά αποτελέσματα, οπότε η ιδέα αυτή εγκατελείφθη νωρίς. Η πρόταση της "επαλήθευσης" των ειδών ανάλογα με τα μουσικά χαρακτηριστικά και με τη παραδοχή που αναφέραμε παραπάνω, υπερίσχυσε ως η επικρατέστερη ιδέα. Οπότε, παρακάτω περιλαμβάνονται μόνο ταξινομητές επιβλεπόμενης μάθησης (μερικοί υλοποιούνται και σε αρχιτεκτονικές ημι-επιβλεπόμενης μάθησης), ενώ όλοι τους είναι ιδιαίτερα δημοφιλείς. Τα μοντέλα που ενσωματώνονται έχουν ως εξής:

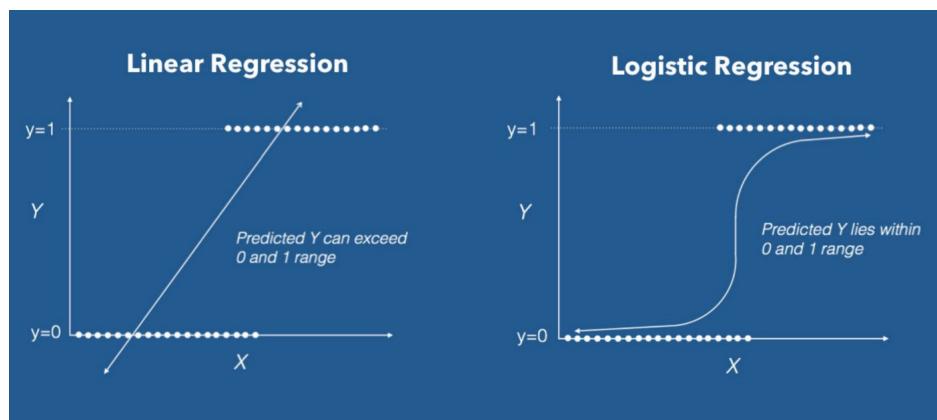
- **Naive Bayes Classifier:** Πρόκειται για μια οικογένεια μοντέλων μηχανικής μάθησης, γνωστών και ως απλών "πιθανοτικών ταξινομητών", οι οποίοι βασίζονται στο θεώρημα Bayes. Χρησιμοποιώντας το θεώρημα αυτό, μπορεί κανείς

να υπολογίζει τη πιθανότητα του να συμβεί το ενδεχόμενο **A**, δοθέντος του γεγονότος ότι έχει συμβεί το ενδεχόμενο **B**. Οπότε, το **B** είναι το αποδεικτικό στοιχείο και το **A** η υπόθεση. Το θεώρημα φέρει τη παραδοχή της ανεξαρτησίας των χαρακτηριστικών, πράγμα που σημαίνει ότι η παρουσία ενός συγκεκριμένου feature δεν επηρεάζει κάποιο. Εξού και η πρόσδοση του επιθέτου "αφελής" (Naive).

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Σχήμα 4.13: Το γνωστό και ιδιαίτερα δημοφιλές Θεώρημα Bayes.

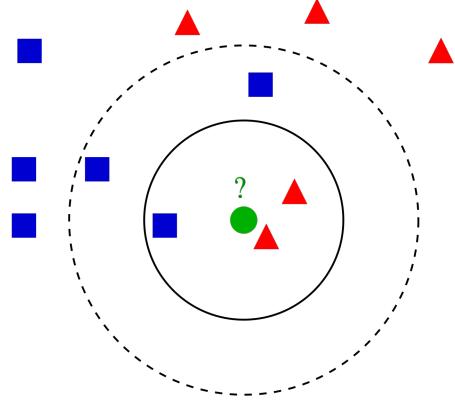
- **Logistic Regression Classifier:** Όπως γίνεται αντιληπτό και από την ονομασία, αποτελεί μέλος της οικογένειας των μοντέλων παλινδρόμησης. Η λογιστική παλινδρόμηση ερευνά το μη γραμμικό αποτέλεσμα μίας εξαρτημένης κατηγορικής μεταβλητής αναφορικά με τη δράση πολλών ανεξάρτητων μεταβλητών. Χαρακτηρίζεται, αναλόγως της φύσης των κατηγοριών της εξαρτημένης μεταβλητής, από τρεις κατηγορίες μοντέλων, τη διωνυμική παλινδρόμηση (με δύο μόνο κατηγορίες), την τακτική (οι κατηγορίες διατάσσονται με αυξητική τάση) και την ονομαστική (ποιοτικές κατηγορίες). Ως ταξινομητής, η λογιστική παλινδρόμηση τείνει να τα πηγαίνει εξαιρετικά καλά.



Σχήμα 4.14: Η διαφορά μεταξύ της γραμμικής (αριστερά) και λογιστικής (δεξιά) παλινδρόμησης, ως προς τη κατανομή των τιμών που τείνουν να πάρουν σε ένα δυαδικό μοντέλο [15].

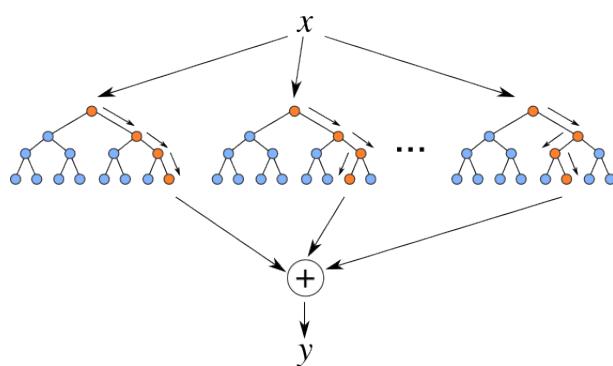
- **k-Nearest Neighbors Classifier (kNN):** Γνωστός στα ελληνικά και ως Κατηγοριοποιητής Εγγύτατου Γείτονα, ο kNN αποτελεί μια από τις πιο απλές μη παραμετρικές μεθοδολογίες ταξινόμησης. Για τη κατηγοριοποίηση οποιουδήποτε δείγματος, πρώτα υπολογίζονται οι αποστάσεις του δείγματος αυτού από όλα

τα πρότυπα εκπαίδευσης, έπειτα κρατούνται τα k πλέον κοντινά πρότυπα (δηλαδή οι γείτονες) και τελικά το δείγμα ταξινομείται στη κλάση που εμφανίζεται με τη μεγαλύτερη συχνότητα μεταξύ των k γειτόνων. Για τον υπολογισμό της απόστασης δυο σημείων θεωρείται η ευκλείδια απόσταση (αν και αυτό μπορεί να αλλάξει).



Σχήμα 4.15: Απεικόνιση της λειτουργίας του αλγορίθμου kNN. Ανάλογα με τον καθορισμό του k, το μοντέλο θα ταξινομήσει το δείγμα (πράσινη κουκίδα) στα κόκκινα ή τα μπλέ τετράγωνα [16].

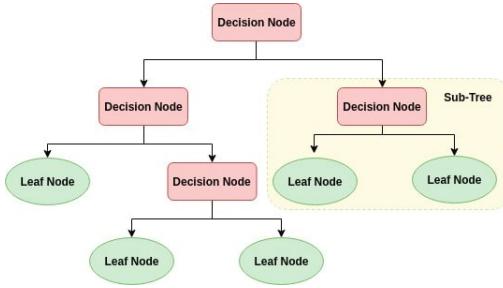
- **Random Forest Classifier:** Όπως οι προηγούμενοι (αλλά και οι επόμενοι) ταξινομητές, πρόκειται για έναν αλγόριθμο επιβλεπόμενης μάθησης. Το "δάσος" το οποίο εκείνος σχηματίζει, είναι στην ουσία ενα σύνολο δέντρων αποφάσεων, τα οποία εκπαιδεύονται συνήθως με τη μέθοδο Bootstrap Aggregating, ώστε να διασφαλίσουν την ακρίβεια και τη σταθερότητα στα αποτελέσματα. Η γενική ιδέα αυτης της μεθόδου εκπαίδευσης που χρησιμοποιεί ο Random Forest είναι πως ο συνδυασμός μοντέλων μάθησης βελτιώνει το τελικό αποτέλεσμα. Εν ολίγοις, ο Random Forest χτίζει πολλαπλά δέντρα αποφάσεων και τελικά τα ενώνει ώστε να αποκτήσει ένα ποιοτικότερο αποτέλεσμα.



Σχήμα 4.16: Η αρχιτεκτονική και ο τρόπος λειτουργίας του Random Forest Classifier.

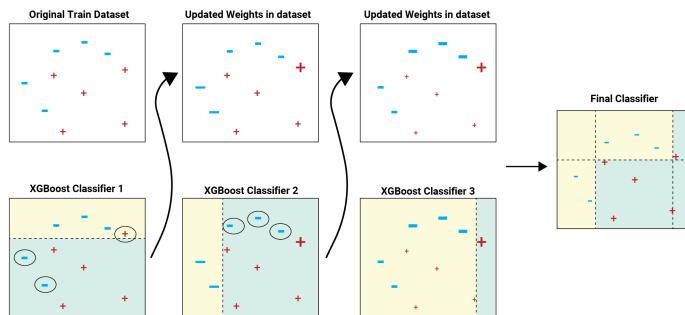
- **Decision Tree Classifier:** Τα Δέντρα Απόφασης, όπως είναι η ελληνική ονομασία τους, αποτελούν με τη σειρά τους μια κατηγορία τεχνικών μη γραμ-

μικών κατηγοριοποιητών στην οποία ο χώρος των χαρακτηριστικών διαιρείται σε περιοχές που αντιστοιχούν στις επιθυμητές κλάσεις. Η απόφαση για τη κατηγοριοποίηση ενός χαρακτηριστικού προκύπτει από τις απαντήσεις σε ερωτήματα που υποβάλλονται σύμφωνα με μία δενδρική δομή. Με λίγα λόγια, τα δέντρα αποφάσεων βασίζονται στις δοθείσες πληροφορίες και ακολουθούν μια συγκεκριμένη δομή, η οποία τους οδηγεί στο τελικό αποτέλεσμα.



Σχήμα 4.17: Τα δέντρα αποφάσεων και η λειτουργία μέσω της οποίας προχωρούν στην ταξινόμηση των χαρακτηριστικών [17].

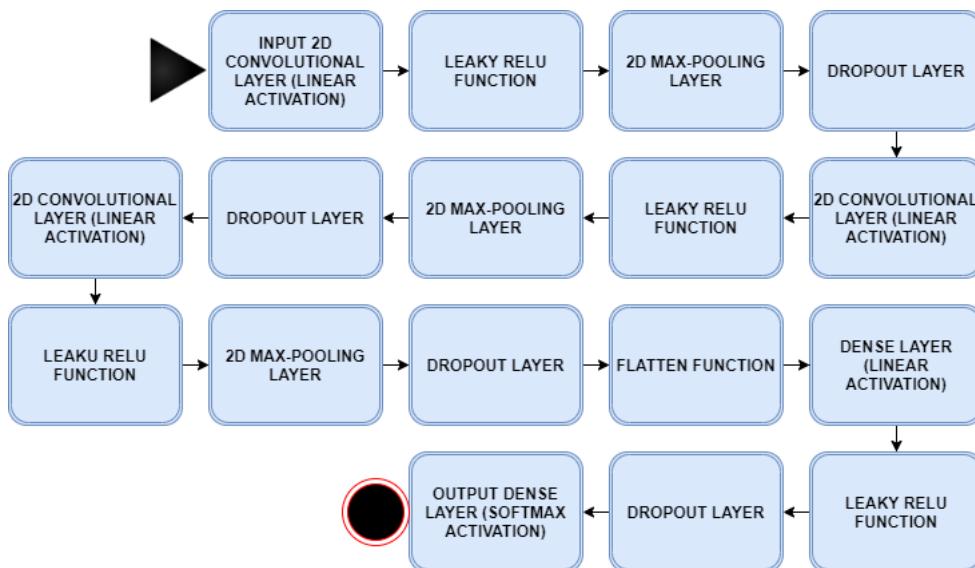
- **XGBoost Classifier:** Ο όρος XGBoost είναι στη πραγματικότητα τα αρχικά της πλήρους ονομασίας eXtreme Gradient Boosting. Ο αλγόριθμος αποτελεί μια σχετικά νέα τάση που έχει καταφέρει να καθιερώθει ανάμεσα στα πιο αποδοτικά και ποιοτικά μοντέλα μηχανικής μάθησης. Στη πραγματικότητα, αυτό που κάνει ο ταξινομητής είναι να ενισχύει την απόδοση ενός κλασικού μοντέλου gradient boosting. Χρησιμοποιεί ένα πιο κανονικοποιημένο αλγόριθμο τυποποίησης που αποτρέπει το πρόβλημα του overfitting (το μοντέλο "μαθαίνει" πάνω στα δοθέντα δεδομένα σε υπερβόλικο βαθμό, μη μπορώντας να "προσαρμοστεί" καλά σε νέα εισερχόμενα χαρακτηριστικά για ταξινόμηση), κάτι που του δίνει καλύτερη απόδοση.



Σχήμα 4.18: Απεικόνιση του τρόπου με τον οποίο δρα ο XGBoost, ενισχύοντας την απόδοση των κλασικών gradient boosting μοντέλων και βελτιώνοντας το τελικό αποτέλεσμα [18].

Πέραν όμως των έξι αυτών ταξινομητών επιβλεπόμενης μάθησης, δημιουργήθηκε και ένα προσαρμοσμένο νευρωνικό δίκτυο, συγκεκριμένα ένα **Convolutional Neural**

**Network (CNN) δυο διαστάσεων**, θέλοντας να υπάρξει και μια απαλή “νότα” του πεδίου της βαθιάς μάθησης. Η ανάπτυξη του έγινε με γνώμονα των τύπων δεδομένων που έχουμε στη κατοχή μας, δηλαδή μουσικά κομμάτια με τα μελωδικά τους χαρακτηριστικά. Αν και τα δίκτυα CNN περισσότερων των μια διαστάσεων κυρίως χρησιμοποιούνται για οπτικά δεδομένα (εικόνες), ένας ανασχηματισμός στο σύνολο μας μπορεί να το καταστήσει μια ιδανική είσοδο σε 2d CNN. Το δίκτυο φέρει τρία convolutional επίπεδα (ένα εκ των οποίων είναι και το πρώτο), τρία max-pooling επίπεδα, τέσσερις χρήσεις της συνάρτησης Leaky ReLU, δύο dense επίπεδα (ένα εκ των οποίων είναι και το τελευταίο) και μια flatten συνάρτηση. Επίσης, έχουν τοποθετηθεί τρία σημεία dropout, μέσω των οποίων διασφαλίζεται η ελαχιστοποίηση των πιθανοτήτων εμφάνισης overfitting. Το συγκεκριμένο μοντέλο αποτελεί το βασικό σύστημα μηχανικής μάθησης της παρούσας διπλωματικής εργασίας, καθώς και τον μοναδικό ταξινομητή που δημιουργήθηκε και παραμετροποιήθηκε από την αρχή, εφόσον τα παραπάνω μοντέλα υπάρχουν έτοιμα μέσω της βιβλιοθήκης Scikit-Learn [60].



Σχήμα 4.19: Η αποτύπωση της διάταξης των επιπέδων στο δημιουργηθέν 2d convolutional νευρωνικό δίκτυο για το μουσικό σύνολο δεδομένων μας.

Να σημειωθεί πως το σύνολο δεδομένων ανασχηματίζεται σε διάταξη (3,3,1), από την αρχική των 9 χαρακτηριστικών για κάθε μουσικό στοιχείο, ώστε να είναι σε θέση να χρησιμοποιηθεί από τον νευρώνα. Έγιναν δοκιμές και με χρήση νευρώνα μιας διάστασης, αλλά τα αποτελέσματα δεν ήταν τόσο καλά, συγκριτικά με το 2d μοντέλο. Για το compilation του μοντέλου, χρησιμοποιείται ο βελτιστοποιητής (optimizer) Adam. Σύμφωνα με την αρχιτεκτονική του συστήματος μας, το δημιουργηθέν νευρωνικό δίκτυο θα εκπαιδευτεί μέσω του μεγάλου συνόλου δεδομένων. Έπειτα, θα γίνει δοκιμή (testing) πάνω στα δύο μικρότερα σύνολα. Να τονιστεί πως ακριβώς το ίδιο θα γίνει και για τους προηγούμενους έξι ταξινομητές. Από εδώ και πέρα όμως,

η λειτουργία του συστήματος αφορά αποκλειστικά τα αποτελέσματα του... custom νευρώνα. Κάθε μουσικό κομμάτι θα λάβει ένα εκ των 19 genres, σύμφωνα με τη κατηγοριοποίηση του CNN. Τελικά, αν το μουσικό είδος που πρόβλεψε το δίκτυο συμφωνεί με κάποιο εκ των μουσικών ειδών που έχουν δωθεί στο κομμάτι από τον καλλιτέχνη του, αυτό το genre θα ενισχύεται (πολλαπλασιάζεται), ώστε να προσδιοθεί σε αυτό βάρος, κάτι που θα αξιοποιηθεί από το content-based σύστημα συστάσεων του συστήματος. Ποια είναι η αρχιτεκτονική αυτού του recommender; Ας δούμε...

## 4.3 Content-Based Σύστημα Συστάσεων

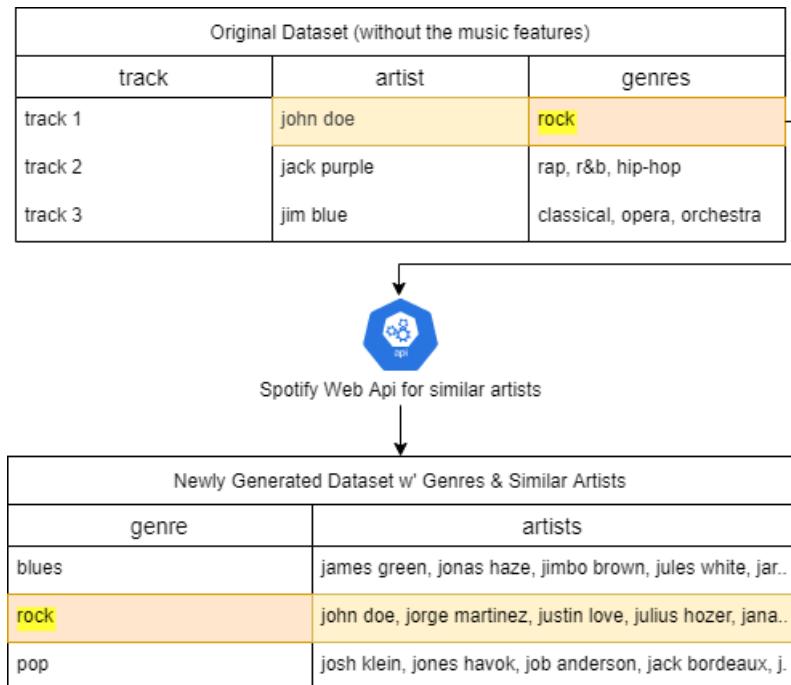
Το σύστημα συστάσεων που έχει αναπτυχθεί στα πλαίσια της παρούσας διπλωματικής εργασίας αφορά έναν "κλασικό" content-based recommender. Βέβαια, η χρήση μοντέλου μηχανικής μάθησης και η αξιοποίηση των ταξινομήσεων του εντός του recommender, μπορεί να καταστήσει το σύστημα ως υβριδικό. Σε κάθε περίπτωση, ο τρόπος λειτουργίας του είναι σύμφωνος με τη content-based λογική & αρχιτεκτονική. Συγκεκριμένα, εδώ γίνεται χρήση των δυο μικρότερων συνόλων δεδομένων. Ουσιαστικά, το σύστημα συστάσεων που έχει αναπτυχθεί, εφαρμόζεται άμεσα πάνω σε οποιοδήποτε μουσικό σύνολο. Μπορούμε να "βαφτίσουμε" τα δυο σύνολα δεδομένων μας ως δυο λίστες αναπαραγώγής, οι οποίες περιέχουν μουσικά κομμάτια που προέρχονται από 19 βασικά genres. Οπότε, ο recommender μας εφαρμόζεται άμεσα πάνω σε οποιαδήποτε λίστα αναπαραγώγής, δημιουργώντας τελικά συστάσεις αποκλειστικές για εκείνη. Πλέον, οι σημαντικές πληροφορίες για κάθε μουσικό κομμάτι δεν είναι τα μελαδικά χαρακτηριστικά του, όπως συνέβαινε στη περίπτωση των ταξινομητών μηχανικής μάθησης. Τώρα, οι μοναδικές ουσιώδεις πληροφορίες είναι το όνομα του καλλιτέχνη που έχει δημιουργήσει το κομμάτι, καθώς και τα μουσικά του είδη, τα οποία τελικά "κληρονομούνται" στο τραγούδι. Η χρήση των αποτελεσμάτων του 2D CNN θα λάβει χώρα στο τελικό στάδιο της λειτουργίας του recommender. Έχοντας απομονώσει τις πληροφορίες που χρειάζονται, το σύστημα συστάσεων προχωράει σε δυο βασικές διαδικασίες. Ουσιαστικά, δημιουργεί δυο νέα βοηθητικά σύνολα δεδομένων. Εκείνα έχουν ως εξής:

- Πρώτα, απομονώνται όσοι καλλιτέχνες "υπάρχουν" εντός του συνόλου δεδομένων / λίστας αναπαραγώγής. Έπειτα, γίνεται χρήση του Spotify Web API endpoint που επιτρέπει τον εντοπισμό παρόμοιων καλλιτεχνών, με βάση το id ενός άλλου που δίνεται ως όρισμα. Με το που συλλεχθούν οι παρόμοιοι καλλιτέχνες (που βρίσκονται δηλαδή στον ίδιο μελαδικό "χώρο" με τον πρώτο), γίνεται αναζήτηση των μουσικών ειδών τους, για κάθε έναν εξ αυτών ξεχωριστά. Τόσο οι νέοι καλλιτέχνες που συλλέχθηκαν, όσο και τα μουσικά είδη τους, αποθηκεύονται σε ένα νέο σύνολο δεδομένων, το οποίο περιέχει αποκλειστικά αυτες τις δυο κατηγορίες πληροφοριών (καλλιτέχνες και τα είδη που τους

αντιστοιχούν).

- Έπειτα, ξεκινάει η διαδικασία σύμπτυξης των συλλεχθέντων πληροφοριών, με βάση το μουσικό είδος. Από τη δημιουργία του παραπάνω νέου συνόλου δεδομένων, είναι σχεδόν βέβαιο ότι θα έχουν σχηματιστεί διπλότυπα καλλιτεχνών (μερικοί καλλιτέχνες θα έχουν... mutual παρόμοιους συ-καλλιτέχνες), ενώ το ίδιο βέβαιη θεωρείται η εμφάνιση διπλότυπων μουσικών ειδών. Έτσι, μέσω μιας συνάρτησης (γραμμένης σε γλώσσα προγραμματισμού Python), τα δεδομένα συγχωνεύονται. Για κάθε μουσικό είδος που εμφανίζεται στο προηγούμενο σύνολο δεδομένων, αντιστοιχούνται όλοι οι καλλιτέχνες που το έχουν. Έτσι, μια σειρά θα έχει ένα μουσικό είδος και όλους τους καλλιτέχνες που το μοιράζονται. Διαγράφονται τα διπλότυπα, δημιουργούνται μοναδικές σειρές ειδών, ενώ στις ομάδες καλλιτεχνών δεν υπάρχει διπλή εμφάνιση κάποιου, γιατί κάτι τέτοιο θα επηρέαζε τη βαρύτητα αργότερα, στον υπολογισμό των συστάσεων.

Οπότε, από το πρώτο σύνολο δεδομένων που έχει δημιουργηθεί μέσω του Spotify Web API (ή το "μεσαίο" ή το "μικρό"), παράγονται δυο νέα σύνολα. Το πρώτο παραχθέν σύνολο έχει ως μοναδικό λόγο ύπαρξης τη παραγωγή του δεύτερου και σημαντικότερου συνόλου, εκείνου δηλαδή που περιέχει ολοκληρωμένα τα στοιχεία μουσικών ειδών με τους καλλιτέχνες τους. Γιατί όμως να γίνει κάτι τέτοιο; Γιατί να λάθει χώρα όλη αυτή η διαδικασία; Ο λόγος είναι ξεκάθρος. Πρόκειται για ενίσχυση του υπαρχόντος συνόλου δεδομένων, στο σκέλος που αξιοποιεί ο content-based recommender. Με την συλλογή σχετικών καλλιτεχνών & των μουσικών ειδών τους, και εν συνεχείᾳ τη δημιουργία συνόλου ειδών με ομαδοποιημένους τους καλλιτέχνες, έχουμε πετύχει τον εμπλουτισμό του dataset, κάτι του ιδιαίτερα ουσιώδες για τη content-based διαδικασία συστάσεων. Με λίγα λόγια, αν δεν είχαμε προχωρήσει σε αυτό το βήμα, τότε τα μόνα δεδομένα που θα είχαμε για τον recommender (για κάθε μουσικό κομμάτι) θα ήταν ένα όνομα του καλλιτέχνη και μερικά genres που του αντιστοιχούσαν. Τώρα όμως, συνδυάζοντας το αρχικό σύνολο με το τελευταίο παραχθέν, για κάθε μουσικό κομμάτι θα έχουμε έναν καλλιτέχνη, αλλά για κάθε ένα από τα είδη που του αντιστοιχούν, θα υπάρχουν (στο νεό σύνολο) ολόκληρες ομάδες σχετικών καλλιτεχνών, κάτι που αποτελεί "τροφή" για το content-based σύστημα. Όλη αυτή η διαδικασία στη πραγματικότητα εμπλούτισε το αρχικό μας σύνολο με πολύ σημαντικές πληροφορίες. Ας διατηρούμε στο νου μας πως τα μελωδικά χαρακτηριστικά δεν είναι χρήσιμα σε αυτή τη φάση. Οπότε, με κάποιον τρόπο έπρεπε να ενισχυθούν τα υπάρχοντα, δηλαδή ο καλλιτέχνης και τα σχετικά μουσικά είδη.



**Σχήμα 4.20:** Η μέθοδος εμπλουτισμού των δεδομένων και η δημιουργία του νέου συνόλου, μέσω χρήσης του endpoint εντοπισμού παρόμοιων καλλιτεχνών από το Spotify Web API. Γίνεται αναζήτηση σχετικών καλλιτεχνών και των μουσικών ειδών τους. Τελικά, ένα νέο σύνολο προκύπτει με τα εμπλουτισμένα είδη και ομαδοποιημένους τους καλλιτέχνες ανά είδος. Έτσι, ένα μουσικό κομμάτι στο αρχικό σύνολο, το οποίο θα αντιστοιχεί σε καποια είδη, θα έχει επιπρόσθετη αντιστοίχηση σε πληροφορίες (καλλιτέχνες), μέσω αυτών των ειδών.

Κάπως έτσι ενισχύεται το σύνολο δεδομένων μας για τη περίπτωση του content-based συστήματος συστάσεων. Ας δούμε τώρα τις τεχνικές που υλοποιεί ο αλγόριθμος για να παράξει τις συστάσεις / προτιμήσεις μουσικών κομματιών. Συγκεκριμένα, γίνεται χρήση δυο συνδυασμών. Ο πρώτος συνδυασμός είναι εκείνος του Count Vectorizer με το Cosine Similarity, ενώ ο δεύτερος είναι εκείνος του Tf-Idf Vectorizer με το Linear Kernel. Οι δυο αυτοί συνδυασμοί αξιοποιούν το νεο-παραχθέν σύνολο δεδομένων με τα μουσικά είδη και τους καλλιτέχνες. Αυτό που κάνουν είναι να επεξεργάζονται τις πληροφορίες των καλλιτεχνών και να δημιουργούν σκορ ομοιότητας μεταξύ των μουσικών ειδών. Ας πάρουμε για παράδειγμα δυο είδη, το Rock και το Metal. Είναι πολύ πιθανό, τα είδη αυτά να περιέχουν κοινούς καλλιτέχνες. Εξίσου πιθανό είναι μεταξύ των ειδών hip-hop και rap. Με βάση αυτές τις κοινές εμφανίσεις καλλιτεχνών, υπολογίζονται οι ομοιότητες μεταξύ των genres. Ας δούμε τα σκέλη των δυο συνδυασμών και τον τρόπο με τον οποίο υπολογίζουν τις ομοιότητες:

- Ο πρώτος συνδυασμός είναι μεταξύ **Count Vectorizer & Cosine Similarity**. Αναφορικά με το Count Vectorizer, πρόκειται για μια μέθοδο η οποία κάνει "tokenize" (δηλαδή διαχωρίζει προτάσεις και ονομασίες σε λέξεις) το διθέν κείμενο, πραγματοποιώντας βασικές λειτουργίες προ-επεξεργασίας των

δεδομένων. Έπειτα, αντιστοιχεί τις λέξεις σε αριθμούς (labels) και μετράει το σύνολο των εμφανίσεων τους στο δοθέν κείμενο. Αυτά αποτελούν και την επιστροφή του στο χρήστη. Τα αποτελέσματα αποτελούν με τη σειρά τους είσοδο στον αλγόριθμο του Cosine Similarity. Πρόκειται για ενα μοντέλο που παράγει μετρική ομοιότητας, σχετικά με το πόσο "μοιάζουν" τα εισερχόμενα στοιχεία σε σχέση με το μέγεθος τους. Να θυμίσουμε πως πλέον έχουμε τις λέξεις (καλλιτέχνες) σε labels και τον αριθμό εμφανίσεων τους ανά genre, κάτι που ερμηνεύεται ως το μέγεθος τους. Με μαθηματική έννοια, υπολογίζει το συνημίτονο της γωνίας μεταξύ δύο vectors σε έναν πολυδιάστατο χώρο. Τα vectors στη προκειμένη είναι οι ομάδες καλλιτεχνών του κάθε μουσικού είδους. Τελικά, επιστρέφει τη μετρική ομοιότητας μεταξύ των δυο, που αποτελεί και τη μετρική ομοιότητας δυο μουσικών ειδών μεταξύ τους.

$$\text{similarity}(A, B) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n A_i^2} \times \sqrt{\sum_{i=1}^n B_i^2}}$$

**Σχήμα 4.21:** Η "κεντρική" μαθηματική εξίσωση μέσω της οποίας εξάγει τις μετρικές του ο αλγόριθμος του cosine similarity.

- Ο δεύτερος συνδυασμός είναι μεταξύ **Tf-Idf Vectorizer & Linear Kernel**. Επιτελούν τον ίδιο ακριβώς σκοπό με τον πρώτο και προαναφερθέντα συνδυασμό, όμως χρησιμοποιούνται για πιοτικότερη εξαγωγή αποτελεσμάτων. Στη προκειμένη περίπτωση, γίνεται χρήση της στατιστικής μεθόδου Term Frequency - Inverse Document Frequency (Tf-Idf) και συγκεκριμένα της συνάρτησης Tf-Idf Vectorizer, ώστε οι δοθείσες πληροφορίες (σειρές ομάδων καλλιτεχνών) να διαχωριστούν, να καταμετρηθούν (πόσες εμφανίσεις μιας λέξης στην ομάδα) και να αποκτήσουν tf-idf τιμές. Έπειτα, μέσω της συνάρτησης linear kernel υπολογίζεται η ομοιότητα των ομάδων καλλιτεχνών, δηλαδή των vectors, που τελικά αντιστοιχούν στα μουσικά είδη. Οπότε, παράγεται μια νέα μετρική εκτίμηση της ομοιότητας μεταξύ των genres του νέου συνόλου δεδομένων.

$$\text{tf}(t, d) = \frac{f_d(t)}{\max_{w \in d} f_d(w)}$$

$$\text{idf}(t, D) = \ln \left( \frac{|D|}{|\{d \in D : t \in d\}|} \right)$$

$$\text{tfidf}(t, d, D) = \text{tf}(t, d) \cdot \text{idf}(t, D)$$

$$\text{tfid}'(t, d, D) = \frac{\text{idf}(t, D)}{|D|} + \text{tfidf}(t, d, D)$$

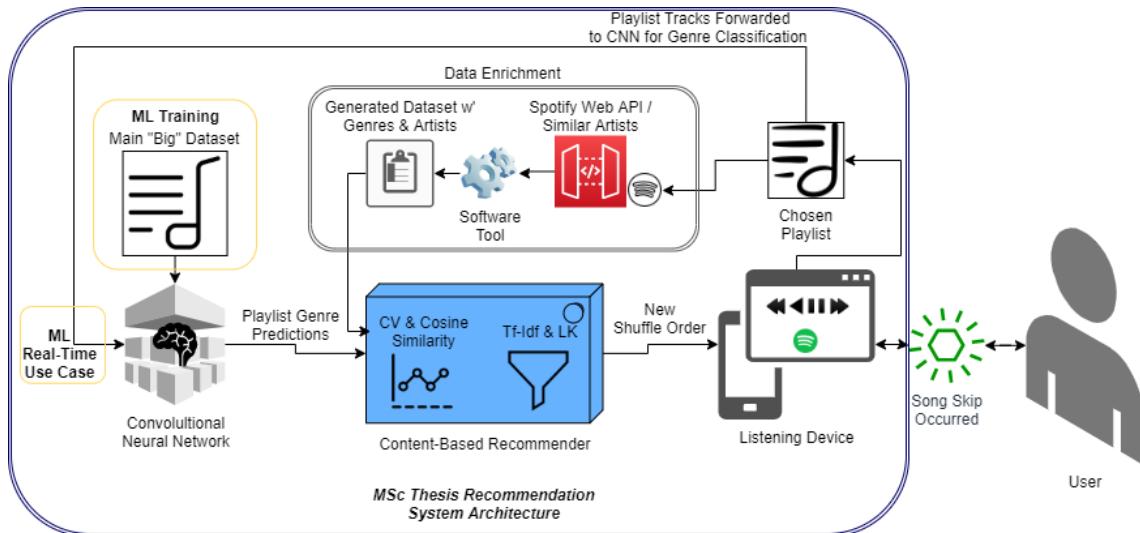
$f_d(t)$  := frequency of term  $t$  in document  $d$

$D$  := corpus of documents

**Σχήμα 4.22:** Οι μαθηματικές εξισώσεις που οδηγούν στον τελικό υπολογισμό της μετρικής  $tf-idf$ .

Τελικά, οι παραχθείσες μετρικές ομοιότητας από τους δύο συνδυασμούς συλλέγονται και υπολογίζεται ο μέσος όρος αυτών, έτσι ώστε η τελική λίστα ομοιοτήτων να είναι όσο το δυνατόν πιο ποιοτική. Αυτή είναι και η τεχνική αρχιτεκτονική λειτουργία του content-based συστήματος συστάσεων της παρούσας διπλωματικής. Στο τελευταίο σκέλος της, χρησιμοποιούνται και οι προβλέψεις του CNN δικτύου που έχει επίσης αναπτυχθεί στα πλαίσια της εργασίας. Αν το προβλεπόμενο μουσικό είδος για ένα κομμάτι "συμφωνεί" με κάποια από τα είδη των καλλιτχενών (που έχουν αντιστοιχίες σε αυτό), τότε αυτό το είδος "ενισχύεται" (πιθανοτικά), κάτι που σημαίνει πως με τη σειρά τους ενισχύονται και οι ομάδες καλλιτεχνών στις οποίες αυτό αντιστοιχεί, κάτι που τελικά σημαίνει ότι προσδίδεται βάρος, ενισχύοντας την πιθανότητα του. **Οπότε, όταν ένας χρήστης προσπεράσει (κάνει skip)** ένα μουσικό κομμάτι από το αρχικό σύνολο δεδομένων μας (αναφερόμαστε πάντα στο "μεσαίο" ή το "μικρό"), θα εντοπίζονται τα μουσικά είδη που του αντιστοιχούν και θα υπολογίζονται οι ομοιότητες εκείνων των ειδών με άλλα (με βάση τη παραχθείσα λίστα ομοιοτήτων από τις ομάδες καλλιτεχνών τους). Τελικά, όσα genres εμφανίζουν θετικές τιμές ομοιότητας με εκείνα του κομματιού αυτού που προσπεράστηκε, θα αφαιρούνται από τη λίστα. Αυτό σημαίνει πως, μαζί τους, θα αφαιρούνται και όλα τα κομμάτια που τα περιλαμβάνουν. Αυτή είναι, με λίγα λόγια, η λειτουργία του **recommender**. Μέσω αρνητικών δειγμάτων (skips), προσπαθεί να συγκλίνει (να αφήσει τελικά) μόνο τα επιθυμητά μουσικά είδη που υποθέτει πως ο χρήστης θέλει να ακούσει εκείνη τη στιγμή. Πόσο γρήγορα μπορεί να καταφέρει τη σύγκλιση στα πλαίσια μιας "μεσαίας" και μιας "μικρής" λίστας αναπαραγωγής; Θα το μάθουμε λίγο παρακάτω, στο κεφάλαιο των αποτε-

λεσμάτων. Τώρα, ας δουμε τα κεντρικά κομμάτια κάθιδια που ενορχηστρώνουν το τεχνικό σκέλος της διπλωματικής εργασίας.



Σχήμα 4.23: Η τελική και ολοκληρωμένη αρχιτεκτονική του συστήματος που παρουσιάζει η παρούσα διπλωματική εργασία. Συνδυασμός των επιλογών του Spotify Web API, των δυνατοτήτων της μηχανικής μάθησης, της λειτουργίας του content-based recommendation και φυσικά της αγαπημένης λειτουργίας του shuffling. Όλα αυτά, σε προσωποποιημένο επίπεδο, για έναν συγκεκριμένη χρήστη, μια συγκεκριμένη χρονική στιγμή και σε πραγματικό χρονο.

## Κεφάλαιο 5

### Υλοποίηση

---

Σε αυτό το κεφάλαιο θα παρουσιαστούν τα βασικά κομμάτια κώδικα που αποτελούν τους κεντρικούς πυλώνες του τεχνικού σκέλους της παρούσας διπλωματικής εργασίας. Αρχικά, παρατίθεται το κομμάτι δημιουργίας του "μεγάλου" και βασικού συνόλου δεδομένων. Έπειτα, ακολουθεί εκείνο του convolutional neural network, δηλαδή του κύριου μοντέλου μηχανικής μάθησης. Τέλος, ολοκληρώνουμε με το σκέλος του content-based συστήματος συστάσεων. Όπως έχει ήδη αναφερθεί, όλος ο τεχνικός τομέας της διπλωματικής έχει αναπτυχθεί σε γλώσσα προγραμματισμού Python, με αξιοποίηση των κατάλληλων βιβλιοθηκών. Φυσικά, το συνολό σύστημα αποτελείται και από αρκετά μικρότερα κομμάτια κώδικα, στα οποία μπορεί κανείς να αποκτήσει πλήρη πρόσβαση κατόπιν επικοινωνίας με τον συγγραφέα της διπλωματικής [61].

#### 5.1 Κώδικας Δημιουργίας Συνόλου Δεδομένων

Ας ξεκινήσουμε με το κομμάτι κώδικα που αφορά τη δημιουργία του μεγάλου συνόλου δεδομένων, μέσω των endpoints του Spotify Web API. Να σημειωθεί πως ο τρόπος συλλογής των δεδομένων έχει ήδη αναλυθεί παραπάνω. Επίσης, η διαδικασία δημιουργίας των δυο μικρότερων datasets (του "μεσαίου" και του "μικρού"), είναι σχεδόν ίδια με τη παρακάτω, με μόνη διαφορά τη ζήτηση συγκεκριμένου αριθμού μουσικών κομματιών και όχι λιστών αναπαραγωγής. Ο λόγος που το μεγάλο σύνολο συλλέχθηκε μέσω της αναζήτησης playlists για κάθε genre, είναι επειδή το Web API του Spotify παρέχει όριο χιλίων μουσικών κομματιών για απευθείας (explicit) ζήτηση, ενώ δεν μπορεί κανείς να διασπάσει το πρόβλημα σε πολλά μικρότερα requests επειδή πάντα θα επιστρέφονται τα ίδια χίλια μουσικά κομμάτια. Τελικά, ο κώδικας διαμορφώνεται ως εξής:

```
1 #the main process for the creation of the big spotify dataset. Request for 80
    ↪ playlists (and their tracks) for every genre.
2 def datasetCreator(genres):
```

```

3     global artist_name
4     global artist_id
5     global track_name
6     global popularity
7     global track_id
8     global checkers
9     global dnc
10    global nrg
11    global ld
12    global spc
13    global acs
14    global ins
15    global lv
16    global vln
17    global tmp
18    global initgenre
19    for gnr in genres:
20        for j in range(0, 80):
21            try:
22                playlist_results = sp.search(q=gnr, type='playlist', market='US
23                                ↪ ', limit=1, offset=j)
24                plist = playlist_results['playlists']['items'][0]
25            except:
26                print("Error occurred at the playlist extraction. Continuing_
27                                ↪ with", len(artist_id), "total elements.")
28            try:
29                plt = 'spotify:playlist:' + plist['id']
30                plst = sp2.playlist(plt)
31                for i, item in enumerate(plst['tracks']['items']):
32                    booler = 0
33                    checkifreached = 0
34                    t = item['track']
35                    if t['artists'][0]['name'] != None:
36                        nn = ''.join(t['artists'][0]['name'])
37                        nn = nn.lower()
38                        artist_name.append(nn)
39                    else:
40                        artist_name.append("")
41                    ider = t['artists'][0]['id']
42                    artist_id.append(ider)

```

```

41         if ider != None:
42             urn = 'spotify:artist:' + ider
43             artist = sp2.artist(urn)
44             ge = ""
45             # addition of empty strings at the begining and the end
46             #   ↪ of each genre, in order to ease the split
47             #   ↪ function later on
48             for genre in artist['genres']:
49                 prep = genre.lower()
50                 prep = prep.replace("_", " ")
51                 ge = ge + "_" + prep
52             ge = ge + "_"
53             checkers.append(ge)
54         else:
55             checkers.append('')
56             tnn = ''.join(t['name'])
57             tnn = tnn.lower()
58             track_name.append(tnn)
59             track_id.append(t['id'])
60             ttt = t['id']
61             gggenre = gnr.lower()
62             gggenre = gggenre.replace("_", " ")
63             initgenre.append(gggenre)
64             checkifreached = 1
65             features = sp2.audio_features(tracks=[ttt])
66             if features[0] != None:
67                 booler = 1
68                 dnc.append(features[0]['danceability'])
69                 nrg.append(features[0]['energy'])
70                 ld.append(features[0]['loudness'])
71                 spc.append(features[0]['speechiness'])
72                 acs.append(features[0]['acousticness'])
73                 ins.append(features[0]['instrumentalness'])
74                 lv.append(features[0]['liveness'])
75                 vln.append(features[0]['valence'])
76                 tmp.append(features[0]['tempo'])
77             else:
78                 booler = 1
79                 dnc.append(mean(dnc))
80                 nrg.append(mean(nrg))

```

```

79         ld.append(mean(ld))
80         spc.append(mean(spc))
81         acs.append(mean(acs))
82         ins.append(mean(ins))
83         lv.append(mean(lv))
84         vln.append(mean(vln))
85         tmp.append(mean(tmp))
86     except:
87         print("An_error_occured_later_on._The_progress_is_being_"
88             "→ continued_with", len(artist_id), "total_elements.")
89         if booler == 0 and checkifreached == 1:
90             dnc.append(mean(dnc))
91             nrg.append(mean(nrg))
92             ld.append(mean(ld))
93             spc.append(mean(spc))
94             acs.append(mean(acs))
95             ins.append(mean(ins))
96             lv.append(mean(lv))
97             vln.append(mean(vln))
98             tmp.append(mean(tmp))
99         continue

```

Αρχικά, γίνεται “σύνδεση” (χρήση της global) των λιστών στις οποίες θα αποθηκεύονται τα δεδομένα για κάθε κομμάτι, σε κάθε επανάληψη. Πρόκειται για λίστες που θα αποθηκεύουν το όνομα και το id του τραγουδιού & του καλλιτέχνη του, τα μουσικά είδη του καλλιτέχνη, καθώς και τα μελωδικά χαρακτηριστικά. Έπειτα, ξεκινάει η κεντρική επανάληψη των 80 λιστών αναπαραγωγής για κάθε ένα από τα δοθέντα genres. Σε κάθε λίστα, ξεκινάει εντός της μια νέα επανάληψη, για όσα μουσικά κομμάτια εκείνη περιέχει. Για κάθε μουσικό κομμάτι, γεμίζουν και οι λίστες που δηλώθηκαν αρχικά. Κάθε λίστα λαβαίνει μια πρόσθεση στοιχείου (εκείνη του ονόματος δέχεται το όνομα, εκείνη του καλλιτέχνη δέχεται τον καλλιτέχνη κλπ), για το κομμάτι που βρίσκεται εντός της επανάληψης. Λαβαίνονται τα κατάλληλα μέτρα για να αποφεχθεί η διακοπή των επαναλήψεων, μέσω try & except, ενώ στο τέλος γίνεται έλεγχος για το αν προκλήθηκε όντως κάποιο error που ενδέχεται να οδήγησε στο γέμισμα μόνο κάποιων εκ των στηλών. Έτσι, εξαλείφεται και ο κίνδυνος ανόμοιου “γεμίσματος” των στηλών, κάτι που φυσικά θα κατέστρεφε την εγγυρότητα και τη ποιότητα του τελικού συνόλου δεδομένων. Τέλος, αν κάποια εκ των μελωδικών χαρακτηριστικών λείπουν ή δεν έχουν καταχωρηθεί για ένα κομμάτι (πχ λόγω error πιο πριν), τότε συμπληρώνονται οι τιμές των μέσων όρων από τις υπάρχουσες τιμές. Όταν ολοκληρωθεί αυτό το σκέλος κώδικα, οι λίστες ενώνονται σε ένα dataframe και

αποθηκεύονται σε ένα .csv αρχείο (αυτά τα δυο βήματα δεν φαίνονται εδώ).

## 5.2 Κώδικας Δημιουργίας Convolutional Neural Network

Προχωρώντας στη δημιουργία του 2D Convolutional Neural Network, παρακάτω παρουσιάζονται τα κομμάτια της προ-επεξεργασίας των δεδομένων για την είσοδο στο δίκτυο, το κεντρικό δίκτυο, η εντολή fitting αυτού, καθώς και η τελική εντολή εκπαίδευσης του. Τα δεδομένα έχουν επεξεργαστεί ήδη σε μεγάλο βαθμό όταν φτάνουμε στο σημείο που φαίνεται παρακάτω, δηλαδή εκείνο της επιπρόσθετης προ-επεξεργασίας. Απλώς το σημείο αυτό αφορά αποκλειστικά τη μετατροπή τους για είσοδο στο CNN. Πριν τον κώδικα, αξίζει να γίνει ιδιαίτερη μνεία στο άρθρο του Datacamp.com, με συγγραφέα τον Aditya Sharma [62], το οποίο παραθέτει αναλυτικά και περιεκτικά τη μεθοδολογία δημιουργίας ενός Convolutional νευρωνικού δικτύου. Υπάρχουν αμέτρητα άρθρα στο διαδίκτυο για τη κατασκευή μοντέλων μηχανικής μάθησης. Επίσης όμως, ο κύκλος σπουδών του μεταπτυχιακού προγράμματος "Επιστήμη Δεδομένων & Μηχανική Μάθηση" (στο οποίο και "ανήκει" η παρούσα διπλωματική) παρείχε τις απαραίτητες γνώσεις για την εξειδικευμένη ανάπτυξη neural networks. Όμως, γίνεται αναφορά στο συγκεκριμένο άρθρο του Datacamp, λόγω της απλότητας με την οποία ανέλυσε κάποια συγκεκριμένα στοιχεία, συμβάλλοντας στην ολοκλήρωση του μοντέλου της διπλωματικής. Παρακάτω λοιπόν μπορείτε να δείτε πως διαμορφώνεται το custom νευρωνικό δίκτυο που αναπτύχθηκε στα πλαίσια της παρούσας εργασίας:

```

1
2 #reshape the sets generated from train_test_split() function
3 train_X = X_train.reshape(-1, 3, 3, 1)
4 test_X = X_test.reshape(-1, 3, 3, 1)
5 train_Y = y_train
6 test_Y = y_test
7
8 #change the type and prepare for proper input to the cnn
9 train_X = train_X.astype('float32')
10 test_X = test_X.astype('float32')
11 train_X = train_X / 255.
12 test_X = test_X / 255.
13
14 # Change the labels from categorical to one-hot encoding
15 train_Y_one_hot = to_categorical(train_Y)
16 test_Y_one_hot = to_categorical(test_Y)

```

```

17
18 # the final sets for training
19 train_X,valid_X,train_label,valid_label = train_test_split(train_X,
20   ↪ train_Y_one_hot, test_size=0.2, random_state=13)
21
22 #initialize batch size, epoches and nclasses
23 bs = 8
24 eps = 60
25 nclasses = len(np.unique(lb))
26
27 #the 2d CNN model
28 fmodel = Sequential()
29 fmodel.add(Conv2D(32, kernel_size=(3, 3),activation='linear',padding='same',
30   ↪ input_shape=(3,3,1)))
31 fmodel.add(LeakyReLU(alpha=0.1))
32 fmodel.add(MaxPooling2D((2, 2),padding='same'))
33 fmodel.add(Dropout(0.25))
34 fmodel.add(Conv2D(64, (3, 3), activation='linear',padding='same'))
35 fmodel.add(LeakyReLU(alpha=0.1))
36 fmodel.add(MaxPooling2D(pool_size=(2, 2),padding='same'))
37 fmodel.add(Dropout(0.25))
38 fmodel.add(Conv2D(128, (3, 3), activation='linear',padding='same'))
39 fmodel.add(LeakyReLU(alpha=0.1))
40 fmodel.add(Flatten())
41 fmodel.add(Dense(128, activation='linear'))
42 fmodel.add(LeakyReLU(alpha=0.1))
43 fmodel.add(Dropout(0.35))
44 fmodel.add(Dense(nclasses, activation='softmax'))
45
46 #compile the model above
47 fmodel.compile(loss=keras.losses.categorical_crossentropy, optimizer=keras.
48   ↪ optimizers.Adam(),metrics=['accuracy'])
49
50 #print the model summary if needed
51
52 #start the training process (and also print it)
53 fmodelTrain = fmodel.fit(train_X, train_label, batch_size=bs,epochs=eps,verbose

```

```

    ↢ =1, validation_data=(valid_X, valid_label))

54
55 #evaluate the model and get the results
56 test_eval = fmodel.evaluate(test_X, test_Y_one_hot, verbose=1)
57 print('Test_loss:', test_eval[0])
58 print('Test_accuracy:', test_eval[1])

```

Σε πρώτη φάση λαβαίνει χώρα το τελικό στάδιο προ-επεξεργασίας των δεδομένων, κατά το οποίο τα train & test sets (παράγωγα της train\_test\_split), μετασχηματίζονται, ενώ γίνεται αλλαγή και στον τύπο τους. Επίσης, αλλάζουν και τα labels ώστε να είναι σύμφωνα με το σύστημα κωδικοποίησης one-hot. Έπειτα, έχουμε τη δημιουργία των τελικών sets για τη διαδικασία της εκπαίδευσης του μοντέλου. Ακολουθεί το 2D Convolutional Neural Network μοντέλο μας, του οποίου η αρχιτεκτονική έχει καταγραφεί και απεικονιστεί στο προηγούμενο κεφάλαιο. Πλέον, μένει το compilation, το fitting και τελικά το evaluation του δημιουργηθέντος μοντέλου. Η προσεκτική παρατήρηση των αποτελεσμάτων του fitting / training και της αξιολόγησης βοήθησε στο να γίνουν οι κατάλληλες παραμετροποιήσεις στο μοντέλο, ώστε να επιτευχθεί το καλύτερο δυνατό αποτέλεσμα. Το κομμάτι της μηχανικής μάθησης ολοκληρώνεται με τη φόρτωση του "μεσαίου" & "μικρού" συνόλου δεδομένων, τη παραγωγή προβλέψεων για αυτά και την αποθήκευση τους, ώστε να χρησιμοποιηθούν στο σύστημα συστάσεων μας, του οποίου ο κώδικας παρουσιάζεται στον ακριβώς επόμενο υπότιτλο (τα τελευταία αυτά βήματα δεν περιλαμβάνονται ως κώδικας παραπάνω).

## 5.3 Κώδικας Δημιουργίας Συστήματος Συστάσεων

Το τελικό κομμάτι είναι εκείνο του συστήματος συστάσεων. Πρόκειται για το σκέλος που αξιοποιεί και τα αποτελέσματα του CNN μοντέλου μηχανικής μάθησης που παρουσιάστηκε παραπάνω. Συγκεκριμένα, παρουσιάζεται η συνάρτηση αξιολόγησης του συστήματος, η οποία δέχεται μια λίστα αναπαραγώγης και επιχειρεί να συγκλίνει σε μια κατηγορία ειδών μουσικής, με βάση τα skips. Πρόκειται για μια προσομοίωση ενός σεναρίου πραγματικού χρόνου που θα γινόταν από έναν άνθρωπο. Η μόνη επιπρόσθετη πληροφορία που παρέχεται στο σύστημα, είναι το μουσικό είδος που "έχει στο νου του" ένας άνθρωπος εκείνη τη στιγμή. Η πληροφορία αυτή παραμένει μόνο στο υψηλό επίπεδο του συστήματος, για να ελέγχεται το αν ένα μουσικό κομμάτι ανήκει σε αυτή η όχι. Άρα, η διαδικασία που πραγματοποιείται έχει ως εξής: **Το σύστημα αξιολόγησης ξεκινάει μια διαδικασία τυχαίας αναπαραγώγης κομματιών (από το "μεσαίο" ή το "μικρό" σύνολο δεδομένων), ελέγχει αν το τραγούδι που προέκυψε "συμφωνεί" μουσικά με το δοθέν είδος που "σκέφτεται" ο υπολογιστής και δρά ανάλογα με τη περίσταση. Αν το τραγούδι ανήκει σε αυτή τη μουσική κατηγορία, συνεχίζει με εκ νέου τυχαία διαλογή.**

Στη περίπτωση που το επιλεχθέν κομμάτι δεν ανήκει στη δοθείσα κατηγορία, τότε γίναι `skip` σε αυτό. Αυτομάτως, για κάθε μουσικό είδος που αντιστοιχεί σε αυτό το κομμάτι, καλείται η συνάρτηση σύστασης. Όσα μουσικά είδη έχουν θετικό ποσοστό ομοιότητας (με βάση το μέσο όρο των `count vectorizer / cosine similarity & tf-idf vectorizer / linear kernel`) με εκείνα που αντιστοιχούν στο κομμάτι, αφαιρούνται από τη λίστα. Οπότε, μαζί τους αφαιρούνται και όλα τα μουσικά κομμάτια που τους αντιστοιχούν. Κάπου εδώ γίνεται και η αξιοποίηση των προβλέψεων του CNN. Αν η πρόβλεψη του νευρωνικού δικτύου για το συγκεκριμένο κομμάτι συμπίπτει με κάποιο είδος από εκείνα που του αντιστοιχούν (που στη πραγματικότητα είναι τα μουσικά είδη του καλλιτέχνη του), τότε η μέθοδος αφαίρεσης των μουσικών ειδών ενισχύεται και επιταχύνεται. Όλα τα παραπάνω λαβαίνουν χώρα μέχρι να υπάρξουν 10 συνεχόμενες αναπαραγωγές κομματιών χωρίς να γίνει `skip` κάπου, ή μέχρι να ακουστούν 17 μουσικά κομμάτια, που ισοδυναμούν περίπου σε μια ώρα αναπαραγωγής. Τελικά, παράγονται κάποιες μετρικές που θα μας βοηθήσουν να αξιολογήσουμε τα δεδομένα. Ο κώδικας είναι ο ακόλουθος:

```

1 def performanceTesting(chosengenre):
2     global mainer
3     global csimilarity
4     global lkernel
5     global playlist
6     plst = playlist
7     totall = len(plst['trackname'])
8     isokay = 0
9     consplays = 0
10    fiveconsecs = 0
11    fccc = 0
12    plays = 0
13    skips = 1
14    unlock = 0
15    reason = ""
16    cgenres = []
17    try:
18        cgenres = get_recommendations(chosengenre, csimilarity, lkernel)
19    except:
20        # when initial genre is not in the enriched list. Rare, but happens
21        sim = [s for s in playlist['genres'] if chosengenre in s]
22        grs = sim[0].split('_')
23        similargenre = grs[1]
24        cgenres = get_recommendations(similargenre, csimilarity, lkernel)

```

```

25
26     while (isokay == 0):
27         ml = 0
28         #smart function for when there is no good progress
29         if(skips == 18 and plays <= 6):
30             if (unlock == 0):
31                 print("UNLOCK_TRIGGERED")
32                 unlock = 1
33                 sim = [s for s in playlist['genres'] if chosengenre in s]
34                 grs = sim[0].split('_')
35                 similargenre = grs[1]
36                 cgenres = get_recommendations(similargenre, csimilarity,
37                     ↪ lkernel)
37                 # get a random track from the playlist and check if it 'agrees' with
38                 ↪ the given genre
38                 ras = list(plst['trackname'])
39                 randtrack = random.choice(ras)
40                 ss = plst.loc[plst['trackname'] == randtrack, 'genres'].iloc[0].split(
41                     ↪ '_')
41                 mg = plst.loc[plst['trackname'] == randtrack, 'main_genre'].iloc[0]
42                 predicted = plst.loc[plst['trackname'] == randtrack, 'predicted'].iloc
43                     ↪ [0]
43                 ss = list(filter(None, ss))
44                 keepit = 1
45                 for cg in cgenres:
46                     if cg in ss:
47                         keepit = 0
48
49                 if mg in cgenres:
50                     keepit = 0
51
52                 # if the random track is of the same genre as the given one, move on.
53                 ↪ If not, then get recommendations for similar genres to that and
54                 ↪ remove their songs
53                 if keepit == 0:
54                     fiveconsecs = fiveconsecs + 1
55                     consplays = consplays + 1
56                     plays = plays + 1
57                     if fiveconsecs == 5:
58                         fccc = fccc + 1

```

```

59             fiveconsecs = 1
60             # conditions that terminate the process
61             if consplays == 10:
62                 reason = "Ten_consecutive_plays_occurred."
63                 isokay = 1
64             if plays == 17:
65                 reason = "Average_listening_session_duration_ended_(1_hour)."
66                 isokay = 1
67             else:
68                 for s in ss:
69                     # check if the machine learning prediction is correct (meaning
                           ↪ if the predicted genre is part of one of the song
                           ↪ artist's genres)
70                     pre = predicted.split("/")
71                     for pr in pre:
72                         if pr in s:
73                             print("Machine_Learning_Prediction_Correct!")
74                         ml = 1
75             # if the prediction is correct, then boost the removal process
76             if ml == 1:
77                 siml = [s for s in playlist['genres'] if 'alternative' in s]
78                 for si in siml:
79                     grss = si.split('_')
80                     for gr in grss:
81                         if predicted in gr:
82                             plst = plst[~plst.genres.str.contains(gr)]
83             # remove the skipped songs genres (which means, their songs as well
                           ↪ ), as well as the... genres similar to those genres (and
                           ↪ their songs)
84             for s in ss:
85                 try:
86                     sets = get_recommendations(s, csimilarity, lkernel)
87                     for st in sets:
88                         gen = "_" + st + "_"
89                         plst = plst[~plst.genres.str.contains(gen)]
90                 except:
91                     continue
92             try:
93                 sets = get_recommendations(mg, csimilarity, lkernel)
94                 for st in sets:

```

```

95             gen = " " + st + " "
96             plst = plst[~plst.genres.str.contains(gen)]
97         except:
98             continue
99
100            fiveconsecs = 0
101            consplays = 0
102            skips = skips + 1
103
104    # print the results
105    print("Performance_testing_ended!_Main_Genre_preference:", chosengenre,
106          "→ with", len(cgenres), "similar_genres._Here_are_the_stats:_->")
107    print("Reason_for_termination_of_the_performance_testing_session:", reason)
108    print("Total_skips:", (skips-1), "._Total_plays:", plays, "._Five_"
109          "→ consecutive_plays_for:", fccc, "times.")
110    print("Play/_Skip_Ratio_per_Hour:", round((plays / skips), 2), "PSh")
111    print("Songs_to_listen_to:", len(plst['trackname']), "from", totall, "total"
112          "→ _tracks." )
113
114    return plays, skips, round((plays / skips), 2), fccc, unlock

```

Η αρχική ανάλυση της λειτουργίας του κώδικα αναλύει και με τη σειρά τον τρόπο εκτέλεσης του. Ας εξηγήσουμε λοιπόν τις παραχθείσες μετρικές, μαζί με έναν επιπρόσθετο μηχανισμό:

- **Unlock Trigger:** Πρόκειται για έναν μηχανισμό επίσπευσης της διαδικασίας αξιολόγησης, για τις περιπτώσεις που η εξέλιξη είναι πάρα πολύ αργή. Αν ο χρήστης έχει στο μυαλό του ενα είδος που είναι πιο σπάνιο, τότε (όπως είναι λογικό) ο αριθμός των skips που απαιτούνται ώστε ο αλγόριθμος να συγκλίνει προς τα εκείνο, θα είναι περισσότερα. Αν λοιπόν η πρόοδος είναι πολύ αργή και τα skips των κομματιών πολλαπλασιάζονται χωρίς ουσιαστικό λόγο, τότε ενεργοποείται ο μηχανισμός "unlock", ο οποίος εντοπίζει μουσικά είδη που λεκτικά ταιριάζουν με το δοθέν (ακόμα και αν στη πραγματικότητα δεν είναι "κοντά" μελωδικά), προσθέτοντας και αυτά στο... παιχνίδι, ως κεντρικά επιλεχθέντα είδη. Έτσι, διασφαλίζεται το γεγονός πως η διαδικασία δεν θα διαιωνίζεται, αλλά θα ολοκληρώνεται μετά από κάποια στιγμή.
- **Plays:** Η πρώτη εκ των τεσσάρων μετρικών που παράγονται. Πρόκειται για τον αριθμό των μουσικών κομματιών που πρόλαβαν να παιχθούν (δεν έγιναν skip επειδή "συμφωνούσαν" με το δοθέν είδος) πριν την ολοκλήρωση της αξιολόγησης. Υπενθυμίζουμε πως 17 τραγούδια αντιστοιχούν σε μια ώρα ακρόασης, οπότε συνήθως αυτή θα είναι και η τιμή που θα επιστρέφεται στο τέλος (εφόσον η ολοκλήρωση μιας ώρας ακρόασης είναι συνθήκη τερματισμού).

- **Skips:** Η δεύτερη εκ των τεσσάρων μετρικών, πρόκειται απλώς για το σύνολο των μουσικών κομματιών που προσπεράστηκαν (έγιναν skip) κατά τη διάρκεια της αξιολόγησης.
- **Play / Skip Ratio Per Hour:** Πρόκειται για το αποτέλεσμα της διαίρεσης των plays και των skips για κάθε γύρο αξιολόγησης, δηλαδή της διαίρεσης των δυο παραπάνω μετρικών. Εφόσον μια από τις δυο συνθήκες τερματισμού είναι η μια ώρα αναπαραγωγής, η αναλογία play / skip θεωρείται ωριαία, εξού και ο ορισμός "per hour". 'Οσο πιο υψηλός είναι ο αριθμός, τόσο το καλύτερο, εφόσον αυτό θα σημαίνει πως χρειάστηκαν λίγα skips για να εμφανίζονται κομμάτια που θέλει ο χρήστης.
- **Five Consecutive Plays:** Η τελευταία μετρική που παράγει η αξιολόγηση μας. Πρόκειται για το σύνολο των περιπτώσεων στις οποίες δεν χρειάστηκε skip για πέντε συνεχόμενες αναπαραγωγές κομματιών. Πρακτικά, αυτό σημαίνει πως για περίπου 15 λεπτά ο χρήστης δεν χρειάστηκε να προσπεράσει ούτε ένα τραγούδι, κάτι το ιδιαίτερα ενθαρρυντικό για την απόδοση του μοντέλου!

Για κάθε ένα εκ των 19 βασικών genres που έχουν οριστεί στην αρχή της παρούσας αναφοράς, η αξιολόγηση περιλαμβάνει δέκα ίδιους γύρους. Οπότε, στο τέλος γίνεται συλλογή δεδομένων για δέκα γύρους αξιολογήσεων. Ποια είναι τελικά τα αποτελέσματα; Ας τα αναλύσουμε παρακάτω, στο 6ο και προ-τελευταίο κεφάλαιο της διπλωματικής εργασίας..!

## Κεφάλαιο 6

### Αποτελέσματα

Σε αυτό το κεφάλαιο παρουσιάζονται τα αποτελέσματα και οι επιδόσεις του συνολικού τεχνικού συστήματος της διπλωματικής εργασίας. Αρχικά θα καταγραφούν οι επιδόσεις των μοντέλων μηχανικής μάθησης, ενώ στη συνέχεια θα εξεταστούν τα αποτελέσματα της αξιολόγησης από το σύστημα συστάσεων.

#### 6.1 Μηχανική Μάθηση

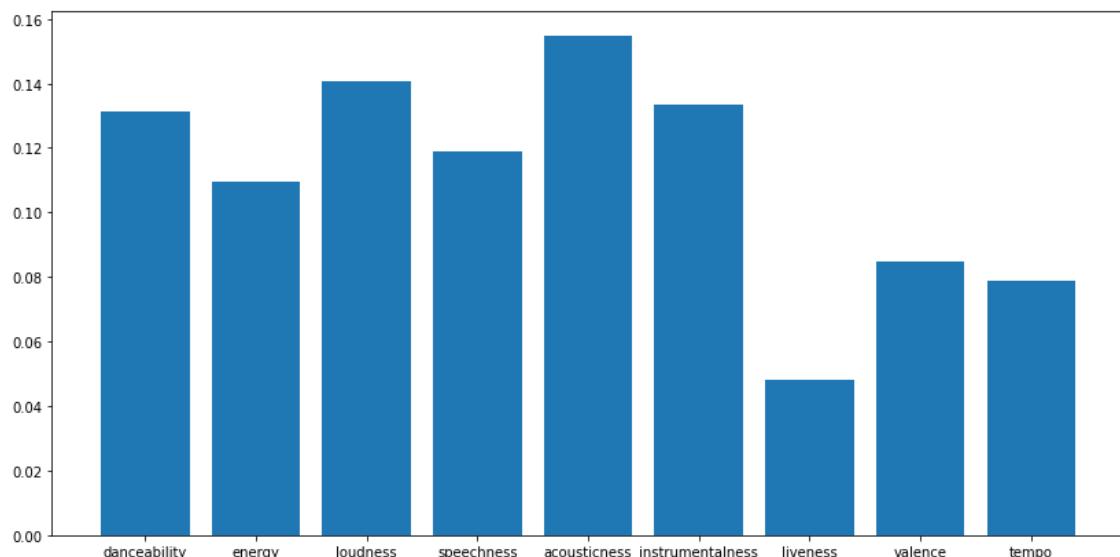
Πρωτού παραθέσουμε τον πίνακα με τις επιδόσεις των ταξινομητών μηχανικής μάθησης που επιλέχθηκαν (και αναλύθηκαν στο τέταρτο κεφάλαιο), ας λάβουμε μερικές επιπρόσθετες πληροφορίες σχετικά με το σύνολο δεδομένων μας. Όπως έχει ήδη αναφερθεί πολλάκις, στο σκέλος της μηχανικής μάθησης έγινε χρήση του "μεγάλου" και βασικού συνόλου δεδομένων. Ας δουμε αν υπάρχουν συσχετίσεις μεταξύ των 9 μελωδικών χαρακτηριστικών, αλλά και την... άποψη του XGBoost ταξινομητή για τη σπουδαιότητα του κάθε feature στο ρόλο έκβασης του τελικού αποτελέσματος.

	danceability	energy	loudness	speechness	acousticness	instrumentalness	liveness	valence	tempo
danceability	1.000000	0.146849	0.286529	0.230233	-0.222778	-0.241231	-0.076341	0.467993	-0.093493
energy	0.146849	1.000000	0.811351	0.131618	-0.770930	-0.399943	0.205253	0.369068	0.223651
loudness	0.286529	0.811351	1.000000	0.141251	-0.669851	-0.572199	0.145716	0.344526	0.186982
speechness	0.230233	0.131618	0.141251	1.000000	-0.123895	-0.170017	0.081364	0.097250	0.059980
acousticness	-0.222778	-0.770930	-0.669851	-0.123895	1.000000	0.378641	-0.127381	-0.271613	-0.188573
instrumentalness	-0.241231	-0.399943	-0.572199	-0.170017	0.378641	1.000000	-0.097986	-0.304595	-0.098052
liveness	-0.076341	0.205253	0.145716	0.081364	-0.127381	-0.097986	1.000000	0.049560	0.040220
valence	0.467993	0.369068	0.344526	0.097250	-0.271613	-0.304595	0.049560	1.000000	0.087599
tempo	-0.093493	0.223651	0.186982	0.059980	-0.188573	-0.098052	0.040220	0.087599	1.000000

Σχήμα 6.1: Πίνακας συσχετίσεων μελωδικών χαρακτηριστικών του κεντρικού dataset.

Παρατηρώντας αρχικά τον πίνακα συσχετίσεων, εντοπίζουμε υψηλό ποσοστό μόνο μεταξύ των energy & loudness. Όλες οι υπόλοιπες τιμές κυμαίνονται κάτω του 0.5,

με μοναδική περίπτωση που τείνει να το αγγίζει να είναι η συσχέτιση μεταξύ valence & danceability. Το υψηλό ποσοστό μεταξύ έντασης και ενέργειας μπορεί να θεωρηθεί ως αρκετά λογικό, εφόσον τα μουσικά κομμάτια με υψηλή ένταση, τείνουν να έχουν και πιο ενεργό μελωδικό ρυθμό. Πρόκειται για ένα ασφαλές συμπέρασμα, που μπορεί να πιστοποιηθεί από τη συγκεκριμένη παρατήρηση. Όμως, δεν υπάρχει κάποια άλλη υπόθεση που μπορούμε να κάνουμε, με βάση πάντα τον παραπάνω πίνακα συσχετίσεων. Ας εξετάσουμε και τη περίπτωση του ιστογράμματος σπουδαιότητας, το οποίο παρέχεται από τον XGBoost ταξινομητή:



**Σχήμα 6.2:** Ιστόγραμμα σπουδαιότητας μελωδικών χαρακτηριστικών του κεντρικού dataset, μέσω του κατηγοριοποιητή XGBoost.

Αν είχαμε να συλλέξουμε τα τέσσερα χαρακτηριστικά με την υψηλότερη σπουδαιότητα, αυτά θα ήταν (με φθίνουσα σειρά) τα acousticness, loudness, instrumentalness & danceability. Φαίνεται πως πρόκειται για τα features που επηρεάζουν περισσότερο τη λειτουργία του ταξινομητή. Επιπρόσθετα, παρατηρούμε πως τα χαρακτηριστικά valence, tempo & liveness απέχουν αισθητά από τις υπόλοιπες τιμές, ιδιαίτερα το τρίτο. Έχοντας στο νου μας το ιδιαίτερα χαμηλό ποσοστό σπουδαιότητας που λαβαίνει το liveness, αλλά και την διαπίστωση πως δεν φέρει καμια υψηλή συσχέτιση με κάποιο άλλο feature, θα μπορούσαμε να το αφαιρέσουμε από το σύνολο δεδομένων μας, μένοντας τελικά με 8 χαρακτηριστικά. Παρόλα αυτά, η εκπαίδευση και αξιολόγηση των μοντέλων μηχανικής μάθησης έγινε με το πλήρες σύνολο, αφού φυσικά έλαβε πρώτα χώρα η απαραίτητη προ-επεξεργασία. Αφού λοιπόν είδαμε τα διαγράμματα συσχετίσεων και σπουδαιότητας, αποκτώντας μια ακόμα μεγαλύτερη οικειότητα με το dataset μας, ας δουμε τον πίνακα αποτελεσμάτων ακρίβειας των machine learning κατηγοριοποιητών. Θυμίζουμε πως οι έξι επιλεχθέντες αλγόριθμοι είναι οι **Naive Bayes**, **Logistic Regression**, **k-Nearest Neighbors**, **Random Forest**, **Decision Tree** και φυσικά το αναπτυχθέν **2D Convolutional Neural Network**.

**work.** Η εκπαίδευση έχει γίνει για πέντε διαφορετικές περιπτώσεις: **i) χρήση 2 τυχαίων μουσικών ειδών (άρα και 2 labels), ii) χρήση 3 τυχαίων μουσικών ειδών (άρα και 3 labels), iii) χρήση 4 τυχαίων μουσικών ειδών (άρα και 4 labels), iv) χρήση συγχωνευμένων μουσικών ειδών και συγκεκριμένα 12 στον αριθμό (άρα και 12 labels), v) χρήση του πλήρους εύρους μουσικών ειδών, δηλαδή 19 στον αριθμό (άρα και 19 labels).** Να σημειωθεί πως, ο τρόπος με τον οποίο έγινε η συγχώνευση των ειδών, έγινε κατόπιν έρευνας σχετικά με τις μελωδικές ομάδες ειδών που μοιάζουν και θα μπορούσαν να ταιριαστούν. Εκ των 19 genres, έγινε συγχώνευση σε 7 περιπτώσεις (r&b / soul, hip-hop / rap, indie / alternative, punk / metal, folk / country, blues / jazz και latin / dance). Ας δούμε τον πίνακα αποτελεσμάτων ακρίβειας:

ML MODELS	AMOUNT OF GENRES TO CLASSIFY				
	TWO*	THREE*	FOUR*	12 (MERGED)	19 (COMPLETE)
NAIVE BAYES	98%	62%	70%	33%	21%
LOGISTIC REGRESSION	98%	66%	73%	34%	26%
k NEAREST NEIGHBOURS	98%	66%	73%	33%	25%
RANDOM FOREST	98%	72%	77%	40%	29%
DECISION TREE	97%	62%	69%	32%	22%
XGBOOST	99%	70%	77%	40%	30%
2D CONVOLUTIONAL NN	98%	69%	76%	34%	25%

	ACCURACY				
	<i>*random genre selection</i>				

Σχήμα 6.3: Ο πλήρης πίνακας αποτελεσμάτων ακρίβειας των υποποιηθέντων μοντέλων μηχανικής μάθησης, με χρήση του πλήρους dataset και εναλλαγή στον αριθμό των μουσικών ειδών (labels) προς κατηγοριοποίηση.

Τα αποτελέσματα που παρατηρούμε στον παραπάνω πίνακα είναι άκρως ενδιαφέροντα. Αρχικά, μπορούμε με ασφάλεια να συμφωνήσουμε ότι ο ταξινομητής XGBoost έχει δίκαια τόσο... καλή φήμη ανά των μοντέλων μηχανικής μάθησης, αφού κατάφερε να πετύχει τα υψηλότερα ποσοστά ακρίβειας (σχεδόν) σε όλες τις δοκιμές ειδών. Επίσης, ο χρόνος ολοκλήρωσης ήταν αρκετά σύντομος, επιβεβαιώνοντας τις αναφορές περι εξαίρετης αποδοτικότητας. Μαζί του στις επιδόσεις των αποτελεσμάτων είναι και ο Random Forest, ένας ταξινομητής γνωστός για τα καλά του στοιχεία. Μάλιστα, στη περίπτωση της κατηγοριοποίησης τριών μουσικών ειδών, κατάφερε να ξεπεράσει τον XGBoost κατά 2 τοις εκατό. Το τρίτο σε απόδοση μοντέλο είναι το δημιουργηθέν convolutional neural network δυο διαστάσεων. Στο σύστημα έγιναν αρκετές δοκιμές και παραμετροποιήσεις, ώστε να επιτευχθούν τα καλύτερα δυνατά αποτελέσματα. Όμως, το CNN είναι εκ φύσεως ένα μοντέλο που αποδίδει καλύτερα με άλλα είδη δεδομένων, κυρίως εικόνων. Η απόδοση του δεν είναι καθόλου κακή, βρίσκεται σταθερά ο τρίτος καλύτερος ταξινομητής σε όλες τις περιπτώσεις (με εξα-

ίρεση τη ταξινόμηση των 19 labels που "έχασε" από το μοντέλο παλινδρόμησης για 1 τοις εκατό), όμως δεν έφτασε τα ποσοστά των XGBoost & Random Forest. Η Λογιστική Παλινδρόμηση (Logistic Regression) παρέμεινε σε ποσοστά ακρίβειας κοντινά του CNN, ενώ ακολουθούν τα k-NN, Naive Bayes & Decision Tree με τις χαμηλότερες συνολικά αποδόσεις. Να σημειωθεί πως οι αποκλίσεις μεταξύ της ακρίβειας των μοντέλων δεν είναι υψηλές. Η μεγαλύτερη εντοπίζεται στη κατηγοριοποίηση των 19 ειδών, μεταξύ του Naive Bayes και του XGBoost. Για άλλη μια φορά, να τονίσουμε πως το Convolutional νευρωνικό δίκτυο έτεινε να παραμείνει πιο κοντά στις επιδόσεις των δυο κορυφαίων μοντέλων.

Τώρα, ας εστιάσουμε στα συμπεράσματα μας για το σύνολο δεδομένων. Αντιλαμβανόμαστε πως οι τιμές που παρέχει η δημοφιλής υπηρεσία ροής μουσικής, Spotify, μέσω των Web API endpoints της, δεν μπορούν από μόνες τους να εγγυηθούν υψηλά αποτελέσματα στην ακρίβεια μοντέλων μηχανικής μάθησης. Τα μελωδικά χαρακτηριστικά των μουσικών κομματιών αποδεικνύονται ένα "δύσκολο" dataset για εκμάθηση, παρά τη λεπτομερή και προσεγμένη προ-επεξεργασία σε αυτά. Μάλιστα, δοκιμάστηκαν διάφορες επιλογές προ-επεξεργασίας, σε μια προσπάθεια αύξησης της ακρίβειας των αποτελεσμάτων. Όμως, κάτι τέτοιο δεν έγινε. Τα δεδομένα αυτά δεν επαρκούν για ένα επιτυχημένο genre recognition. Μάλιστα, τα αποτελέσματα του πίνακα απεικονίζουν τη, όχι και τόσο καλή, ποιότητα των δεδομένων. Στη πρώτη περίπτωση ταξινόμησης, εκείνη των δυο ειδών, όλοι οι αλγόριθμοι πετυχαίνουν υψηλά επίπεδα ακρίβειας. Προσθέτοντας όμως μόλις ένα genre (label), παρατηρείται μια αισθητή πτώση των αποτελεσμάτων σε όλους τους ταξινομητές. Και ενώ θα περίμενε κανείς πως η περαιτέρω αύξηση των labels θα μείωνει ακόμα περισσότερο την ακρίβεια, το παράξενο είναι πως η κατάσταση μοιάζει να διαφοροποείται με τη προσθήκη ενός ακόμα μουσικού είδους. Τα ποσοστά ακρίβειας ανεβαίνουν, όντας καλύτερα σε όλα τα μοντέλα. Σε αυτό το σημείο, να υπενθυμίσουμε για άλλη μια φορά πως έχει γίνει τυχαία επιλογή μουσικών ειδών για κάθε περίπτωση. Οι πολλαπλές δοκιμές με τυχαία είδη φαίνεται να ενισχύουν το συμπέρασμα πως η ποιότητα των δεδομένων δεν είναι καλή. Κάθε φορά τα αποτελέσματα ακρίβειας μπορεί να είχαν αρκετές διαφορές μεταξύ τους, κάτι που σημαίνει πως μερικά μουσικά είδη επηρεάζουν το αποτέλεσμα αρκετά περισσότερο από κάποια άλλα.

Κατόπιν περαιτέρω έρευνας πάνω σε αυτό το συμπέρασμα, εκείνο μοιάζει πλέον να θεωρείται βέβαιο. Ουσιαστικά, η έρευνα έδωσε απάντηση στο ερώτημα "Γιατί δεν είναι καλή η ποιότητα των δεδομένων;" Πέραν του γεγονότος πως άλλες μορφές μουσικών datasets μπορούν να έχουν καλύτερα αποτελέσματα στην κατηγοριοποίηση (όπως αναφέρθηκε στο τρίτο κεφάλαιο), τι μπορεί να κυλάει λάθος για το συγκεκριμένο σύνολο; Η απάντηση είναι απλή: Η έλλειψη αυθεντικών μουσικών ειδών για κάθε κομμάτι και το είδος των μελωδικών χαρακτηριστικών. Για το πρώτο σκέλος, θα μπορούσε κανείς να πει πως τα δεδομένα έχουν συλλεχθεί μέσω αναζήτησης με βάση συγκεκριμένα μουσικά είδη. Οπότε, τι θα μπορούσε να πάει στραβά; Δυστυχώς, το

πρόβλημα φαίνεται να είναι πως η αναζήτηση με βάση ένα "κεντρικό" μουσικό είδος (που θα μπορούσε δηλαδή να χαρακτηριστεί και οικογένεια ειδών) είναι αναζήτηση " υψηλού επιπέδου " (high level), σε σημείο που ενδέχεται χάνεται η ποιότητα αυτής. Το ακόλουθο παράδειγμα αποτυπώνει απόλυτα τη κατάσταση ως έχει: Το μουσικό είδος alternative, αντιπροσωπεύει μια συγκεκριμένη μελωδική κατηγορία, που μάλιστα οι περισσότεροι άνθρωποι έχουν έμμεσα ή άμεσα διατυπωμένη στο νου τους. Κατά την αναζήτηση μουσικών κομματιών (ή λιστών αναπαραγωγής) που να σχετίζονται με το alternative genre, κανείς θα περίμενε πως θα συνέλλεγε μουσικά κομμάτια της μελωδικής έννοιας που έχει στο μυαλό του και ταυτίζει με τον όρο alternative. Όμως, το Spotify μοιάζει να μην λειτουργεί έτσι. Κατά την αναζήτηση λιστών αναπαραγωγής με βάση το alternative, θα συλλεχθούν πολλά και διαφορετικά μουσικά κομμάτια, τα οποία θεωρούνται όντως εναλλακτικά, αλλά, μεταξύ άλλων, και για διαφορετικά μουσικά είδη. Δηλαδή, θα συλλεχθούν και κομμάτια alternative metal, alternative pop, alternative hip-hop και άλλα, που στη πραγματικότητα όμως, μελωδικά, δεν μοιάζουν καθόλου μεταξύ τους. Αυτό δυστυχώς οδηγεί σε μείωση της ποιότητας των δεδομένων, αφού δυσκολεύει το έργο των ταξινομητών για ουσιαστική αξιοποίηση των μελωδικών χαρακτηριστικών ώστε να μπορέσουν να εκπαιδευτούν ορθά.

Αυτή η παρατήρηση δεν ισχύει σε όλα τα μουσικά είδη. Είναι εμφανώς μειωμένη σε genres όπως το rock ή το hip-hop. Είναι όμως έκδηλη στο classical, το οποίο καταλήγει να αποτελεί συλλογή κλασικών μεν κομματιών, αλλά όχι μόνο εκείνων της κλασικής μουσικής που έχουμε όλοι στο νου μας. Για παράδειγμα, περιλαμβάνονται και τραγουδια classical rock. Κάπως έτσι, καθίσταται ιδιαίτερα δύσκολη η δημιουργία ενός πραγματικά ποιοτικού συνόλου δεδομένων με τη χρήση του Spotify Web API. Βέβαια, θα μπορούσε κανείς να προχωρήσει σε εξαντλητική έρευνα, να καταγράψει μερικές χιλιάδες κομμάτια από κάθε είδος που να το αντιπροσωπεύουν πραγματικά, εν συνεχείᾳ να "περάσει" τις καταγραφές του σε ένα κομμάτι προγραμματιστικού κώδικα και τελικά να αναζητήσει τα κομμάτια αυτά μέσω του Web API. Όμως, στο ενδιάμεσο θα πρέπει να έχει μάθει και το id τους στο Spotify, ώστε να μπορέσει να τα αναζητήσει. Πρόκειται για μια τρομακτικά χρονοθόρα διαδικασία, που ενδεχομένως να αποβεί και πάλι άχρηστη, λόγω της δυσκολίας που φέρουν εκ φύσεως τα δεδομένα αυτά για σωστή κατηγοριοποίηση (όπως αναφέρθηκε ήδη) και αποτελεί το δεύτερο σκέλος της απάντησης. Τα δεδομένα που λαβαίνουμε για κάθε μουσικό κομμάτι, δηλαδή τα μελωδικά του χαρακτηριστικά, δεν επαρκούν για ορθή και επιτυχημένη κατηγοριοποίηση. Χρειαζόμαστε καλύτερου (ή διαφορετικού) είδους δεδομένα που να περιγράφουν το κάθε τραγούδι, όπως για παράδειγμα είναι τα ηχητικά αρχεία, τα σπεκτογράμματα, ο συνδυασμός αυτών κλπ. Παρατηρώντας στον πίνακα αποτελεσμάτων τη σταδιακή κατακρύμνη των ποσοστών ακρίβειας των μοντέλων, αντιλαμβανόμαστε πως τα μελωδικά χαρακτηριστικά δεν παρέχουν " σταθερότητα " τέτοια ώστε με αυτά να εκπαιδευτούν ταξινομητές για genre recognition.

Αν και μας φτάνουν σε ένα σημείο, για να πάμε παραπέρα μάλλον θα πρέπει να αλλάξουμε τύπο δεδομένων, κάτι το οποίο θα μπορούσε να εξεταστεί από το Spotify.

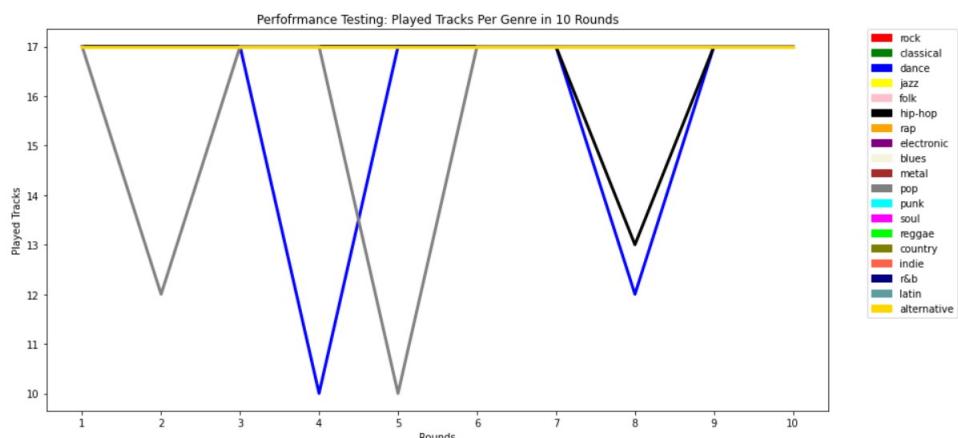
Δεν μπορούμε να πούμε με απόλυτη βεβαιότητα ποιο από τα δυο σκέλη ευθύνεται κυρίως για τα χαμηλά ποσοστά επιτυχίας, αλλά έχουμε μια ισχυρή ένδειξη πως είναι το είδος των χαρακτηριστικών (μελωδικά). Διεξάγοντας κάποιες δοκιμές και χρησιμοποιώντας πιο "συγκεκριμένα" genres για να λάβουμε δεδομένα, αλλά αισθητά λιγότερα σε μέγεθος, υπήρξε μια μέση βελτίωση της τάξης του 5% (στο καλύτερο σενάριο) από κάποια ML μοντέλα, άλλα δεν βελτιώθηκαν καθόλου, ενώ κάποια τα πήγαν χειρότερα. Άρα, φαίνεται πως τα "high level" μουσικά είδη συμβάλλουν στο πρόβλημα, αλλά δεν έχουν το μεγαλύτερο μερίδιο ευθύνης. Στη προκειμένη περίπτωση όμως, τουλάχιστον για το πρώτο σκέλος της απάντησης του "Τι δεν κυλάει ομαλά ;", η λύση που θα μπορούσε να δοθεί είναι μόνο μία: Κάποια στιγμή, το Spotify θα πρέπει να μεριμνήσει και να προμηθεύσει τουλάχιστον ένα genre (ως πληροφορία) σε κάθε τραγούδι. Αυτό θα συμβάλλει κατά τα μέγιστα στη βελτίωση της ποιότητας μελλοντικών ερευνών & υλοποίησεων τεχνικών μηχανικής μάθησης πάνω στα σύνολα δεδομένων που δημιουργούνται μέσω του Web API της δημοφιλούς υπηρεσίας ροής μουσικής. Αν γνωρίζαμε το πραγματικό μουσικό είδος για κάθε κομμάτι, τότε το έργο της παρούσας διπλωματικής εργασίας θα ήταν απείρως πιο εύκολο. Θα προσπερνούσαμε το κομμάτι του genre recognition (οπότε αυτομάτως θα λυνόταν και το δεύτερο σκέλος της απάντηση, δηλαδή εκείνο της χρήσης μελωδικών χαρακτηριστικών στη κατηγοριοποίηση, αφού δεν θα χρειαζόμασταν καν classification), θα μπορούσαμε να κάνουμε πιο σωστή ανάλυση πάνω στα μελωδικά χαρακτηριστικά, ακόμα και αν πλέον δεν χρειαζόταν, ενώ φυσικά θα αναμέναμε να λάβουμε προφανώς καλύτερα αποτελέσματα ακρίβειας από τους ταξινομητές. Πάλι δεν θα λαβαίναμε ποσοστά της τάξεως του 90 τοις εκατό, όπως πιθανόν να συνέβαινε με καλύτερα σύνολα δεδομένων, σαν τα μουσικά datasets ηχητικών αρχείων, σίγουρα όμως αισθητά βελτιωμένες αποδόσεις.

Γιατί αναφέρουμε τη φράση "αισθητά βελτιωμένες αποδόσεις"; Τελικά, όπως φαίνεται στις δυο τελευταίες δοκιμές, δηλαδή εκείνες με το μερικώς συγχωνευμένο (12 είδη) και το πλήρες (19 είδη) σύνολο δεδομένων, τα αποτελέσματα ακρίβειας είναι αποκαρδωτικά. Για τα 12 labels, το καλύτερο ποσοστό ανήκει στους XGBoost & Random Forest με 40%. Είναι φυσικά καλύτερο από μια τυχαία ταξινόμηση (1 / 12 ισούται με 8%), αλλά σίγουρα δεν μπορεί να προσφέρει ποιοτικές και αξιοποιήσιμες προβλέψεις. Όσον αφορά το πλήρες σύνολο με 19 labels, καλύτερη επίδοση ήταν (ξανά) του XGBoost με 30%. Φυσικά, εξακουθεί να είναι καλύτερο αποτέλεσμα από τη τυχαία κατηγοριοποίηση (1 / 19 ισούται με 5%), αλλά ο δρόμος βελτίωσης είναι εμφανέστατος. Σε κάθε περίπτωση, τα δυο μικρότερα σύνολα δεδομένων, το "μεσαίο" και το "μικρό", φορτώθηκαν στο σύστημα και, μέσω του CNN (που δεν είχε τη καλύτερη επίδοση από τα μοντέλα αλλά προτιμήθηκε λόγω του ότι είναι... σπιτικό), έγινε παραγωγή προβλέψεων μουσικών ειδών. Οι προβλέψεις αποθηκεύτηκαν και

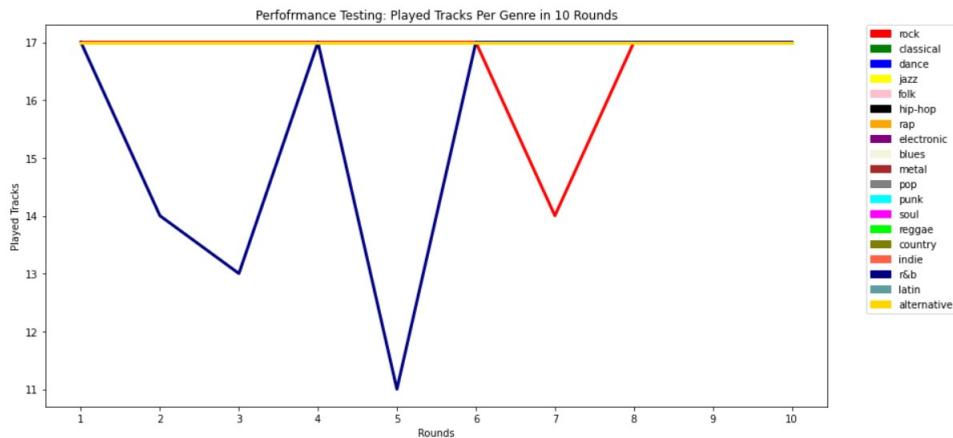
χρησιμοποιήθηκαν στο τελικό σύστημα συστάσεων. Ας ολοκληρώσουμε το κεφάλαιο των αποτελεσμάτων, αλλά και τα βασικά μέρη της παρούσας διπλωματικής εργασίας, βλέποντας σε διαγράμματα τα αποτελέσματα της αξιολόγησης του συστήματος συστάσεων. Οι υποθέσεις και οι έρευνες που κάναμε και προαναφέρονται σε αυτές τις παραγράφους, επιβεβαιώνονται επίσης από τις επιδόσεις μερικών genres στον recommender. Πριν περάσουμε λοιπόν στον επίλογο με τα τελικά συμπεράσματα και τις μελλοντικές επεκτάσεις, ας δουμε τα παραχθέντα αποτελέσματα του τελικού μας συστήματος. **Τελικά, μπορεί ένα σύστημα συστάσεων να συγκλίνει στις επιθυμητές μουσικές προτιμήσεις του χρήστη, σε πραγματικό χρόνο, με αξιοποίηση μόνο των αρνητικών δειγμάτων;** Ας το μάθουμε...

## 6.2 Σύστημα Συστάσεων

Φτάσαμε στη... στιγμή της αλήθειας. Ας εξετάσουμε τις επιδόσεις του συστήματος συστάσεων μας, στο οποίο χρησιμοποιήθηκαν μόνο αρνητικά δείγματα, δηλαδή τα skips του χρήστη. Υπενθυμίζουμε πως κάθε αξιολόγηση είχε δέκα γύρους, εξού και οι τιμές [1,10] στον άξονα των xx' των διαγραμμάτων. Οπότε, κάθε μουσικό είδος πέρασε από το μοντέλο αξιολόγησης δέκα φορές. Υπενθυμίζουμε επίσης πως τα αποτελέσματα αφορούν μια περίοδο ακρόασης μουσικής, διάρκειας μιας ώρας. Αυτό το όριο τέθηκε διότι έτσι μπορεί να γίνει λόγος σε αποτελέσματα πραγματικού χρόνου, αλλά και επειδή 60 λεπτά μπορούν να θεωρηθούν ως η μέση διάρκεια περιόδου ακρόασης μουσικής από ενα χρήστη. Βέβαια, αν ο χρόνος επιμυκηνθεί, το σύστημα συστάσεων θα εξακολουθεί να δουλεύει. Απλώς έπρεπε να τεθεί ένα όριο στα πλαίσια της αξιολόγησης. Θα γίνεται παρουσίαση δυο διαγραμμάτων, ένα του "μικρού" συνόλου δεδομένων και ένα του "μεσαίου", για κάθε μια από τις μετρικές Plays, Skips, Play / Skip Ratio & Five Consecutive Skips. Ξεκινάμε με τη πρώτη μετρική, εκείνη των Plays.

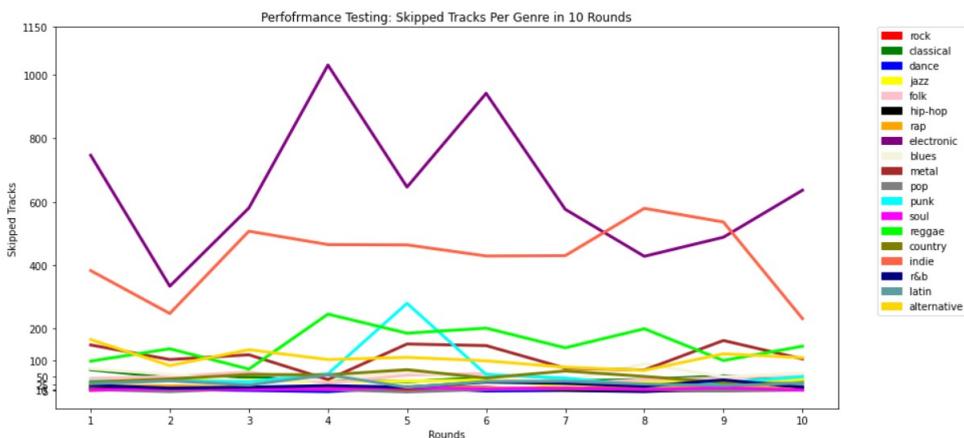


Σχήμα 6.4: Διάγραμμα Συνολικών Αναπαραγωγών "μικρού" συνόλου.

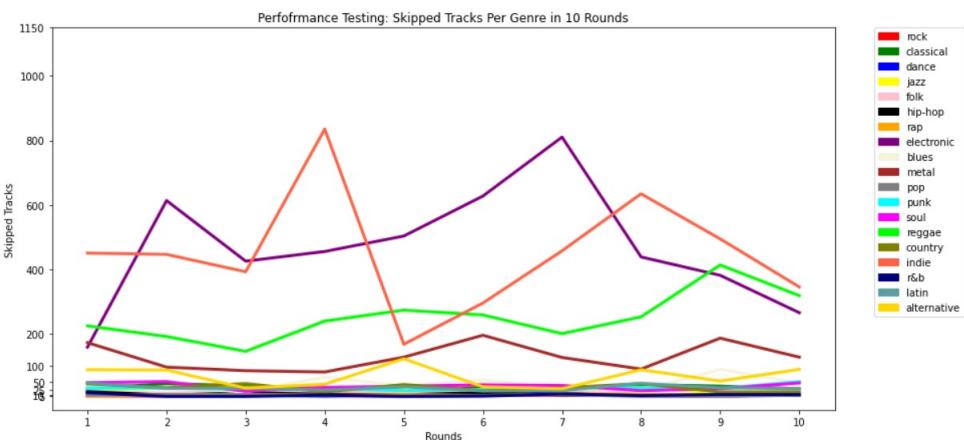


Σχήμα 6.5: Διάγραμμα Συνομικών Αναπαραγωγών "μεσαίου" συνόλου.

Στη συγκεκριμένη μετρική, τα λιγότερα μουσικά κομμάτια (δηλαδή τα λιγότερα plays - αναπαραγωγές) είναι και το καλύτερο. Γιατί αυτό; Όπως έχουμε αναφέρει ήδη αρκετές φορές, μια ώρα μουσικής ακρόασης ισοδυναμεί περίπου σε 17 μουσικά κομμάτια. Οπότε, όσα genres από τα 19 που φαίνονται και στο διάγραμμα έχουν πετύχει 17 αναπαραγωγές, αυτό σημαίνει πως εξάντλησαν τη μια ώρα αναπαραγωγής. Μέχρι εδώ, όλα καλά. Όμως, αυτή ήταν μια από τις δυο συνθήκες τερματισμού, η δεύτερη συνθήκη ήταν οι δέκα συνεχόμενες αναπαραγωγές, χωρίς να προκύψει ούτε ένα skip στο ενδιάμεσο τους. Αν κάτι τέτοιο συμβεί, αυτό σημαίνει πως ο recommender κατάφερε να συγκλίνει στο επιθυμητό μουσικό είδος εξαιρετικά γρήγορα, οπότε ο χρήστης δεν χρειαζόταν πλέον να προσπερνάει μουσικά κομμάτια που δεν ήθελε. Μπορεί λοιπόν (για παράδειγμα) να ακουσει δυο μουσικά κομμάτια, να έκανε δυο skips και μετά να μην χρειάστηκε να παρέμβει ποτέ. Οπότε, 2 τα αρχικά τραγούδια που άκουσε και 10 τα επόμενα (στα οποία και ενεργοποιείται η συνθήκη τερματισμού), ισοδυναμούν με 12, ένα αποτέλεσμα που στη πραγματικότητα είναι καλύτερο του 17! Παρατηρώντας λοιπόν τα διαγράμματα, αριθμούμε μόνο πέντε genres που κατάφεραν να ενεργοποιήσουν τη "δύσκολη" δεύτερη συνθήκη τερματισμού. Στο μικρό σύνολο, έχουμε το pop (γκρι χρώμα) σε δυο γύρους, το dance (μπλέ χρώμα) επίσης σε δυο γύρους και το hip-hop (μαύρο χρώμα) σε έναν γύρο. Στο μεσαίο σύνολο, έχουμε το r&b (navy μπλέ χρώμα - τα χρώματα παραμένουν ίδια για τα genres σε όλα τα διαγράμματα), σε τρεις γύρους, ενώ την εμφάνιση της κάνει και η rock (κόκκινο χρώμα) σε έναν γύρο. Μπορούμε να εξάγουμε κάποια γρήγορα συμπεράσματα, τα οποία τείνουν να συμφωνήσουν με εκείνα που αναφέραμε παραπάνω, αλλά πρώτα ας προχωρήσουμε και στα διαγράμματα των skips.



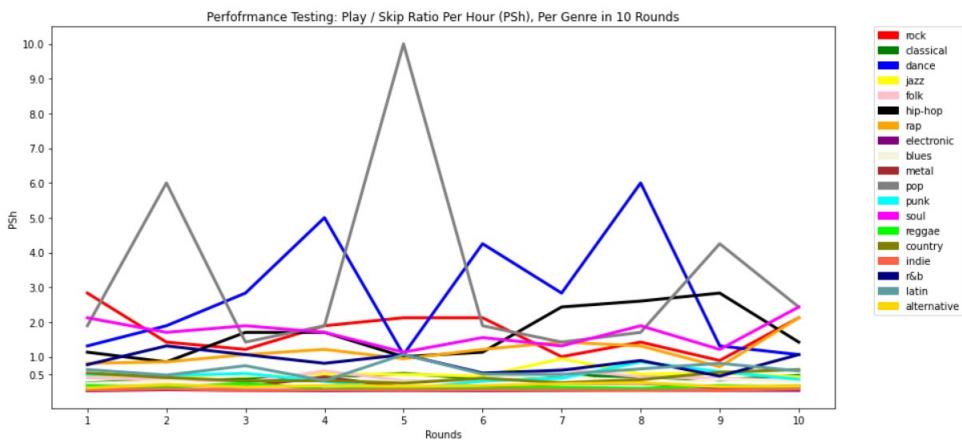
Σχήμα 6.6: Διάγραμμα Συνολικών Skips “μικρού” συνόλου.



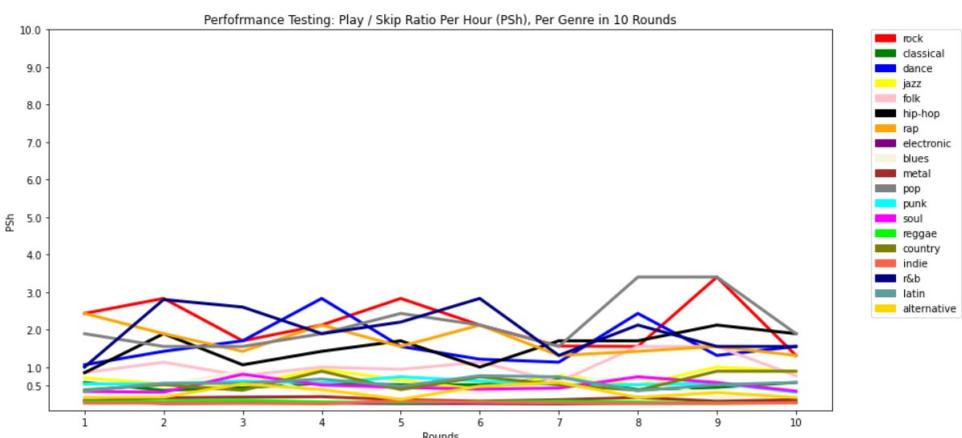
Σχήμα 6.7: Διάγραμμα Συνολικών Skips “μεσαίου” συνόλου.

Εδώ μπορούμε να αποκτήσουμε μια πολύ καλύτερη εικόνα της απόδοσης του κάθε μουσικού είδους. Εξίσου βοηθητικά θα είναι και τα διαγράμματα του Play / Skip Ratio παρακάτω. Παρατηρώντας προσεκτικά, θα δούμε πως οι επιδόσεις των μουσικών ειδών κυμαίνονται στα ίδια επίπεδα, ανεξαρτήτως συνόλου δεδομένων. Για παράδειγμα, τα είδη electronic & indie (μωβ και “ντοματινί” χρώμα αντίστοιχα) κάνουν... υψηλές πτήσεις και στα δυο διαγράμματα, απαιτώντας εκατοντάδες skips σε κάθε γύρο για να μπορέσει να λειτουργήσει η αξιολόγηση αυτών. Επιπρόσθετα, η γραμμή της reggae (ανοιχτό πράσινο χρώμα) μοιάζει ως η τρίτη χειρότερη, υποδεικνύοντας το αυξημένο πλήθος skips που επίσης χρειάστηκε το σύστημα για να κάνει προσπάθειες σύγκλισης. Αναφορικά με τη punk, ενώ στο μικρό σύνολο ο πέμπτος της γύρος δεν ήταν καθόλου καλός, σε γενικές γραμμές διατηρήθηκε κάπως χαμηλά, ενώ στο μεσαίο σύνολο έμεινε μονίμως σε τιμές κάτω των 50 προσπεράσεων. Άλλα genres που δεν τα πανε καλά στο διάγραμμα του μεσαίου dataset είναι τα metal (κεραμιδί χρώμα) και alternative (ανοιχτό πορτοκαλί χρώμα). Οι υποθέσεις και τα συμπεράσματα που έχουμε ήδη κάνει, φαίνεται να ισχυροποιούνται ακόμα

περισσότερο. Ας τα πιστοποιήσουμε και με τα διαγράμματα του Play / Skip Ratio για τα δυο σύνολα δεδομένων μας.



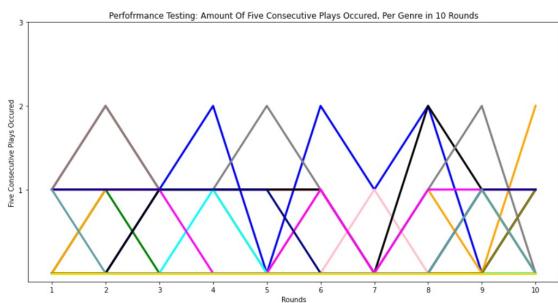
Σχήμα 6.8: Διάγραμμα Play / Skip Ratio Ανά 'Ωρα (PSh) "μικρού" συνόλου.



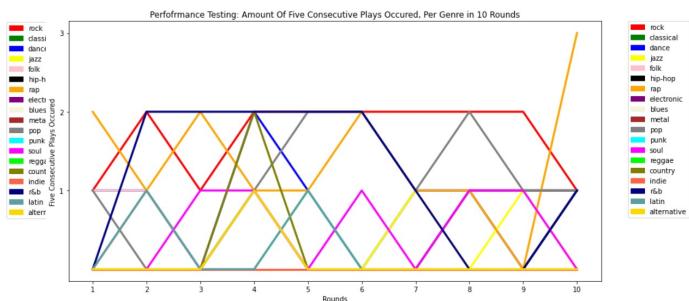
Σχήμα 6.9: Διάγραμμα Play / Skip Ratio Ανά 'Ωρα (PSh) "μεσαίου" συνόλου.

Σε αυτά τα διαγράμματα αντιλαμβανόμαστε πως, όσο υψηλότερη είναι η τιμή, τόσο το καλύτερο. Ένα υψηλό PSh υποδεικνύει πως, σε διάστημα μιας ώρας, χρειάστηκαν λίγα skips κομματιών, αναλογικά με το σύνολο εκείνων που αναπαράχθηκαν κανονικά. Όπως γίνεται εύκολα αντιληπτό, αν ο αριθμός PSh είναι μικρότερος της μονάδας (1), αυτό σημαίνει πως (σε διάστημα μιας ώρας) τα skips ήταν περισσότερα από τα plays, κάτι όχι και τόσο καλό. Στο διάγραμμα του μικρού συνόλου δεδομένων, παρατηρούμε τα pop, dance, rock, hip-hop, soul & rap να τα πηγαίνουν αρκετά καλά, με τα δυο πρώτα μάλιστα να αυξάνουν τον πήχη κατά πολύ. Η πλειοψηφία των μουσικών ειδών όμως κυμαίνεται κοντά στο 1, ενώ λίγα τείνουν στο μηδέν. Φυσικά, μια PSh μετρική κοντά στο 1 εξακολουθεί να είναι καλό αποτέλεσμα. Πάντως, μπορούμε να είμαστε σίγουροι τα genres που τείνουν στο μηδέν με το ελάχιστο ποσοστό είναι και εκείνα με τον ιδιαίτερα υψηλό αριθμό skips προηγουμένως, δηλαδή τα electronic & indie. Στο διάγραμμα του μεσαίου συνόλου, παρατηρούνται

χαμηλότερες τιμές, αλλά τα είδη που πριν είχαν καλές επιδόσεις, φαίνεται πως τις διατηρούν. Το genre που φαίνεται να τα πηγαίνει καλύτερα σε αυτό το dataset είναι η r&b μουσική. Από εκεί και πέρα, τα υπόλοιπα είδη κυμαίνονται στις ίδιες χαμηλές τιμές. Όμως, είναι αρκετά ενθαρρυντικό το γεγονός πως έχουμε, έστω λίγα, genres με "υπολογίσιμες" PSh μετρικές. Πλέον, μετά και την παρατήρηση των δυο αυτών διαγραμμάτων, μπορούμε να καταλήξουμε και να επικυρώσουμε τα συμπεράσματα μας. Ας το κάνουμε αυτό αμέσως μετά τη παρουσίαση των διαγραμμάτων Five Consecutive Plays, το οποίο δεν αποκαλύπτει πολλά, αλλά επιβαθμίνει πως κάποια (για την ακρίβεια αρκετά) μουσικά είδη κατάφεραν να αποσπάσουν περιόδους στις οποίες δεν χρειάστηκε skip για τουλάχιστον πέντε κομμάτια.



Σχήμα 6.10: Διάγραμμα Πέντε Συνεχόμενων Αναπαραγωγών "μικρού" συνόλου.



Σχήμα 6.11: Διάγραμμα Πέντε Συνεχόμενων Αναπαραγωγών "μεσαίου" συνόλου.

Οι υποθέσεις και τα συμπεράσματα που έγιναν μέσα από την ανάλυση του "μεγάλου" συνόλου δεδομένων και των αποτελεσμάτων ακρίβειας των μοντέλων μηχανικής μάθησης, επικυρώνονται με την παρατήρηση των διαγραμμάτων της αξιολόγησης του recommender. Κάποια μουσικά είδη πετυχαίνουν εξαιρετικές επιδόσεις, κάτι μάλιστα που αποδεικνύει πως το σύστημα συστάσεων της παρούσας διπλωματικής δουλεύει κανονικά! Μάλιστα, οι επιδόσεις των περισσότερων genres είναι "υπέρ" της θετικής ψήφου στο μοντέλο που αναπτύχθηκε μέσα από αυτή την εργασία. Όμως, η κακή ποιότητα δεδομένων σε κάποια μουσικά είδη, λόγω των διαφόρων κομματιών που συλλέχθηκαν "κάτω από την ομπρέλα" τους, δημιουργεί ένα ιδιαίτερα θολό τοπίο για αυτά, κάτι που αποτυπώνεται και στις επιδόσεις τους. Οπότε, ενώ στη περίπτωση της Μηχανικής Μάθησης αυτό δεν επηρεάζει τόσο τα αποτελέσματα (μάλλον μεγαλύτερη επιρροή έχει το είδος των χαρακτηριστικών που χρησιμοποιούνται για τη κατηγοριοποίηση), φαίνεται πως στον content-based recommender παίζει αισθητά πιο σημαντικό ρόλο. Για παράδειγμα, η electronic μουσική που είχε ίσως τη χειρότερη επίδοση στο σύστημα συστάσεων, περιέχει κομμάτια αρκετών διαφορετικών ειδών, τα οποία απλώς είχαν κοινά "ηλεκτρονικά" στοιχεία. Το ίδιο ισχύει και για τη reggae. Οπότε, αυτά τα είδη θα μπορούσαν να θεωρηθούν ως μεγάλα "outliers", που στρέφουν τη προσοχή πάνω τους και μας απομακρύνουν από την επικύρωση του αναπτυχθέντος συστήματος συστάσεων ως μια επιτυχία. Γιάυτό, θα τα θεωρήσουμε

ακριβώς ως τέτοια. **Εφόσον η πλειοψηφία των μουσικών ειδών υποδεικνύει πως ο recommender μπορεί να λειτουργήσει με χρήση αρνητικών δειγμάτων και τελικά όντως να συγκλίνει στο επιθυμητό είδος (σε πραγματικό χρόνο), ενώ γνωρίζουμε πως ο λόγος κακής συμπεριφοράς μερικών genres είναι λόγω προβληματικής ποιότητας των δεδομένων τους (και όχι κάποιο πρόβλημα του συστήματος) τότε μπορούμε με σιγουριά να πουμε πως η έρευνα μας πέτυχε !**

## Μέρος **III**

### Επίλογος

---



## Κεφάλαιο 7

# Συμπεράσματα & Επεκτάσεις

---

### 7.1 Συμπεράσματα

Τελικά ; Τα... καταφέραμε ; Ας θυμηθούμε το βασικό ερώτημα που πραγματεύεται η παρούσα διπλωματική εργασία : **Μπορεί ένα σύστημα συστάσεων, υλοποιηθέν εντός μιας λίστας αναπαραγωγής, να συγκλίνει στις επιθυμητές μουσικές προτιμήσεις του χρήστη, σε πραγματικό χρόνο, με αξιοποίηση μόνο των αρνητικών δειγμάτων;** Η απάντηση είναι πως **Ναι!**. Οι δοκιμές που διεξήχθησαν απέδειξαν πως είναι εφικτό να υπάρξει σύγκλιση στην επιθυμητή μουσική κατηγορία μέσω αρνητικών δειγμάτων και μάλιστα με σχετικά λίγα skips (από 0 έως 20), και όλα αυτά εντός 60 λεπτών (στη χειρότερη περίπτωση). Ναι, ένα σύστημα συστάσεων μπορεί να λειτουργήσει μόνο με αρνητικά δείγματα και να παράξει αξιόπιστες προτιμήσεις στον χρήστη, σε μια συγκεκριμένη χρονική στιγμή ! Όμως, υπάρχει ένας βασικός παράγοντας για να μπορέσει να λειτουργήσει σωστά η διαδικασία. Το σύνολο δεδομένων, δηλαδή οι πληροφορίες που συνοδεύουν τα μουσικά κομμάτια, θα πρέπει να είναι σωστές, επαρκείς και ποιοτικές. Παρατηρήσαμε genres στα οποία επιτεύχθηκε σύγκλιση με εξαιρετικά αποτελέσματα. Άλλα είδη όμως είχαν απογοητευτικές επιδόσεις. Κατά πάσα πιθανότητα, αυτό συνέβη επειδή το μουσικό είδος μέσω του οποίου συλλέχθηκαν τα κομμάτια, τελικά έφερε τραγούδια πολλών διαφορετικών ειδών, τα οποία όμως θα μπορούσαν να συγκαταλεχθούν στην ίδια "οικογένεια" με το αρχικό. Για τα είδη που είναι εκ φύσεως καλύτερα ορισμένα όμως, όπως είναι για παράδειγμα τα rap & hip-hop, οι επιδόσεις ήταν εξαιρετικές. Επίσης, ένα άλλο συμπέρασμα είναι πως, δυστυχώς, τα μελωδικά χαρακτηριστικά που παρέχει το Spotify Web API δεν μπορούν να θεωρηθούν ως η καλύτερη επιλογή για genre recognition. Άλλα είδη μουσικών συνόλων δεδομένων, τα οποία περιέχουν ηχητικά αρχεία ή σπεκτογράμματα (και τα οποία φυσικά χαίρουν διαφορετικής επεξεργασίας), μπορούν να αποφέρουν καλύτερα και ποιοτικότερα αποτελέσματα, σύμφωνα με σχετικές υλοποιήσεις & δημοσιεύσεις. Τα 9 χαρακτηριστικά που παρέχει η διάσημη υπηρεσία ροής μουσικής για τα τραγούδια της, δυστυχώς δεν φέρνουν την... άνοιξη. Σε κάθε περίπτωση, τα αποτελέσματα ταξινόμησης του υλοποιηθέντος

2D Convolutional Neural Network, ακόμα και αν υστερούσαν σε ακρίβεια, αξιοποιήθηκαν εντός του κεντρικού συστήματος συστάσεων, έτσι ώστε αν η πρόβλεψη genre του CNN για ένα μουσικό κομμάτι συμφωνούσε με κάποιο είδος από εκείνα που του αντιστοιχούσαν μέσω του καλλιτέχνη του, να επιταχύνεται η διαδικασία του recommendation. Οπότε, παρά τη χαμηλή ακρίβεια, υπήρξε ουσιαστική χρήση μηχανικής μάθησης εντός του recommender, κάτι που σημαίνει πως μπορεί κάλλιστα να χαρακτηριστεί και **υθριδικός!**

## 7.2 Μελλοντικές Επεκτάσεις

Ένα άλλο βασικό συμπέρασμα της έρευνας που διεξήχθη στα πλαίσια της παρούσας διπλωματικής εργασίας, είναι πως **κρίνεται επιτακτική η ανάγκη ενσωμάτωσης ετικετών μουσικών ειδών αποκλειστικά για τα κομμάτια**, από το Spotify, ώστε οι μελλοντικές έρευνες και δημοσιεύσεις να μπορέσουν να εξάγουν ακόμα πιο ασφαλή συμπεράσματα, εφόσον θα έχουν συλλέξει πιο ποιοτικά δεδομένα. Έτσι, κάτι 1διαίτερα σημαντικό, θα μπορέσουν να γίνουν και ορθότεροι πειραματισμοί με μοντέλα & αλγορίθμους μηχανικής μάθησης. Όμως, θα απαλειφθεί και η ανάγκη του genre recognition, οπότε οι μελλοντικές machine learning δοκιμές θα μπορούν να εστιάσουν αλλού. Αναφορικά με το σύστημα συστάσεων που αναπτύχθηκε, μελλοντικά θα μπορούσε να εξεταστεί το ενδεχόμενο ενσωμάτωσης του σε commercial προϊόντα, όπως είναι το σύστημα αναπαραγωγής του ίδιου του Spotify, κάποια άλλη υπηρεσία ροής μουσικής ή μια απλή εφαρμογή λογισμικού music / media player. Η αποδοτικότητα και η λειτουργία του τον καθιστούν χρήσιμο. Θα μπορούσε να υπάρχει εντός του συστήματος αναπαραγωγής, ως μια επιλογή που έχουν στη διάθεση τους οι χρήστες, εκείνη δηλαδή του smart shuffling, το οποίο θα προσπαθούσε να συγκλίνει στις μελαδικές προτιμήσεις του χρήστη, όσο εκείνος άκουγε τυχαία μουσικά κομμάτια από τη λίστα αναπαραγωγής του και θα προσπερνούσε όσα δεν ήθελε εκείνη τη στιγμή. Επιπρόσθετα, θα μπορούσε να εξεταστεί η ενσωμάτωση του ίδιου συστήματος συστάσεων, αλλά δεχόμενος το αντίθετο είδος δεδομένων, δηλαδή μόνο τα θετικά δείγματα, ή έναν συνδυασμό θετικών & αρνητικών δειγμάτων. Σε μια γρήγορη δοκιμή που διεξήχθη πάνω στον recommender της παρούσας διπλωματικής, **η χρήση ενός και μόνο θετικού δείγματος θα ήταν αρκετή ώστε εκείνος να αντιληφθεί πλήρως τη προτίμηση του χρήστη και να αναπροσαρμώσει ολόκληρη τη λίστα αναπαραγωγής με εξαιρετική επιτυχία!** Δεν θα υπήρχε ανάγκη επανάληψης ή δεύτερου θετικού δείγματος. Αρκεί ο χρήστης να ξεκινούσε την ακρόαση με ένα μουσικό κομμάτι που θέλει να ακούσει εκείνη τη στιγμή και **το σύστημα συστάσεων που αναπτύχθηκε σε αυτή τη διπλωματική εργασία θα εντόπιζε απευθείας τα επόμενα μουσικά κομμάτια**, αντιλαμβανόμενος τις προτιμήσεις του.

Τελικά, ναι, είναι εφικτό για τα συστήματα συστάσεων να συνδυαστούν με αλ-

γορίθμους μηχανικής μάθησης, να αξιοποιήσουν αρνητικά δείγματα και να παράξουν αξιόλογα αποτελέσματα προτάσεων! Η παρούσα μεταπτυχιακή διπλωματική εργασία αποτελεί περίτρανη απόδειξη αυτού. Φυσικά, υπάρχει χώρος για βελτίωση! Άλλα όταν υπάρχει χώρος για βελτίωσει, υπάρχει και θέληση για περαιτέρω έρευνα, ανάπτυξη και υλοποίηση...! **Τελικά, recommenders υπάρχουν πολλοί. Recommender μουσικών δεδομένων όμως που να αξιοποιεί μόνο αρνητικά δείγματα και να προβαίνει σε συστάσεις ανάλογα με τη μελωδική προτίμηση του χρήστη, σε πραγματικό χρόνο, ενώ ταυτόχρονα αξιοποιεί και προβλέψεις μοντέλου μηχανικής μάθησης, μόνο ένας!**



## Παρατήματα

---



## Βιβλιογραφία

---

- [1] *Wvau.org's article: "Feedback: Are Musical Genres Meaningless?"*. <http://wvau.org/3046/archives/feedback-are-musical-genres-meaningless/>. Image.
- [2] *MIT News's article: "Try this! Researchers devise better recommendation algorithm"*. <https://news.mit.edu/2017/better-recommendation-algorithm-1206>. Image.
- [3] *Towards Data Science's article: "Build your own Recommender System within 5 minutes!"*. <https://doi.org/10.26782/jmcms.spl.7/2020.02.00006>. Image.
- [4] *BBVA AI Factory's article: "Recommender systems: Marketing gets personal"*. <https://www.bbvaaifactory.com/recommender-systems-marketing-gets-personal/>. Image.
- [5] Folasade Olubusola Isinkaye, YO Folajimi και Bolande Adefowoke Ojokoh.  
*Image from: "Recommendation systems: Principles, methods and evaluation". Egyptian Informatics Journal, 16(3):261-273, 2015.*
- [6] Sindhu V,Nivedha S,Prakash M, "AN EMPIRICAL SCIENCE RESEARCH ON BIOINFORMATICS IN MACHINE LEARNING", 2020, Journal Special Issue. <https://doi.org/10.26782/jmcms.spl.7/2020.02.00006>. Image.
- [7] Sebastian Raschka, "What Are Machine Learning And Deep Learning? An Overview.", Lecture 01, Course STAT 453, UW-Madison. [https://github.com/rasbt/stat453-deep-learning-ss20/blob/master/L01-intro/L01-intro\\_slides.pdf](https://github.com/rasbt/stat453-deep-learning-ss20/blob/master/L01-intro/L01-intro_slides.pdf). Image.
- [8] ThermoFisher Scientific's article, "Machine Learning: A Primer to Laboratory Applications". <https://www.thermofisher.com/blog/connectedlab/machine-learning-a-primer-to-laboratory-applications/>. Image.
- [9] Oracle Blog's article, "Types of Machine Learning and Top 10 Algorithms Everyone Should Know". <https://blogs.oracle.com/datascience/types-of-machine-learning-and-top-10-algorithms-everyone-should-know-v2>. Image.

- [10] *9To5Mac's article "What's new in the Apple Music app for iOS 14: Listen Now tab, endless autoplay, iPad redesign"*. <https://9to5mac.com/2020/06/24/ios-14-apple-music/>. Image.
- [11] Dongsheng Li, Qin Lv, Xing Xie, Li Shang, Huanhuan Xia, Tun Lu και Ning Gu. *Interest-based real-time content recommendation in online social communities*. *Knowledge-based systems*, 28:1–12, 2012.
- [12] Medium's article "Item-based Collaborative Filtering for Music Recommender System". <https://medium.com/unstructured-data-service/item-based-collaborative-filtering-for-music-recommender-system-ff3b60d1bc80>. Image.
- [13] DataFlair's article "Python Project - Music Genre Classification". <https://data-flair.training/blogs/python-project-music-genre-classification/>. Image.
- [14] Michaël Defferrard, Kirell Benzi, Pierre Vandergheynst και Xavier Bresson. *Fma: A dataset for music analysis*. *arXiv preprint arXiv:1612.01840*, 2016.
- [15] Towards Data Science's article "Introduction to Logistic Regression". <https://towardsdatascience.com/introduction-to-logistic-regression-66248243c148>. Image.
- [16] Wikipedia's article about the k-Nearest Neighbors algorithm. [https://en.wikipedia.org/wiki/K-nearest\\_neighbors\\_algorithm](https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm). Image.
- [17] Datacamp's article "Decision Tree Classification in Python". <https://www.datacamp.com/community/tutorials/decision-tree-classification-python>. Image.
- [18] Quantinsti.com's article "Introduction to XGBoost in Python". <https://blog.quantinsti.com/xgboost-python/>. Image.
- [19] James Blades. *Percussion instruments and their history*. Bold Strummer Limited, 1992.
- [20] Michael Bull. *Investigating the culture of mobile listening: From Walkman to iPod. Consuming music together*, σελίδες 131–149. Springer, 2006.
- [21] Dan D Carman. *Device and method for CD shuffle play*. Δίπλωμα Ευρεσιτεχνίας, 1995. US Patent 5,408,448.
- [22] Katie Rose M Sanfilippo, Neta Spiro, Miguel Molina-Solana και Alexandra Lamont. *Do the shuffle: Exploring reasons for music listening through shuffled play*. *PLoS one*, 15(2):e0228457, 2020.

- [23] Francesco Ricci, Lior Rokach και Bracha Shapira. *Recommender systems: introduction and challenges*. *Recommender systems handbook*, σελίδες 1–34. Springer, 2015.
- [24] Uri Hanani, Bracha Shapira και Peretz Shoval. *Information filtering: Overview of issues, research and systems. User modeling and user-adapted interaction*, 11(3):203–259, 2001.
- [25] Jonathan L Herlocker, Joseph A Konstan, Loren G Terveen και John T Riedl. *Evaluating collaborative filtering recommender systems*. *ACM Transactions on Information Systems (TOIS)*, 22(1):5–53, 2004.
- [26] Dietmar Jannach, Markus Zanker, Alexander Felfernig και Gerhard Friedrich. *Recommender systems: an introduction*. Cambridge University Press, 2010.
- [27] John S Breese, David Heckerman και Carl Kadie. *Empirical analysis of predictive algorithms for collaborative filtering*. *arXiv preprint arXiv:1301.7363*, 2013.
- [28] Xenonstack.com's article: "Next Generation Recommender Systems Overview". <https://www.xenonstack.com/blog/recommender-systems/>. Article about recommendation systems and their generations through the recent years.
- [29] Tom M Mitchell και others. *Machine learning*. 1997. Burr Ridge, IL: McGraw Hill, 45(37):870–877, 1997.
- [30] Arthur L Samuel. *Some studies in machine learning using the game of checkers*. *IBM Journal of research and development*, 3(3):210–229, 1959.
- [31] Ethem Alpaydin. *Introduction to machine learning*. MIT press, 2020.
- [32] Jürgen Schmidhuber. *Deep learning in neural networks: An overview*. *Neural networks*, 61:85–117, 2015.
- [33] Taiwo Oladipupo Ayodele. *Types of machine learning algorithms*. *New advances in machine learning*, 3:19–48, 2010.
- [34] Stevo Bozinovski και others. *A self-learning system using secondary reinforcement*. *Cybernetics and Systems Research*, σελίδες 397–402, 1982.
- [35] Yoshua Bengio, Aaron Courville και Pascal Vincent. *Representation learning: A review and new perspectives*. *IEEE transactions on pattern analysis and machine intelligence*, 35(8):1798–1828, 2013.

- [36] Xianke Zhou, Sai Wu, Chun Chen, Gang Chen και Shanshan Ying. *Real-time recommendation for microblogs*. *Information Sciences*, 279:301–325, 2014.
- [37] Sofiane Abbar, Sihem Amer-Yahia, Piotr Indyk και Sepideh Mahabadi. *Real-time recommendation of diverse related articles*. *Proceedings of the 22nd international conference on World Wide Web*, σελίδες 1–12, 2013.
- [38] Jin cheng Zhang και Yasufumi Takama. *Proposal of Context-Aware Music Recommender System Using Negative Sampling*. *Annual Conference of the Japanese Society for Artificial Intelligence*, σελίδες 114–125. Springer, 2019.
- [39] Geoffray Bonnin και Dietmar Jannach. *A comparison of playlist generation strategies for music recommendation and a new baseline scheme*. *Workshops at the twenty-seventh AAAI conference on artificial intelligence*, 2013.
- [40] Markus Schedl, Peter Knees και Fabien Gouyon. *New paths in music recommender systems research*. *Proceedings of the Eleventh ACM Conference on Recommender Systems*, σελίδες 392–393, 2017.
- [41] Bruce Ferwerda, Emily Yang, Markus Schedl και Marko Tkalcic. *Personality traits predict music taxonomy preferences*. *Proceedings of the 33rd annual acm conference extended abstracts on human factors in computing systems*, σελίδες 2241–2246, 2015.
- [42] Marius Kaminskas και Francesco Ricci. *Emotion-based matching of music to places*. *Emotions and Personality in Personalized Services*, σελίδες 287–310. Springer, 2016.
- [43] Dmitry Bogdanov και Perfecto Herrera. *How Much Metadata Do We Need in Music Recommendation? A Subjective Evaluation Using Preference Sets*. *ISMIR*, σελίδες 97–102, 2011.
- [44] Leyi Wei, Pengwei Xing, Jianchang Zeng, JinXiu Chen, Ran Su και Fei Guo. *Improved prediction of protein-protein interactions using novel negative samples, features, and an ensemble classifier*. *Artificial Intelligence in Medicine*, 83:67–74, 2017.
- [45] Peifeng Wang, Shuangyin Li και Rong Pan. *Incorporating gan for negative sampling in knowledge representation learning*. *Proceedings of the AAAI Conference on Artificial Intelligence*, τόμος 32, 2018.
- [46] Junjie Chen, Xiaolong Wang και Bin Liu. *IMiRNA-SSF: improving the identification of MicroRNA precursors by combining negative sets with different distributions*. *Scientific reports*, 6(1):1–10, 2016.

- [47] Tao Li, Mitsunori Ogihara και Qi Li. *A comparative study on content-based music genre classification*. Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval, σελίδες 282–289, 2003.
- [48] Changsheng Xu, Namunu C Maddage, Xi Shao, Fang Cao και Qi Tian. *Musical genre classification using support vector machines*. 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03)., τόμος 5, σελίδες “-429. IEEE, 2003.
- [49] Hareesh Bahuleyan. *Music genre classification using machine learning techniques*. arXiv preprint arXiv:1804.01149, 2018.
- [50] Roberto Basili, Alfredo Serafini και Armando Stellato. *Classification of musical genre: a machine learning approach*. ISMIR, 2004.
- [51] Prasenjeet Fulzele, Rajat Singh, Naman Kaushik και Kavita Pandey. *A hybrid model for music genre classification using LSTM and SVM*. 2018 Eleventh International Conference on Contemporary Computing (IC3), σελίδες 1–3. IEEE, 2018.
- [52] Sergio Oramas, Oriol Nieto, Francesco Barbieri και Xavier Serra. *Multi-label music genre classification from audio, text, and images using deep features*. arXiv preprint arXiv:1707.04916, 2017.
- [53] Brendan OConnor Dawen Liang, Haijie Gu. *Music Genre Classification with the Million Song Dataset*. Free publication available on the web, 2011.
- [54] Anneloes Bal. *Music genre classification using audio features and lyrics sentiment features with multinomial logistic regression*. Μεταπτυχιακή διπλωματική εργασία, Tilburg University, 2018.
- [55] Kehan Luo. *Machine Learning Approach for Genre Prediction on Spotify Top Ranking Songs*. 2018.
- [56] *Million Song Dataset*. <http://millionsongdataset.com/>. Official Website.
- [57] *Free Music Archive Dataset*. <https://github.com/mdeff/fma>. Official Website.
- [58] *Python*. <https://www.python.org/>. Official Website.
- [59] *Python spotify Library*. <https://spotipy.readthedocs.io/en/2.17.1/>. Official Website.
- [60] *Scikit-Learn*. <https://scikit-learn.org/stable/>. Official Website.

- [61] *Anastasios (Tasos) Nikolakopoulos*, [tasosnikolakop@mail.ntua.gr](mailto:tasosnikolakop@mail.ntua.gr). <https://www.linkedin.com/in/anastasios-nikolakopoulos/>. Contact Info.
- [62] *Datacamp's article "Convolutional Neural Networks in Python with Keras"*. <https://www.datacamp.com/community/tutorials/convolutional-neural-networks-python>. Web Article.