



Σχολή Εφαρμοσμένων Τεχνών & Βιώσιμου Σχεδιασμού

Ακουστικός Σχεδιασμός και Ψηφιακός Ήχος

Διπλωματική Εργασία

Πρόβλεψη και Σημασιολογική Αναπαράσταση Μουσικού
Συναισθήματος από Αντικειμενικά Ηχητικά Χαρακτηριστικά

Λάζαρος Γιαννακός

Επιβλέπων καθηγητής: Ρήγας Κωτσάκης

Πάτρα, Μάιος 2026

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του φοιτητή («συγγραφέας/δημιουργός») που την εκπόνησε. Στο πλαίσιο της πολιτικής ανοικτής πρόσβασης ο συγγραφέας/δημιουργός εκχωρεί στο ΕΑΠ, μη αποκλειστική άδεια χρήσης του δικαιώματος αναπαραγωγής, προσαρμογής, δημόσιου δανεισμού, παρουσίασης στο κοινό και ψηφιακής διάχυσής τους διεθνώς, σε ηλεκτρονική μορφή και σε οποιοδήποτε μέσο, για διδακτικούς και ερευνητικούς σκοπούς, άνευ ανταλλάγματος και για όλο το χρόνο διάρκειας των δικαιωμάτων πνευματικής ιδιοκτησίας. Η ανοικτή πρόσβαση στο πλήρες κείμενο για μελέτη και ανάγνωση δεν σημαίνει καθ' οιονδήποτε τρόπο παραχώρηση δικαιωμάτων διανοητικής ιδιοκτησίας του συγγραφέα/δημιουργού ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αντιγραφή, αποθήκευση, πώληση, εμπορική χρήση, μετάδοση, διανομή, έκδοση, εκτέλεση, «μεταφόρτωση» (downloading), «ανάρτηση» (uploading), μετάφραση, τροποποίηση με οποιονδήποτε τρόπο, τμηματικά ή περιληπτικά της εργασίας, χωρίς τη ρητή προηγούμενη έγγραφη συναίνεση του συγγραφέα/δημιουργού. Ο συγγραφέας/δημιουργός διατηρεί το σύνολο των ηθικών και περιουσιακών του δικαιωμάτων.

Πρόβλεψη και Σημασιολογική Αναπαράσταση Μουσικού
Συναισθήματος από Αντικειμενικά Ηχητικά Χαρακτηριστικά

Λάζαρος Γιαννακός

Επιτροπή Επίβλεψης Διπλωματικής Εργασίας

Επιβλέπων Καθηγητής

Ρήγας Κωτσάκης

Αναπληρωτής Καθηγητής ΔΙΠΑΕ

Τμήμα Μηχανικών Πληροφορικής και
Ηλεκτρονικών Συστημάτων

Συν-Επιβλέπων Καθηγητής

Χαράλαμπος Δημούλας

Καθηγητής ΑΠΘ

Τμήμα Δημοσιογραφίας και Μέσων
Μαζικής Επικοινωνίας

Πάτρα, Μάιος 2026

Ευχαριστίες

Θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες στον επιβλέποντα καθηγητή μου, κ. Ρήγα Κωτσάκη, για την υποστήριξη και τις πολύτιμες συμβουλές του κατά τη διάρκεια εκπόνησης της Διπλωματικής Εργασίας.

Επίσης, θα ήθελα να ευχαριστήσω τους καθηγητές του μεταπτυχιακού προγράμματος για τις γνώσεις, τα ερεθίσματα και τη συνολική συμβολή τους στην ακαδημαϊκή μου πορεία.

Τέλος, ευχαριστώ θερμά την οικογένειά μου για την στήριξη και την ενθάρρυνσή της σε όλη τη διάρκεια των σπουδών μου.

*Σε σύγκριση με τη μουσική,
όλα τα φαινόμενα είναι απλώς σύμβολα.*

Friedrich Nietzsche

Περίληψη

Η παρούσα Διπλωματική Εργασία διερευνά τη δυνατότητα πρόβλεψης συναισθηματικών αξιολογήσεων μουσικών αποσπασμάτων με βάση αντικειμενικά ηχητικά χαρακτηριστικά. Η εργασία εντάσσεται στο πεδίο της Ανάκτησης Μουσικής Πληροφορίας και ειδικότερα της Αναγνώρισης Μουσικού Συναισθήματος, εξετάζοντας τόσο διαστασιακές συναισθηματικές αναπαραστάσεις όσο και βασικά συναισθήματα.

Για τον σκοπό αυτό, χρησιμοποιήθηκε σύνολο δεδομένων μουσικών αποσπασμάτων με διαθέσιμες ανθρώπινες αξιολογήσεις ως προς τις διαστάσεις Valence, Energy και Tension, καθώς και ως προς τα βασικά συναισθήματα Happy, Sad, Anger, Fear και Tender. Από τα μουσικά αποσπάσματα εξήχθησαν αντικειμενικά ηχητικά χαρακτηριστικά που αφορούν την ένταση και τη δυναμική, τον ρυθμό, την τονικότητα, το φασματικό περιεχόμενο και το ηχόχρωμα. Τα χαρακτηριστικά οργανώθηκαν σε εναλλακτικά σύνολα εισόδου και αξιοποιήθηκαν για την εκπαίδευση μοντέλων παλινδρόμησης.

Στο πειραματικό μέρος αξιολογήθηκαν διαφορετικά μοντέλα μηχανικής μάθησης και σύνολα χαρακτηριστικών. Ο αλγόριθμος GradientBoostingRegressor παρουσίασε την καλύτερη συνολική απόδοση και αποτέλεσε το βασικό μοντέλο των τελικών πειραμάτων. Επιπλέον, εφαρμόστηκαν διαδικασίες επιλογής χαρακτηριστικών και ερμηνευσιμότητας με χρήση της μεθόδου SHAP και ανάλυσης σταθερότητας, οι οποίες οδήγησαν σε μειωμένα σύνολα χαρακτηριστικών με βελτιωμένη απόδοση και μεγαλύτερη ερμηνευσιμότητα.

Τα αποτελέσματα ανέδειξαν τη σημασία κυρίως τονικών/αρμονικών και ηχοχρωματικών/φασματικών χαρακτηριστικών για την πρόβλεψη των συναισθηματικών αξιολογήσεων. Παράλληλα, εξετάστηκε η χαρτογράφηση μεταξύ συναισθηματικών διαστάσεων και βασικών συναισθημάτων. Η άμεση πρόβλεψη βασικών συναισθημάτων από τα ηχητικά χαρακτηριστικά παρουσίασε καλύτερη συνολική απόδοση, ενώ η έμμεση προσέγγιση μέσω των προβλεπόμενων διαστάσεων προσέφερε ένα ενδιάμεσο ερμηνευτικό πλαίσιο.

Συνολικά, η εργασία δείχνει ότι τα αντικειμενικά ηχητικά χαρακτηριστικά μπορούν να αξιοποιηθούν για την πρόβλεψη και ερμηνεία μουσικού συναισθήματος. Επιπλέον, η αναπαράσταση των αποτελεσμάτων ως συναισθηματικών προφίλ συμβάλλει σε μια πιο κατανοητή σημασιολογική παρουσίαση του συναισθηματικού χαρακτήρα των μουσικών αποσπασμάτων.

Λέξεις – Κλειδιά

Ανάκτηση Μουσικής Πληροφορίας, Αναγνώριση Μουσικού Συναισθήματος, αντικειμενικά ηχητικά χαρακτηριστικά, μηχανική μάθηση, επιλογή χαρακτηριστικών, SHAP

Prediction and Semantic Representation of Musical Emotion from Objective Audio Features

Lazaros Giannakos

Abstract

This Diploma Thesis investigates the possibility of predicting emotional ratings of musical excerpts based on objective audio features. The study falls within the field of Music Information Retrieval (MIR) and, more specifically, Music Emotion Recognition (MER), examining both dimensional emotional representations and basic emotions.

For this purpose, a dataset of musical excerpts with available human ratings was used, including ratings for the dimensions Valence, Energy and Tension, as well as for the basic emotions Happy, Sad, Anger, Fear and Tender. Objective audio features related to loudness and dynamics, rhythm, tonality, spectral content and timbre were extracted from the musical excerpts. These features were organised into alternative input sets and used to train regression models.

In the experimental part, different machine learning models and feature sets were evaluated. GradientBoostingRegressor algorithm achieved the best overall performance and was selected as the main model for the final experiments. In addition, feature selection and interpretability procedures were applied using SHAP and stability analysis, leading to reduced feature sets with improved performance and greater interpretability.

The results highlighted the importance of mainly tonal/harmonic and timbral/spectral features for predicting emotional ratings. Moreover, the mapping between emotional dimensions and basic emotions was examined. Direct prediction of basic emotions from audio features achieved better overall performance, while the indirect approach through predicted dimensions provided an intermediate interpretative framework.

Overall, the study shows that objective audio features can be used for the prediction and interpretation of musical emotion. Furthermore, representing the results as emotional profiles contributes to a more understandable semantic presentation of the emotional character of musical excerpts.

Keywords

Music Information Retrieval, Music Emotion Recognition, objective audio features, machine learning, feature selection, SHAP

Περιεχόμενα

Περίληψη.....	vi
Abstract	viii
Περιεχόμενα	x
Κατάλογος Εικόνων	xii
Κατάλογος Πινάκων	xiii
Συντομογραφίες & Ακρωνύμια.....	xiv
1. Εισαγωγή.....	1
1.1 Αντικείμενο και Σκοπός της Διπλωματικής Εργασίας.....	1
1.2 Ερευνητικά Ερωτήματα και Στόχοι	2
1.3 Συνοπτική Μεθοδολογική Προσέγγιση	3
1.4 Δομή της Διπλωματικής Εργασίας.....	4
2. Θεωρητικό Υπόβαθρο.....	6
2.1 Ανάκτηση Μουσικής Πληροφορίας.....	6
2.2 Συναισθηματική Αναπαράσταση στη Μουσική	7
2.3 Διαστασιακά και Κατηγορικά Μοντέλα Συναισθήματος	8
2.4 Αντικειμενικά Ηχητικά Χαρακτηριστικά	10
2.5 Μηχανική Μάθηση για Πρόβλεψη Συναισθήματος στη Μουσική.....	11
3. Βιβλιογραφική Επισκόπηση.....	14
3.1 Πρόβλεψη Μουσικού Συναισθήματος από Ηχητικά Χαρακτηριστικά.....	14
3.2 Διαστασιακές και Κατηγορικές Προσεγγίσεις.....	16
3.3 Επιλογή Χαρακτηριστικών και Ερμηνευσιμότητα	17
3.4 Χαρτογράφηση μεταξύ Συναισθηματικών Χώρων.....	19
3.5 Σημασιολογική Αναπαράσταση Μουσικού Συναισθήματος	20
3.6 Κενά της Βιβλιογραφίας και Σημασία της Διπλωματικής Εργασίας.....	21
4. Εξαγωγή και Ανάλυση Ηχητικών Παραμέτρων	23
4.1 Περιγραφή Συνόλου Δεδομένων.....	23
4.2 Συναισθηματικές Αξιολογήσεις και Στόχοι Πρόβλεψης	24
4.3 Διαδικασία Εξαγωγής και Επιλογής Αντικειμενικών Χαρακτηριστικών	25
4.4 Προεπεξεργασία και Τελική Διαμόρφωση των Δεδομένων	28
5. Υλοποίηση Μοντέλων και Πειραματική Διαδικασία	30
5.1 Συνοπτική Περιγραφή της Πειραματικής Διαδικασίας.....	30
5.2 Επιλογή Μοντέλων Πρόβλεψης και Παράμετροι.....	31
5.3 Μέτρα Αξιολόγησης	33
5.4 SHAP και Ανάλυση Σταθερότητας.....	34
5.5 Κοινή και ανά Στόχο Επιλογή Αντικειμενικών Χαρακτηριστικών	35
5.6 Τελικά Μοντέλα Πρόβλεψης για Διαστάσεις και Βασικά Συναισθήματα	36
5.7 Στατιστικός Έλεγχος Βελτίωσης των Τελικών Μοντέλων	37
5.8 Χαρτογράφηση Συναισθηματικών Διαστάσεων σε Βασικά Συναισθήματα.....	38
5.9 Άμεση Πρόβλεψη Συναισθηματικού Προφίλ και Σύγκριση με Χαρτογράφηση.....	40
6. Πειραματικά Αποτελέσματα	41
6.1 Αρχική Σύγκριση Μοντέλων και Συνόλων Χαρακτηριστικών.....	41
6.2 Τελικά Μοντέλα για Συναισθηματικές Διαστάσεις	43
6.3 Τελικά Μοντέλα Πρόβλεψης Βασικών Συναισθημάτων	44
6.4 Αποτελέσματα Επιλογής Χαρακτηριστικών και Ερμηνευσιμότητας	46
6.5 Στατιστική Σύγκριση Τελικών Μοντέλων	52
6.6 Χαρτογράφηση μεταξύ Συναισθηματικών Χώρων.....	54

6.7	Σύγκριση Άμεσης και Έμμεσης Πρόβλεψης Βασικών Συναισθημάτων	57
6.8	Σύνοψη και Ερμηνεία Αποτελεσμάτων	59
7.	Σημασιολογική Ερμηνεία και Αναπαράσταση Συναισθηματικού Προφίλ Μουσικών Αποσπασμάτων	61
7.1	Στόχος της Σημασιολογικής Ερμηνείας	61
7.2	Αναπαράσταση Συναισθηματικού Προφίλ	61
7.3	Σημασιολογική Ερμηνεία μέσω Χαρτογράφησης Διαστάσεων	63
7.4	Σημασιολογικές Περιγραφές Μουσικών Αποσπασμάτων	64
7.5	Σύνοψη και Περιορισμοί της Σημασιολογικής Αναπαράστασης	65
8.	Συζήτηση, Συμπεράσματα και Μελλοντική Έρευνα	67
8.1	Συζήτηση	67
8.2	Συμπεράσματα	69
8.3	Μελλοντική Έρευνα	70
	Βιβλιογραφικές Αναφορές	72
	Παράρτημα Α: Πίνακας Μουσικών Αποσπασμάτων και Συναισθηματικών Αξιολογήσεων	77

Κατάλογος Εικόνων

Εικόνα 5.1: Διάγραμμα ροής της πειραματικής διαδικασίας.....	31
Εικόνα 6.1: SHAP bar plot για την πρόβλεψη του συναισθήματος Happy	48
Εικόνα 6.2: SHAP beeswarm plot για την πρόβλεψη του συναισθήματος Happy	48
Εικόνα 6.3: SHAP bar plot για την πρόβλεψη της διάστασης Valence.....	50
Εικόνα 6.4: SHAP beeswarm plot για την πρόβλεψη της διάστασης Valence.....	50
Εικόνα 6.5: Κέντρα βάρους των βασικών συναισθημάτων στον διαστασιακό χώρο: α) τρειςδιάστατη αναπαράσταση Valence-Energy-Tension, β) προβολή στο επίπεδο Valence- Tension και γ) προβολή στο επίπεδο Valence-Energy.	55
Εικόνα 7.1: Ενδεικτικά συναισθηματικά προφίλ μουσικών αποσπασμάτων.	62
Εικόνα 7.2: Ενδεικτική σημασιολογική αναπαράσταση μουσικού αποσπάσματος μέσω της θέσης του στον διαστασιακό χώρο και της σύγκρισης πραγματικού, άμεσα προβλεπόμενου και έμμεσα χαρτογραφημένου συναισθηματικού προφίλ.....	64

Κατάλογος Πινάκων

Πίνακας 4.1: Αντικειμενικά χαρακτηριστικά, τρόπος υπολογισμού και περιγραφή.....	27
Πίνακας 6.1: Αρχική σύγκριση RandomForestRegressor και GradientBoostingRegressor στο σύνολο των 48 χαρακτηριστικών	42
Πίνακας 6.2: Αρχική σύγκριση συνόλων χαρακτηριστικών με GradientBoostingRegressor	42
Πίνακας 6.3: Απόδοση τελικών μοντέλων για την πρόβλεψη συναισθηματικών διαστάσεων.....	43
Πίνακας 6.4: Απόδοση τελικών μοντέλων για την πρόβλεψη βασικών συναισθημάτων...	45
Πίνακας 6.5: Κοινά μειωμένα σύνολα χαρακτηριστικών για τα τελικά μοντέλα.....	47
Πίνακας 6.6: Ανά στόχο μειωμένα σύνολα χαρακτηριστικών για τα τελικά μοντέλα	49
Πίνακας 6.7: Στατιστική σύγκριση τελικών μοντέλων συναισθηματικών διαστάσεων με Wilcoxon signed-rank test.....	53
Πίνακας 6.8: Στατιστική σύγκριση τελικών μοντέλων βασικών συναισθημάτων με Wilcoxon signed-rank test.....	53
Πίνακας 6.9: Χαρτογράφηση από πραγματικές συναισθηματικές διαστάσεις σε πραγματικά βασικά συναισθήματα	54
Πίνακας 6.10: Απόδοση χαρτογράφησης από προβλεπόμενες διαστάσεις σε βασικά συναισθήματα.....	56
Πίνακας 6.11: Απόδοση χαρτογράφησης ως προς το κυρίαρχο συναίσθημα.....	56
Πίνακας 6.12: Πίνακας σύγχυσης για τη χαρτογράφηση με Euclidean απόσταση	57
Πίνακας 6.13: Σύγκριση άμεσης και έμμεσης πρόβλεψης βασικών συναισθημάτων	58
Πίνακας 7.1: Ενδεικτικές σημασιολογικές περιγραφές μουσικών αποσπασμάτων.....	65
Πίνακας A.1: Μουσικά αποσπάσματα και συναισθηματικές αξιολογήσεις.....	77

Συντομογραφίες & Ακρωνύμια

ΔΕ	Διπλωματική Εργασία
BPM	Beats Per Minute / Χτύποι ανά λεπτό
FFT	Fast Fourier Transform / Ταχύς Μετασχηματισμός Fourier
HPCP	Harmonic Pitch Class Profile
MAE	Mean Absolute Error / Μέσο Απόλυτο Σφάλμα
MER	Music Emotion Recognition / Αναγνώριση Μουσικού Συναισθήματος
MFCC	Mel-Frequency Cepstral Coefficients
MIR	Music Information Retrieval / Ανάκτηση Μουσικής Πληροφορίας
MIREX	Music Information Retrieval Evaluation eXchange
MLR	Multiple Linear Regression / Πολλαπλή Γραμμική Παλινδρόμηση
PLS	Partial Least Squares Regression / Μερική Παλινδρόμηση Ελαχίστων Τετραγώνων
R^2	Coefficient of Determination / Συντελεστής προσδιορισμού
RMS	Root Mean Square / Ρίζα Μέσης Τετραγωνικής Τιμής
SHAP	Shapley Additive exPlanations
STD	Standard Deviation / Τυπική απόκλιση
ZCR	Zero-Crossing Rate / Ρυθμός Διέλευσης από το Μηδέν

1. Εισαγωγή

Η μουσική συνδέεται στενά με την έκφραση, την αντίληψη και την πρόκληση συναισθηματικών καταστάσεων, αποτελώντας σημαντικό αντικείμενο μελέτης στη μουσική ψυχολογία και στην Ανάκτηση Μουσικής Πληροφορίας (Music Information Retrieval, MIR). Η δυνατότητα αυτόματης ανάλυσης του συναισθηματικού περιεχομένου της μουσικής έχει οδηγήσει στην ανάπτυξη του πεδίου της Αναγνώρισης Μουσικού Συναισθήματος (Music Emotion Recognition, MER), το οποίο επιχειρεί να συνδέσει μουσικά δεδομένα και ηχητικά χαρακτηριστικά με ανθρώπινες συναισθηματικές αξιολογήσεις (Kim et al., 2010; Yang & Chen, 2012).

Στο πλαίσιο αυτό, τα αντικειμενικά ηχητικά χαρακτηριστικά αποτελούν σημαντικό μέσο αναπαράστασης του μουσικού σήματος, καθώς επιτρέπουν τη μελέτη πτυχών όπως η ένταση, ο ρυθμός, η τονικότητα και το ηχόχρωμα. Παράλληλα, η πρόβλεψη μουσικού συναισθήματος παραμένει σύνθετο πρόβλημα, καθώς οι συναισθηματικές αξιολογήσεις επηρεάζονται από πολλούς μουσικούς και υποκειμενικούς παράγοντες, ενώ η σχέση μεταξύ ακουστικών ιδιοτήτων και συναισθηματικού περιεχομένου δεν είναι πάντα άμεση ή γραμμική (Panda et al., 2023; Sloboda & Juslin, 2010).

1.1 Αντικείμενο και Σκοπός της Διπλωματικής Εργασίας

Αντικείμενο της παρούσας Διπλωματικής Εργασίας (ΔΕ) είναι η πρόβλεψη συναισθηματικών αξιολογήσεων μουσικών αποσπασμάτων με βάση αντικειμενικά ηχητικά χαρακτηριστικά. Πιο συγκεκριμένα, εξετάζεται κατά πόσο χαρακτηριστικά που εξάγονται από το ηχητικό σήμα, όπως χαρακτηριστικά έντασης, ρυθμού, τονικότητας και ηχοχρώματος, μπορούν να αξιοποιηθούν για την πρόβλεψη τόσο συναισθηματικών διαστάσεων όσο και βασικών συναισθημάτων.

Σκοπός της ΔΕ είναι η ανάπτυξη και αξιολόγηση μιας μεθοδολογικής προσέγγισης που συνδυάζει την εξαγωγή αντικειμενικών ηχητικών χαρακτηριστικών με μοντέλα μηχανικής μάθησης για την πρόβλεψη μουσικού συναισθήματος. Παράλληλα, η ΔΕ στοχεύει στη διερεύνηση της σημασίας των επιμέρους χαρακτηριστικών, στην επιλογή αποδοτικότερων υποσυνόλων χαρακτηριστικών και στην ερμηνεία της σχέσης μεταξύ μουσικών ιδιοτήτων και συναισθηματικών αξιολογήσεων.

Επιπλέον, η ΔΕ εξετάζει τη σχέση μεταξύ δύο διαφορετικών μορφών συναισθηματικής αναπαράστασης: των διαστασιακών αξιολογήσεων και των βασικών συναισθημάτων. Μέσω της χαρτογράφησης μεταξύ των δύο συναισθηματικών χώρων, διερευνάται κατά πόσο οι προβλεπόμενες συναισθηματικές διαστάσεις μπορούν να αξιοποιηθούν για την εκτίμηση ενός πλήρους συναισθηματικού προφίλ βασικών συναισθημάτων. Με αυτόν τον τρόπο, η ΔΕ επιδιώκει μια πιο εύκολα αντιληπτή και κατανοητή σημασιολογική αναπαράσταση του μουσικού συναισθήματος.

1.2 Ερευνητικά Ερωτήματα και Στόχοι

Η ΔΕ οργανώνεται γύρω από το βασικό ερευνητικό ερώτημα του κατά πόσο αντικειμενικά ηχητικά χαρακτηριστικά μουσικών αποσπασμάτων μπορούν να αξιοποιηθούν για την πρόβλεψη συναισθηματικών αξιολογήσεων.

Το βασικό αυτό ερώτημα αναλύεται στα ακόλουθα επιμέρους ερευνητικά ερωτήματα:

- Ποια σύνολα αντικειμενικών ηχητικών χαρακτηριστικών οδηγούν σε καλύτερη πρόβλεψη συναισθηματικών αξιολογήσεων;
- Ποια μοντέλα μηχανικής μάθησης είναι καταλληλότερα για την πρόβλεψη συναισθηματικών διαστάσεων και βασικών συναισθημάτων;
- Ποια ηχητικά χαρακτηριστικά συμβάλλουν περισσότερο στις προβλέψεις των μοντέλων;
- Μπορεί η επιλογή μειωμένων συνόλων χαρακτηριστικών να βελτιώσει την απόδοση και την ερμηνευσιμότητα των μοντέλων;
- Σε ποιο βαθμό μπορούν οι προβλεπόμενες συναισθηματικές διαστάσεις να χαρτογραφηθούν σε προφίλ βασικών συναισθημάτων;
- Πώς μπορεί το προβλεπόμενο συναισθηματικό προφίλ να παρουσιαστεί με πιο κατανοητό και σημασιολογικά ερμηνεύσιμο τρόπο;

Αντίστοιχα, οι βασικοί στόχοι της ΔΕ είναι:

- Η εξαγωγή και οργάνωση αντικειμενικών ηχητικών χαρακτηριστικών από μουσικά αποσπάσματα.
- Η εκπαίδευση και αξιολόγηση μοντέλων μηχανικής μάθησης για την πρόβλεψη συναισθηματικών διαστάσεων και βασικών συναισθημάτων.
- Η σύγκριση διαφορετικών συνόλων χαρακτηριστικών και μοντέλων πρόβλεψης.

- Η εφαρμογή τεχνικών επιλογής χαρακτηριστικών και ερμηνευσιμότητας, με έμφαση στη μέθοδο SHAP (Shapley Additive exPlanations).
- Η διερεύνηση της σχέσης μεταξύ διαστασιακών και κατηγορικών συναισθηματικών αναπαραστάσεων μέσω χαρτογράφησης.
- Η διαμόρφωση συναισθηματικών προφίλ και ενδεικτικών σημασιολογικών περιγραφών μουσικών αποσπασμάτων.

1.3 Συνοπτική Μεθοδολογική Προσέγγιση

Η μεθοδολογική προσέγγιση της ΔΕ συνοψίζεται στα εξής βήματα. Αρχικά, συγκεντρώθηκε το σύνολο δεδομένων που χρησιμοποιείται στην εργασία, το οποίο περιλαμβάνει μουσικά αποσπάσματα συνοδευόμενα από αξιολογήσεις ως προς συναισθηματικές διαστάσεις και βασικά συναισθήματα.

Στη συνέχεια, πραγματοποιήθηκε εξαγωγή αντικειμενικών ηχητικών χαρακτηριστικών από τα μουσικά αποσπάσματα με χρήση εργαλείων της βιβλιοθήκης Essentia και συμπληρωματικής επεξεργασίας. Τα χαρακτηριστικά αυτά οργανώθηκαν σε τρία διαφορετικά σύνολα εισόδου και συνδυάστηκαν με τις διαθέσιμες συναισθηματικές αξιολογήσεις, ώστε να διαμορφωθούν τα δεδομένα που χρησιμοποιήθηκαν στα πειράματα πρόβλεψης.

Στο πειραματικό στάδιο, εκπαιδεύτηκαν και αξιολογήθηκαν μοντέλα μηχανικής μάθησης για την πρόβλεψη τόσο συναισθηματικών διαστάσεων όσο και βασικών συναισθημάτων. Η αξιολόγηση πραγματοποιήθηκε με τεχνικές διασταυρούμενης επικύρωσης και κατάλληλες μετρικές παλινδρόμησης. Παράλληλα, εφαρμόστηκαν διαδικασίες επιλογής χαρακτηριστικών και ερμηνευσιμότητας με χρήση της μεθόδου SHAP και ανάλυσης σταθερότητας, με στόχο τη διαμόρφωση αποδοτικότερων και πιο ερμηνεύσιμων μειωμένων συνόλων χαρακτηριστικών.

Τέλος, διερευνήθηκε η σχέση μεταξύ διαστασιακών και κατηγορικών συναισθηματικών αναπαραστάσεων μέσω διαδικασίας χαρτογράφησης. Οι προβλέψεις των συναισθηματικών διαστάσεων χρησιμοποιήθηκαν για την εκτίμηση προφίλ βασικών συναισθημάτων, ενώ τα αποτελέσματα συγκρίθηκαν με την άμεση πρόβλεψη των βασικών συναισθημάτων από τα ηχητικά χαρακτηριστικά. Επιπλέον, τα τελικά συναισθηματικά προφίλ αξιοποιήθηκαν για τη σημασιολογική αναπαράσταση και ερμηνεία του συναισθηματικού χαρακτήρα των μουσικών αποσπασμάτων.

1.4 Δομή της Διπλωματικής Εργασίας

Η ΔΕ οργανώνεται σε οκτώ Κεφάλαια. Στο Κεφάλαιο 1 παρουσιάζονται το αντικείμενο, ο σκοπός, τα ερευνητικά ερωτήματα και οι στόχοι, καθώς και η συνοπτική μεθοδολογική προσέγγιση της ΔΕ.

Στο Κεφάλαιο 2 παρουσιάζεται το θεωρητικό υπόβαθρο της ΔΕ. Εισάγεται το πεδίο της MIR, εξετάζεται η συναισθηματική αναπαράσταση στη μουσική και παρουσιάζονται τα διαστασιακά και κατηγορικά μοντέλα συναισθήματος. Επιπλέον, περιγράφεται ο ρόλος των αντικειμενικών ηχητικών χαρακτηριστικών και της μηχανικής μάθησης στην πρόβλεψη μουσικού συναισθήματος.

Στο Κεφάλαιο 3 πραγματοποιείται βιβλιογραφική επισκόπηση σχετικών ερευνών στο πεδίο της πρόβλεψης μουσικού συναισθήματος. Εξετάζονται προσεγγίσεις πρόβλεψης από ηχητικά χαρακτηριστικά, μέθοδοι επιλογής χαρακτηριστικών και ερμηνευσιμότητας, καθώς και η χαρτογράφηση μεταξύ διαφορετικών συναισθηματικών χώρων. Τέλος, αναδεικνύονται τα βασικά κενά της βιβλιογραφίας και η σημασία της παρούσας ΔΕ.

Στο Κεφάλαιο 4 περιγράφεται το σύνολο δεδομένων και η διαδικασία εξαγωγής και προεπεξεργασίας των αντικειμενικών ηχητικών χαρακτηριστικών. Παρουσιάζονται τα σύνολα χαρακτηριστικών που διαμορφώθηκαν και η τελική οργάνωση των δεδομένων για τα πειράματα.

Στο Κεφάλαιο 5 παρουσιάζεται η πειραματική μεθοδολογία. Περιγράφονται τα μοντέλα πρόβλεψης, οι μετρικές αξιολόγησης, η διαδικασία επιλογής χαρακτηριστικών, η διαδικασία χαρτογράφησης μεταξύ συναισθηματικών χώρων και το συνολικό πλαίσιο αξιολόγησης των τελικών μοντέλων.

Στο Κεφάλαιο 6 παρουσιάζονται και αναλύονται τα πειραματικά αποτελέσματα. Συγκρίνονται διαφορετικά μοντέλα και σύνολα χαρακτηριστικών, αξιολογούνται τα τελικά μοντέλα πρόβλεψης και παρουσιάζονται τα αποτελέσματα της επιλογής χαρακτηριστικών, της ερμηνευσιμότητας, της σύγκρισης τελικών μοντέλων και της χαρτογράφησης.

Στο Κεφάλαιο 7 παρουσιάζεται η σημασιολογική ερμηνεία και αναπαράσταση των αποτελεσμάτων. Τα συναισθηματικά προφίλ που προκύπτουν από τα μοντέλα αξιοποιούνται για την πιο κατανοητή περιγραφή του συναισθηματικού χαρακτήρα των μουσικών αποσπασμάτων.

Στο Κεφάλαιο 8 γίνεται συνολική αποτίμηση των αποτελεσμάτων της ΔΕ, παρουσιάζονται τα βασικά συμπεράσματα και προτείνονται κατευθύνσεις για μελλοντική έρευνα.

Στο τέλος της ΔΕ παρατίθεται Παράρτημα, στο οποίο περιλαμβάνεται ο πλήρης πίνακας των μουσικών αποσπασμάτων και των αντίστοιχων συναισθηματικών αξιολογήσεων που αξιοποιήθηκαν στο πλαίσιο της ΔΕ.

2. Θεωρητικό Υπόβαθρο

Στο παρόν Κεφάλαιο παρουσιάζεται το θεωρητικό υπόβαθρο της ΔΕ. Αρχικά, εισάγεται το πεδίο της Ανάκτησης Μουσικής Πληροφορίας (MIR) και η σημασία του για την ανάλυση και οργάνωση μουσικών δεδομένων. Στη συνέχεια, εξετάζεται η συναισθηματική αναπαράσταση στη μουσική, με έμφαση στα διαστασιακά και κατηγορικά μοντέλα συναισθήματος. Επιπλέον, παρουσιάζεται ο ρόλος των αντικειμενικών ηχητικών χαρακτηριστικών και της μηχανικής μάθησης στην πρόβλεψη μουσικού συναισθήματος. Με τον τρόπο αυτό, διαμορφώνεται το θεωρητικό πλαίσιο πάνω στο οποίο βασίζεται η μεθοδολογία της παρούσας ΔΕ.

2.1 Ανάκτηση Μουσικής Πληροφορίας

Η Ανάκτηση Μουσικής Πληροφορίας (Music Information Retrieval, MIR) αποτελεί ένα διεπιστημονικό πεδίο το οποίο περιλαμβάνει διάφορες δραστηριότητες έρευνας και ανάπτυξης με κοινό παρονομαστή την πρόσβαση στη μουσική. Πιο συγκεκριμένα, αφορά κυρίως την εξαγωγή σημαντικών χαρακτηριστικών από τη μουσική, την οργάνωση της μουσικής με βάση αυτά τα χαρακτηριστικά και την ανάπτυξη μηχανών και σχημάτων αναζήτησης και ανάκτησης μουσικών πόρων, όπως αναζήτηση βάσει περιεχομένου (content-based search), συστήματα προτάσεων μουσικής (music recommendation systems) και διεπαφές χρήστη για περιήγηση σε μεγάλες μουσικές συλλογές (user interfaces for browsing large music collections) (Downie, 2003; Orio, 2006; Schedl et al., 2014).

Ως εκ τούτου, οι εφαρμογές της MIR καλύπτουν ένα ευρύ φάσμα εργασιών, όπως η αναγνώριση μουσικού είδους, η αναζήτηση με βάση την ομοιότητα, η αναγνώριση μελωδίας ή ρυθμού, η αυτόματη εξαγωγή μεταδεδομένων και η εκτίμηση συναισθηματικού ή σημασιολογικού περιεχομένου (Orio, 2006; Schedl et al., 2014). Ανάλογα με το είδος της πληροφορίας που αξιοποιείται, οι προσεγγίσεις MIR μπορούν να βασίζονται σε μεταδεδομένα, σε συμβολικές αναπαραστάσεις της μουσικής ή στο ίδιο το ηχητικό σήμα. Στην περίπτωση της ανάλυσης βάσει περιεχομένου (content-based MIR), η έμφαση δίνεται στην εξαγωγή χαρακτηριστικών που περιγράφουν διαφορετικές πτυχές του μουσικού σήματος (Casey et al., 2008).

Στο πλαίσιο της content-based MIR, και ειδικότερα της ανάλυσης του ηχητικού σήματος, το ηχητικό σήμα μετατρέπεται σε αριθμητικές αναπαραστάσεις, οι οποίες μπορούν να χρησιμοποιηθούν από υπολογιστικά μοντέλα για εργασίες ανάλυσης, ταξινόμησης ή

πρόβλεψης. Τα χαρακτηριστικά αυτά μπορεί να αφορούν την ενέργεια, τον ρυθμό, την τονικότητα, το φασματικό περιεχόμενο ή την ηχοχρωματική δομή του σήματος. Με αυτόν τον τρόπο, η μουσική πληροφορία μετατρέπεται σε ένα σύνολο μετρήσιμων μεταβλητών, κατάλληλων για στατιστική ανάλυση και μηχανική μάθηση (Casey et al., 2008; Müller, 2015).

Η παρούσα ΔΕ εντάσσεται στο πεδίο της MIR, καθώς αξιοποιεί αντικειμενικά ηχητικά χαρακτηριστικά για την πρόβλεψη συναισθηματικών αξιολογήσεων μουσικών αποσπασμάτων.

2.2 Συναισθηματική Αναπαράσταση στη Μουσική

Η μελέτη του μουσικού συναισθήματος είναι μια σύνθετη διεργασία, καθώς το συναίσθημα αποτελεί πολυδιάστατο φαινόμενο, το οποίο συνδέεται με υποκειμενικά βιώματα, γνωστικές διεργασίες, εκφραστικές αντιδράσεις και σωματικές αποκρίσεις (Sloboda & Juslin, 2010).

Το συναισθηματικό περιεχόμενο ενός μουσικού αποσπάσματος μπορεί να μελετηθεί είτε ως συναίσθημα που εκφράζεται από τη μουσική ή γίνεται αντιληπτό από τον ακροατή (perceived emotion) είτε ως συναίσθημα που προκαλείται βιωματικά κατά την ακρόαση (induced emotion). Η διάκριση μεταξύ αντίληψης και βιωματικής απόκρισης είναι σημαντική, καθώς ένα μουσικό κομμάτι μπορεί να αναγνωρίζεται ως χαρούμενο ή λυπημένο χωρίς απαραίτητα να προκαλεί τα αντίστοιχα συναισθήματα στον ακροατή, ή ακόμη και να προκαλεί διαφορετική συναισθηματική απόκριση. Η σχέση μεταξύ αντίληψης και απόκρισης είναι σύνθετη, καθώς και οι δύο παράμετροι επηρεάζονται από μουσικούς, προσωπικούς και παράγοντες του πλαισίου ακρόασης (Gabrielsson, 2001; Sloboda & Juslin, 2010).

Στη σχετική βιβλιογραφία έχουν προταθεί διαφορετικοί τρόποι αναπαράστασης του μουσικού συναισθήματος. Οι Eerola & Vuoskoski (2013) διακρίνουν τέσσερις βασικές κατηγορίες μοντέλων συναισθήματος: τα διακριτά ή κατηγορικά μοντέλα (discrete models), τα διαστασιακά μοντέλα (dimensional models), τα λοιπά ή εναλλακτικά μοντέλα (miscellaneous models) και τα μουσικά εξειδικευμένα μοντέλα (music-specific models). Τα κατηγορικά μοντέλα βασίζονται σε συγκεκριμένες συναισθηματικές κατηγορίες, όπως χαρά, λύπη, φόβος ή θυμός, ενώ τα διαστασιακά μοντέλα περιγράφουν το συναίσθημα μέσω συνεχών αξόνων, όπως η θετικότητα/αρνητικότητα (valence) και ο βαθμός ενεργοποίησης

(arousal). Τα λοιπά μοντέλα περιλαμβάνουν έννοιες όπως η ένταση, η προτίμηση, η ομοιότητα ή η εκφραστικότητα, ενώ τα μουσικά εξειδικευμένα μοντέλα επιχειρούν να αποτυπώσουν συναισθηματικές ποιότητες που εμφανίζονται συχνά στη μουσική εμπειρία και δεν καλύπτονται πλήρως από τα γενικά μοντέλα συναισθήματος.

Στην παρούσα ΔΕ, το ενδιαφέρον εστιάζεται στη συναισθηματική αξιολόγηση μουσικών αποσπασμάτων, δηλαδή στον τρόπο με τον οποίο οι ακροατές αποδίδουν συναισθηματικές ιδιότητες στη μουσική. Οι αξιολογήσεις αυτές αναπαρίστανται μέσω διαστασιακών και κατηγορικών μοντέλων, τα οποία παρουσιάζονται αναλυτικότερα στην επόμενη Ενότητα.

2.3 Διαστασιακά και Κατηγορικά Μοντέλα Συναισθήματος

Τα διαστασιακά και τα κατηγορικά μοντέλα αποτελούν δύο από τις πιο διαδεδομένες προσεγγίσεις για την αναπαράσταση του συναισθήματος στη μουσική. Οι δύο προσεγγίσεις διαφοροποιούνται ως προς τον τρόπο περιγραφής του συναισθηματικού περιεχομένου, αλλά μπορούν να λειτουργήσουν συμπληρωματικά, καθώς αποτυπώνουν διαφορετικές όψεις της συναισθηματικής εμπειρίας και αντίληψης (Eerola & Vuoskoski, 2013).

Τα διαστασιακά μοντέλα αποτυπώνουν το συναίσθημα ως σημείο ή περιοχή μέσα σε έναν συνεχή συναισθηματικό χώρο. Η πιο γνωστή προσέγγιση είναι το circumplex model του Russell (1980), σύμφωνα με το οποίο τα συναισθήματα μπορούν να αναπαρασταθούν με βάση δύο βασικές διαστάσεις: τη συναισθηματική θετικότητα ή αρνητικότητα (valence) και τον βαθμό ενεργοποίησης (arousal). Η διάσταση valence εκφράζει κατά πόσο ένα συναίσθημα γίνεται αντιληπτό ως θετικό ή αρνητικό, ενώ η διάσταση arousal περιγράφει το επίπεδο διέγερσης ή ενεργοποίησης που το συνοδεύει. Για παράδειγμα, η χαρά χαρακτηρίζεται συνήθως από θετική valence και υψηλότερη arousal, ενώ η λύπη από αρνητική valence και χαμηλότερη arousal.

Στο πεδίο της μουσικής, τα διαστασιακά μοντέλα έχουν χρησιμοποιηθεί ευρέως, επιτρέποντας την περιγραφή μουσικών αποσπασμάτων ως προς τον βαθμό θετικότητας, ενεργοποίησης ή έντασης που εκφράζει. Σε ορισμένες μουσικές εφαρμογές, η διάσταση arousal διακρίνεται περαιτέρω σε επιμέρους διαστάσεις, όπως Energy και Tension, ώστε να αποτυπωθούν πιο λεπτές διαφοροποιήσεις μεταξύ ενεργητικότητας και έντασης (Eerola & Vuoskoski, 2011).

Τα κατηγορικά ή διακριτά μοντέλα, αντίθετα, περιγράφουν το συναίσθημα μέσω συγκεκριμένων συναισθηματικών κατηγοριών. Η προσέγγιση αυτή συνδέεται με τη θεωρία

των βασικών συναισθημάτων, σύμφωνα με την οποία ορισμένα συναισθήματα, όπως η χαρά, η λύπη, ο φόβος, ο θυμός, η έκπληξη και η αηδία, θεωρούνται διακριτές και σχετικά καθολικές συναισθηματικές κατηγορίες (Ekman, 1992). Στη μουσική, ωστόσο, δεν χρησιμοποιούνται πάντα όλες οι κλασικές κατηγορίες βασικών συναισθημάτων, καθώς ορισμένες από αυτές, όπως η αηδία ή η έκπληξη, είναι λιγότερο κατάλληλες για την περιγραφή μουσικού περιεχομένου. Αντίθετα, συχνά αξιοποιούνται κατηγορίες όπως η τρυφερότητα ή η γαλήνη (Eerola & Vuoskoski, 2013; Scherer, 2004). Πρακτικά, οι κατηγορικές μεταβλητές δεν χρησιμοποιούνται μόνο ως μοναδικές ετικέτες ταξινόμησης, αλλά μπορούν να αποτυπωθούν και ως συνεχείς αξιολογήσεις του βαθμού στον οποίο κάθε συναίσθημα εκφράζεται σε ένα μουσικό απόσπασμα (Eerola & Vuoskoski, 2011).

Η σχέση μεταξύ διαστασιακών και κατηγορικών μοντέλων έχει εξεταστεί στο πλαίσιο της μουσικής από τους Eerola & Vuoskoski (2011) και Eerola et al. (2009). Στις έρευνές τους χρησιμοποιήθηκαν τόσο διαστασιακές αξιολογήσεις, όπως Valence, Energy και Tension, όσο και αξιολογήσεις βασικών συναισθημάτων, όπως Happy, Sad, Anger, Fear και Tender. Τα αποτελέσματα έδειξαν ότι οι δύο μορφές αναπαράστασης παρουσιάζουν σημαντική αντιστοιχία, γεγονός που υποστηρίζει την άποψη ότι μπορούν να θεωρηθούν συμπληρωματικοί τρόποι περιγραφής του μουσικού συναισθήματος.

Στην παρούσα ΔΕ αξιοποιούνται και οι δύο προσεγγίσεις. Οι συναισθηματικές διαστάσεις Valence, Energy και Tension χρησιμοποιούνται ως συνεχείς μεταβλητές-στόχοι, ενώ τα βασικά συναισθήματα Happy, Sad, Anger, Fear και Tender χρησιμοποιούνται για την αποτύπωση του συναισθηματικού προφίλ κάθε μουσικού αποσπάσματος. Με αυτόν τον τρόπο, εξετάζεται τόσο η δυνατότητα πρόβλεψης κάθε συναισθηματικού χώρου από αντικειμενικά ηχητικά χαρακτηριστικά όσο και η σχέση μεταξύ των δύο μορφών συναισθηματικής αναπαράστασης. Η χαρτογράφηση μεταξύ διαστασιακού και κατηγορικού χώρου επιτρέπει να εξεταστεί κατά πόσο οι συνεχείς συναισθηματικές διαστάσεις μπορούν να μεταφραστούν σε προφίλ βασικών συναισθημάτων. Παράλληλα, η αναπαράσταση ενός μουσικού αποσπάσματος ως συναισθηματικού προφίλ μπορεί να υποστηρίξει πιο ερμηνεύσιμες και σημασιολογικά κατανοητές περιγραφές του συναισθηματικού του χαρακτήρα.

2.4 Αντικειμενικά Ηχητικά Χαρακτηριστικά

Τα μουσικά αποσπάσματα μπορούν να αναπαρασταθούν μέσω αντικειμενικών ηχητικών χαρακτηριστικών, δηλαδή αριθμητικών μεγεθών που εξάγονται από το ηχητικό σήμα. Τα χαρακτηριστικά αυτά δεν αντιστοιχούν άμεσα σε συναισθηματικές έννοιες, αλλά αποτελούν μετρήσιμες ιδιότητες του ήχου, οι οποίες μπορούν να αξιοποιηθούν από στατιστικά ή υπολογιστικά μοντέλα για εργασίες ταξινόμησης, πρόβλεψης ή ανάκτησης πληροφορίας (Müller, 2015; Panda et al., 2023).

Τα αντικειμενικά χαρακτηριστικά μπορούν να περιγράψουν διαφορετικά επίπεδα της μουσικής δομής. Ενδεικτικά, χαρακτηριστικά έντασης και δυναμικής αποτυπώνουν την ενεργειακή στάθμη και τις μεταβολές της έντασης του σήματος, ενώ ρυθμικά χαρακτηριστικά σχετίζονται με την ταχύτητα και την οργάνωση του ρυθμικού παλμού. Αντίστοιχα, τονικά και αρμονικά χαρακτηριστικά περιγράφουν πτυχές όπως το τονικό κέντρο, το ύψος και η αρμονική σταθερότητα, ενώ φασματικά και ηχοχρωματικά χαρακτηριστικά αποτυπώνουν την κατανομή της ενέργειας στο φάσμα και τη χροιά του ήχου. Τέτοιες κατηγορίες χαρακτηριστικών έχουν χρησιμοποιηθεί εκτενώς σε εφαρμογές MIR και ιδιαίτερα σε μελέτες αναγνώρισης ή πρόβλεψης μουσικού συναισθήματος (Panda et al., 2023).

Διαφορετικές μουσικές ιδιότητες μπορούν να σχετίζονται με διαφορετικές συναισθηματικές πτυχές μουσικών αποσπασμάτων. Για παράδειγμα, χαρακτηριστικά που αφορούν τον ρυθμό και την ενεργειακή στάθμη συνδέονται συχνά με την αντίληψη της ενεργοποίησης ή της έντασης, ενώ χαρακτηριστικά που σχετίζονται με την τονικότητα, την αρμονία και το ηχόχρωμα μπορούν να συμβάλλουν στην αποτύπωση διαφορών ως προς τη συναισθηματική ποιότητα ενός αποσπάσματος (Lartillot, 2017). Ωστόσο, η σχέση μεταξύ ηχητικών χαρακτηριστικών και συναισθήματος δεν είναι απλή, καθώς το συναισθηματικό περιεχόμενο της μουσικής είναι απόρροια πολλών μουσικών παραμέτρων και επηρεάζεται από τον τρόπο αντίληψης των ακροατών (Panda et al., 2023; Sloboda & Juslin, 2010).

Διάφορα εργαλεία λογισμικού έχουν αναπτυχθεί για την εξαγωγή τέτοιων χαρακτηριστικών στο πεδίο της MIR. Το MIRtoolbox αποτελεί ένα από τα γνωστά εργαλεία για την εξαγωγή μουσικών χαρακτηριστικών από αρχεία ήχου σε περιβάλλον MATLAB, παρέχοντας δυνατότητες ανάλυσης χαρακτηριστικών σχετικών με τον ρυθμό, την τονικότητα, το ηχόχρωμα και τη μορφή (Lartillot & Toivianen, 2007; Lartillot, 2017). Αντίστοιχα, η

Essentia αποτελεί βιβλιοθήκη ανάλυσης ήχου και εξαγωγής μουσικής πληροφορίας, υλοποιημένη σε C++ με διαθέσιμα Python bindings, η οποία περιλαμβάνει αλγόριθμους για την εξαγωγή φασματικών, ρυθμικών, τονικών και υψηλότερου επιπέδου μουσικών χαρακτηριστικών (Bogdanov et al., 2013).

Στην παρούσα ΔΕ, τα αντικειμενικά ηχητικά χαρακτηριστικά χρησιμοποιούνται ως είσοδοι σε μοντέλα μηχανικής μάθησης για την πρόβλεψη συναισθηματικών αξιολογήσεων. Με τον τρόπο αυτό, το μουσικό σήμα μετατρέπεται σε ένα σύνολο μετρήσιμων μεταβλητών, οι οποίες επιτρέπουν τη διερεύνηση της σχέσης μεταξύ ακουστικών ιδιοτήτων και αντιληπτού συναισθηματικού περιεχομένου. Η αναλυτική διαδικασία εξαγωγής και διαμόρφωσης των χαρακτηριστικών που χρησιμοποιήθηκαν στη ΔΕ παρουσιάζεται στο Κεφάλαιο 4.

2.5 Μηχανική Μάθηση για Πρόβλεψη Συναισθήματος στη Μουσική

Η πρόβλεψη συναισθήματος στη μουσική αποτελεί ένα από τα βασικά αντικείμενα της MIR και συχνά αναφέρεται στη βιβλιογραφία ως Music Emotion Recognition (MER). Στο πλαίσιο αυτό, τα μουσικά αποσπάσματα αναπαρίστανται μέσω ηχητικών χαρακτηριστικών, τα οποία χρησιμοποιούνται ως είσοδοι σε υπολογιστικά μοντέλα με στόχο την εκτίμηση του συναισθηματικού τους περιεχομένου. Η προσέγγιση αυτή συνδυάζει την ανάλυση του ηχητικού σήματος με τεχνικές μηχανικής μάθησης (machine learning), επιτρέποντας την αυτόματη πρόβλεψη συναισθηματικών αξιολογήσεων (Kim et al., 2010; Yang & Chen, 2012).

Η σχετική βιβλιογραφία περιλαμβάνει διαφορετικές προσεγγίσεις MER, όπως ταξινόμηση συναισθήματος από συμβολικά αρχεία ή αρχεία ήχου, πολυετικετική ταξινόμηση (multi-label classification), διαστασιακή πρόβλεψη μέσω παλινδρόμησης, ανάλυση μεταβολής συναισθήματος στον χρόνο, αξιοποίηση στίχων και πολυτροπικές προσεγγίσεις. Οι προσεγγίσεις αυτές έχουν υλοποιηθεί τόσο με κλασικές μεθόδους μηχανικής μάθησης όσο και με νεότερες μεθόδους βαθιάς μάθησης (Panda et al., 2023).

Ανάλογα με τον τρόπο αναπαράστασης του συναισθήματος, το πρόβλημα πρόβλεψης ή αναγνώρισης συναισθήματος μπορεί να προσεγγιστεί είτε ως πρόβλημα ταξινόμησης (classification) είτε ως πρόβλημα παλινδρόμησης (regression) (Huq et al., 2010). Στις κατηγορικές προσεγγίσεις, κάθε μουσικό απόσπασμά μπορεί να αντιστοιχεί σε μία ή περισσότερες συναισθηματικές κατηγορίες, όπως χαρά, λύπη ή φόβος. Αντίθετα, στις διαστασιακές προσεγγίσεις, το συναίσθημα αποτυπώνεται μέσω συνεχών τιμών σε δύο ή

περισσότερες διαστάσεις, όπως η Valence, η Energy ή η Tension, και συνεπώς η πρόβλεψη πραγματοποιείται με μεθόδους παλινδρόμησης. Επίσης, οι βασικές συναισθηματικές κατηγορίες μπορούν να μοντελοποιηθούν ως συνεχείς αξιολογήσεις, με στόχο την εκτίμηση του βαθμού στον οποίο κάθε συναίσθημα εκφράζεται στο μουσικό απόσπασμα (Eerola et al., 2009; Eerola & Vuoskoski, 2011).

Ένα τυπικό πλαίσιο μηχανικής μάθησης για MER περιλαμβάνει την εξαγωγή ηχητικών χαρακτηριστικών, τη διαμόρφωση ενός συνόλου δεδομένων όπου κάθε μουσικό απόσπασμα περιγράφεται από ένα σύνολο μεταβλητών, την εκπαίδευση μοντέλων πρόβλεψης και την αξιολόγησή τους σε δεδομένα που δεν χρησιμοποιήθηκαν κατά την εκπαίδευση. Η αξιολόγηση πραγματοποιείται συνήθως με τεχνικές διασταυρούμενης επικύρωσης (cross-validation) και κατάλληλες μετρικές, αναλόγως αν το πρόβλημα αντιμετωπίζεται ως ταξινόμηση ή παλινδρόμηση (Huq et al., 2010; Yang & Chen, 2012).

Τα τελευταία χρόνια, μέθοδοι βαθιάς μάθησης (deep learning) και πολυτροπικές προσεγγίσεις, δηλαδή προσεγγίσεις που αξιοποιούν περισσότερες από μία πηγές πληροφορίας, όπως ήχο, στίχους ή MIDI (Musical Instrument Digital Interface), έχουν αποκτήσει σημαντική παρουσία στο πεδίο της MER, ιδιαίτερα σε περιπτώσεις όπου είναι διαθέσιμα μεγάλα σύνολα δεδομένων. Ωστόσο, οι κλασικές προσεγγίσεις μηχανικής μάθησης που βασίζονται σε εξαγόμενα ηχητικά χαρακτηριστικά εξακολουθούν να είναι σημαντικές, ιδιαίτερα όταν το διαθέσιμο σύνολο δεδομένων είναι περιορισμένου μεγέθους ή όταν στόχος της ανάλυσης είναι η ερμηνευσιμότητα των αποτελεσμάτων (Panda et al., 2023; Kang & Herremans, 2025a).

Παρά τη σημαντική πρόοδο στο πεδίο, η πρόβλεψη μουσικού συναισθήματος παραμένει δύσκολο πρόβλημα. Οι ανθρώπινες συναισθηματικές αξιολογήσεις χαρακτηρίζονται από υποκειμενικότητα, ενώ η σχέση μεταξύ ηχητικών χαρακτηριστικών και συναισθηματικού περιεχομένου είναι ιδιαίτερα σύνθετη. Επίσης, το μέγεθος και η ποιότητα των διαθέσιμων συνόλων δεδομένων, καθώς και η δυνατότητα γενίκευσης των μοντέλων σε διαφορετικά μουσικά είδη αποτελούν σημαντικές προκλήσεις (Kang & Herremans, 2025a).

Στην παρούσα ΔΕ, η πρόβλεψη συναισθηματικών αξιολογήσεων αντιμετωπίζεται ως πρόβλημα παλινδρόμησης. Τα αντικειμενικά ηχητικά χαρακτηριστικά χρησιμοποιούνται ως είσοδοι σε μοντέλα μηχανικής μάθησης, με στόχο την πρόβλεψη τόσο συναισθηματικών

διαστάσεων όσο και βασικών συναισθημάτων. Η αναλυτική περιγραφή των μοντέλων, των μετρικών αξιολόγησης και της πειραματικής διαδικασίας παρουσιάζεται στο Κεφάλαιο 5.

3. Βιβλιογραφική Επισκόπηση

Στο παρόν Κεφάλαιο παρουσιάζεται η βιβλιογραφική επισκόπηση της ΔΕ, με έμφαση στις προσεγγίσεις πρόβλεψης μουσικού συναισθήματος από ηχητικά χαρακτηριστικά. Εξετάζονται οι διαστασιακές και κατηγορικές προσεγγίσεις, οι μέθοδοι επιλογής χαρακτηριστικών και ερμηνευσιμότητας, καθώς και η χαρτογράφηση μεταξύ συναισθηματικών χώρων. Τέλος, αναδεικνύονται τα βασικά κενά της βιβλιογραφίας και η σημασία της παρούσας ΔΕ.

3.1 Πρόβλεψη Μουσικού Συναισθήματος από Ηχητικά Χαρακτηριστικά

Οι πρώτες προσεγγίσεις MER βασίστηκαν σε χειροποίητα ηχητικά χαρακτηριστικά (handcrafted audio features), τα οποία περιγράφουν πτυχές όπως η ένταση, ο ρυθμός, η τονικότητα, το φασματικό περιεχόμενο και το ηχόχρωμα. Τα χαρακτηριστικά αυτά χρησιμοποιήθηκαν ως είσοδοι σε μοντέλα μηχανικής μάθησης για την ταξινόμηση ή παλινδρόμηση συναισθηματικών αξιολογήσεων. Η προσέγγιση αυτή επιτρέπει τη διερεύνηση της σχέσης μεταξύ συγκεκριμένων μουσικών ιδιοτήτων και συναισθηματικής αντίληψης, ενώ παράλληλα παρέχει μεγαλύτερη ερμηνευσιμότητα σε σχέση με πιο σύνθετα μοντέλα που λειτουργούν ως μαύρο κουτί (black box) (Yang & Chen, 2012; Panda et al., 2023).

Σημαντικό ρόλο στην εξέλιξη του πεδίου έχουν συντελέσει και οι προσπάθειες δημιουργίας κοινών διαδικασιών αξιολόγησης. Το Music Information Retrieval Evaluation eXchange (MIREX) Audio Mood Classification task αποτελεί μια ετήσια benchmark διαδικασία για την αξιολόγηση αλγορίθμων MER, επιτρέποντας τη σύγκριση διαφορετικών προσεγγίσεων πάνω σε κοινά δεδομένα και με κοινές μετρικές (Panda et al., 2023). Ωστόσο, η πρόοδος δεν ήταν πάντα θεαματική. Για την περίοδο 2007-2019, οι Panda et al. (2023) επισημαίνουν ότι οι βελτιώσεις παρέμειναν περιορισμένες, με την καλύτερη επίδοση να φτάνει περίπου στο 69.8% ακρίβεια (accuracy) σε πρόβλημα πέντε κατηγοριών. Το γεγονός αυτό επιβεβαιώνει τη δυσκολία του προβλήματος και την ανάγκη για χαρακτηριστικά που αποτυπώνουν αποτελεσματικά το συναισθηματικό περιεχόμενο της μουσικής.

Παράλληλα με τις κατηγορικές προσεγγίσεις, σημαντική κατεύθυνση αποτέλεσε η πρόβλεψη συνεχών συναισθηματικών αξιολογήσεων. Οι Eerola et al. (2009) διερεύνησαν την πρόβλεψη πολυδιάστατων συναισθηματικών αξιολογήσεων από ηχητικά

χαρακτηριστικά με χρήση μοντέλων παλινδρόμησης. Η εργασία αυτή είναι ιδιαίτερα σημαντική για την παρούσα ΔΕ, καθώς συνδέει άμεσα τα αντικειμενικά χαρακτηριστικά μουσικού σήματος με διαστασιακές και κατηγορικές αξιολογήσεις συναισθήματος, προσφέροντας ένα πλαίσιο σύγκρισης για την πρόβλεψη συναισθηματικών διαστάσεων και βασικών συναισθημάτων.

Τα τελευταία χρόνια, το πεδίο έχει επεκταθεί με τη χρήση βαθιάς μάθησης και πολυτροπικών δεδομένων. Οι Kang & Herremans (2025a) πραγματοποίησαν μια επισκόπηση διαθέσιμων συνόλων δεδομένων και μοντέλων που έχουν αναπτυχθεί για πρόβλεψη μουσικού συναισθήματος μετά το 2020, αναδεικνύοντας τη μετάβαση από αποκλειστικά *handcrafted* χαρακτηριστικά σε πιο σύνθετες αρχιτεκτονικές και αναπαραστάσεις, λιγότερο εξαρτημένες από τη χειροκίνητη διαμόρφωση χαρακτηριστικών. Παρόλα αυτά, παραμένουν σημαντικές προκλήσεις, όπως η υποκειμενικότητα των συναισθηματικών αξιολογήσεων, η ανομοιογένεια των συνόλων δεδομένων, η δυσκολία γενίκευσης σε διαφορετικά μουσικά είδη και η ανάγκη για ερμηνεύσιμα μοντέλα.

Στη βιβλιογραφία έχουν χρησιμοποιηθεί διαφορετικοί αλγόριθμοι μηχανικής μάθησης, όπως Support Vector Machines, k-Nearest Neighbors, Random Forests, μοντέλα παλινδρόμησης και νευρωνικά δίκτυα, ανάλογα με τη φύση του προβλήματος και το διαθέσιμο σύνολο δεδομένων (Kim et al., 2010; Yang & Chen, 2012; Panda et al., 2015; Kang & Herremans, 2025a). Σε νεότερες εργασίες, ιδιαίτερη έμφαση δίνεται σε συνελκτικά και αναδρομικά νευρωνικά δίκτυα, transformers και embeddings, ειδικά σε μεγάλα ή πολυτροπικά σύνολα δεδομένων (Kang & Herremans, 2025a).

Στο πλαίσιο αυτό, οι προσεγγίσεις που βασίζονται σε αντικειμενικά ηχητικά χαρακτηριστικά εξακολουθούν να παρουσιάζουν ιδιαίτερο ενδιαφέρον. Αν και οι τεχνικές βαθιάς μάθησης μπορούν να επιτύχουν υψηλές επιδόσεις σε μεγάλα σύνολα δεδομένων, τα μοντέλα που αξιοποιούν *handcrafted* χαρακτηριστικά επιτρέπουν πιο άμεση ερμηνεία της σχέσης μεταξύ μουσικών ιδιοτήτων και συναισθηματικών αξιολογήσεων. Η παρούσα ΔΕ εντάσσεται σε αυτήν τη γραμμή έρευνας, καθώς αξιοποιεί αντικειμενικά ηχητικά χαρακτηριστικά και μοντέλα μηχανικής μάθησης για την πρόβλεψη τόσο συναισθηματικών διαστάσεων όσο και βασικών συναισθημάτων.

3.2 Διαστασιακές και Κατηγορικές Προσεγγίσεις

Με βάση τα θεωρητικά μοντέλα που παρουσιάστηκαν στο προηγούμενο Κεφάλαιο, η βιβλιογραφία της MER έχει ακολουθήσει δύο βασικές γραμμές: την κατηγορική αναγνώριση συναισθήματος και τη διαστασιακή πρόβλεψη συνεχών συναισθηματικών αξιολογήσεων. Οι δύο προσεγγίσεις διαφοροποιούνται ως προς τον τύπο συναισθηματικής αναπαράστασης και, επομένως, ως προς τον τρόπο διατύπωσης του προβλήματος μηχανικής μάθησης (Yang & Chen, 2012).

Στις κατηγορικές προσεγγίσεις, το μουσικό συναίσθημα αναπαρίσταται μέσω διακριτών ετικετών ή ομάδων ετικετών. Η λογική αυτή έχει χρησιμοποιηθεί σε προβλήματα ταξινόμησης, όπου κάθε μουσικό απόσπασμα αντιστοιχίζεται σε μία ή περισσότερες συναισθηματικές κατηγορίες. Χαρακτηριστικό παράδειγμα αποτελεί το MIREX Audio Mood Classification Task, στο οποίο τα μουσικά αποσπάσματα οργάνωθηκαν σε πέντε ομάδες διάθεσης και η αξιολόγηση των μοντέλων πραγματοποιήθηκε ως πρόβλημα ταξινόμησης (Hu et al., 2008; Panda et al., 2023). Οι κατηγορικές προσεγγίσεις μπορεί να προσφέρουν εύκολα ερμηνεύσιμες ετικέτες, ωστόσο η αποτύπωση σύνθετων ή μικτών συναισθηματικών αποχρώσεων παραμένει πρόκληση στο πεδίο.

Αντίθετα, οι διαστασιακές προσεγγίσεις αντιμετωπίζουν το συναίσθημα ως συνεχές μέγεθος σε έναν συναισθηματικό χώρο, συνήθως με άξονες όπως valence και arousal. Στο πλαίσιο αυτό, η πρόβλεψη αντιμετωπίζεται ως πρόβλημα παλινδρόμησης. Η προσέγγιση αυτή επιτρέπει την αποτύπωση λεπτότερων συναισθηματικών διαφορών μεταξύ μουσικών αποσπασμάτων και έχει χρησιμοποιηθεί σε αρκετά σύνολα δεδομένων (Yang & Chen, 2012; Kang & Herremans, 2025a).

Η σχέση μεταξύ των δύο προσεγγίσεων έχει εξεταστεί ειδικά στη μουσική από τους Eerola & Vuoskoski (2011), οι οποίοι συνέκριναν διακριτά και διαστασιακά μοντέλα συναισθήματος σε μουσικά αποσπάσματα κινηματογραφικής μουσικής. Η μελέτη τους έδειξε ότι και οι δύο τύποι αναπαράστασης μπορούν να περιγράψουν σε σημαντικό βαθμό το ίδιο συναισθηματικό περιεχόμενο, αλλά με διαφορετικά πλεονεκτήματα. Τα διαστασιακά μοντέλα παρέχουν πιο συνεχή περιγραφή του συναισθήματος, ενώ τα κατηγορικά μοντέλα προσφέρουν πιο άμεσες και κατανοητές συναισθηματικές ετικέτες. Το σύνολο δεδομένων που προέκυψε από την έρευνα τους αξιοποιήθηκε επίσης από τους Eerola et al. (2009), οι οποίοι διερεύνησαν την πρόβλεψη συναισθηματικών αξιολογήσεων

από ηχητικά χαρακτηριστικά τόσο σε διαστασιακό όσο και σε κατηγορικό επίπεδο, χρησιμοποιώντας μοντέλα παλινδρόμησης. Οι συγκεκριμένες έρευνες είναι ιδιαίτερα σχετικές με την παρούσα ΔΕ, καθώς η πρώτη συνδέεται με το σύνολο δεδομένων που χρησιμοποιήθηκε, ενώ η δεύτερη αποτελεί βασικό σημείο αναφοράς για τη σύγκριση των τελικών μοντέλων.

Στη σύγχρονη βιβλιογραφία, η συνύπαρξη κατηγορικών και διαστασιακών τύπων συναισθηματικής αναπαράστασης παραμένει ενεργό ζήτημα. Οι Kang & Herremans (2025a) επισημαίνουν ότι τα διαθέσιμα σύνολα δεδομένων διαφέρουν σημαντικά ως προς τον τρόπο αναπαράστασης του συναισθήματος, καθώς άλλα χρησιμοποιούν κατηγορικές ετικέτες και άλλα διαστασιακές αξιολογήσεις. Πρόσφατες προσεγγίσεις επιχειρούν ακόμη και την ενοποίηση των δύο τύπων συναισθηματικής αναπαράστασης σε κοινά πλαίσια μάθησης, με στόχο την αξιοποίηση πολλαπλών συνόλων δεδομένων και τη βελτίωση της γενίκευσης των μοντέλων (Kang & Herremans, 2025b). Η εργασία των Kang & Herremans (2025b) βρίσκεται σε στάδιο προδημοσίευσης και αποτελεί μια ένδειξη σύγχρονης ερευνητικής τάσης.

Η ΔΕ εντάσσεται σε αυτήν την ερευνητική κατεύθυνση, επιχειρώντας την πρόβλεψη τόσο συναισθηματικών διαστάσεων όσο και βασικών συναισθημάτων, καθώς και τη διερεύνηση της σχέσης των δύο χώρων μέσω χαρτογράφησης.

3.3 Επιλογή Χαρακτηριστικών και Ερμηνευσιμότητα

Η επιλογή χαρακτηριστικών αποτελεί σημαντικό στάδιο σε προσεγγίσεις MER. Στόχος είναι ο εντοπισμός των χαρακτηριστικών που συμβάλλουν περισσότερο στην πρόβλεψη, η μείωση της πολυπλοκότητας των μοντέλων και η ενίσχυση της ερμηνευσιμότητας των αποτελεσμάτων. Οι Panda et al. (2023) υπογραμμίζουν ότι η πρόοδος στο MER δεν εξαρτάται μόνο από τη χρήση πιο σύνθετων μοντέλων, αλλά και από τον σχεδιασμό και την επιλογή χαρακτηριστικών που αποτυπώνουν ουσιαστική συναισθηματική πληροφορία.

Στη βιβλιογραφία έχουν αξιοποιηθεί διαφορετικές στρατηγικές για τη μείωση του αριθμού ή της πολυπλοκότητας των χαρακτηριστικών. Οι Zhang et al. (2016) διακρίνουν τρεις βασικές κατηγορίες μεθόδων στο πλαίσιο της MER: μεθόδους φίλτρου (filter methods), wrapper μεθόδους και μεθόδους μείωσης διαστατικότητας (dimension reduction methods). Οι μέθοδοι φίλτρου επιλέγουν χαρακτηριστικά με βάση κριτήρια ανεξάρτητα από το τελικό μοντέλο, όπως η συσχέτιση ή η πληροφοριακή συνεισφορά κάθε χαρακτηριστικού.

Ενδεικτικές μέθοδοι είναι οι Regressional Relief-F (RReliefF), Information Gain Evaluator και Correlation-based Feature Selection (CFS). Οι wrapper μέθοδοι, όπως είναι η Forward Selection (FS), η Backward Elimination (BE) και η Hybrid Forward and Backward Stepwise Selection (HS), αξιολογούν υποσύνολα χαρακτηριστικών μέσω της απόδοσης ενός συγκεκριμένου αλγορίθμου μάθησης. Αντίθετα, οι μέθοδοι μείωσης διαστατικότητας, όπως η Principal Component Analysis (PCA), μετασχηματίζουν το αρχικό σύνολο μεταβλητών σε νέες συνιστώσες. Αν και οι μέθοδοι αυτές μπορούν να βελτιώσουν την απόδοση ή να αντιμετωπίσουν προβλήματα πολυσυγγραμμικότητας, ενδέχεται να μειώσουν την άμεση ερμηνευσιμότητα των αποτελεσμάτων, καθώς οι νέες συνιστώσες δεν αντιστοιχούν σε μεμονωμένα ηχητικά χαρακτηριστικά (Eerola et al., 2009; Zhang et al., 2016).

Η ερμηνευσιμότητα αποτελεί επίσης σημαντικό ζήτημα στο πεδίο της MER. Ένα μοντέλο μπορεί να παρουσιάζει ικανοποιητική απόδοση, αλλά η χρησιμότητά του περιορίζεται αν δεν είναι σαφές ποια ηχητικά χαρακτηριστικά συμβάλλουν στις προβλέψεις του. Σε αυτό το πλαίσιο, οι Chowdhury et al. (2019) πρότειναν τη χρήση χαρακτηριστικών μεσαίου επιπέδου (mid-level perceptual features), δηλαδή χαρακτηριστικών που βρίσκονται πιο κοντά στον τρόπο με τον οποίο οι ακροατές περιγράφουν τη μουσική, με στόχο τη γεφύρωση του χάσματος μεταξύ χαμηλού επιπέδου ηχητικών ιδιοτήτων και υψηλού επιπέδου συναισθηματικών εννοιών.

Επιπλέον, τεχνικές ερμηνευσιμότητας όπως η μέθοδος SHAP, μπορούν να χρησιμοποιηθούν για την ποσοτικοποίηση της συμβολής των επιμέρους χαρακτηριστικών στις προβλέψεις ενός μοντέλου. Η μέθοδος SHAP βασίζεται στις τιμές Shapley από τη θεωρία παιγνίων και επιτρέπει την εκτίμηση της επίδρασης κάθε χαρακτηριστικού τόσο σε μεμονωμένες προβλέψεις όσο και σε συνολικό επίπεδο μοντέλου (Lundberg & Lee, 2017; Shapley, 1953). Οι Li et al. (2025) χρησιμοποίησαν τη μέθοδο SHAP για την ερμηνεία προβλέψεων μουσικού συναισθήματος. Παράλληλα, στη γενικότερη βιβλιογραφία της μηχανικής μάθησης, οι τιμές SHAP έχουν αξιοποιηθεί και για επιλογή χαρακτηριστικών, καθώς επιτρέπουν την κατάταξη και αξιολόγηση της συνεισφοράς των αρχικών μεταβλητών (Kraev et al., 2024).

Η παρούσα ΔΕ αξιοποιεί τη μέθοδο SHAP σε συνδυασμό με διαδικασία ανάλυσης σταθερότητας μέσω διασταυρούμενης επικύρωσης (cross-validation) για την επιλογή καταλληλότερων μειωμένων συνόλων χαρακτηριστικών, καθώς και για την ερμηνεία των τελικών μοντέλων.

3.4 Χαρτογράφηση μεταξύ Συναισθηματικών Χώρων

Η διερεύνηση της σχέσης μεταξύ διαστασιακών και κατηγορικών μοντέλων αποτελεί επίσης ερευνητικό αντικείμενο στο πεδίο της MER. Η σχέση μεταξύ των δύο συναισθηματικών χώρων έχει εξεταστεί από τους Eerola & Vuoskoski (2011), οι οποίοι συνέκριναν διακριτά και διαστασιακά μοντέλα συναισθήματος στη μουσική. Τα αποτελέσματα έδειξαν υψηλή αντιστοιχία μεταξύ των δύο προσεγγίσεων, ιδιαίτερα σε βασικές διαστάσεις που σχετίζονται με τη θετικότητα/αρνητικότητα και την ενεργοποίηση. Επισημαίνουν, ωστόσο, ότι τα κατηγορικά μοντέλα μπορεί να έχουν χαμηλότερη ανάλυση στην περιγραφή συναισθηματικά ασαφών παραδειγμάτων, καθώς ορισμένα μουσικά αποσπάσματα δεν αντιστοιχούν καθαρά σε μία μόνο συναισθηματική κατηγορία.

Σε υπολογιστικό επίπεδο, οι Eerola et al. (2009) αξιοποίησαν ηχητικά χαρακτηριστικά για την πρόβλεψη τόσο διαστασιακών όσο και κατηγορικών συναισθηματικών αξιολογήσεων. Η εργασία τους δείχνει ότι οι δύο χώροι μπορούν να εξεταστούν παράλληλα στο ίδιο πειραματικό πλαίσιο και αποτελεί βασικό σημείο αναφοράς για τη σύγκριση μοντέλων που επιχειρούν να συνδέσουν αντικειμενικά ηχητικά χαρακτηριστικά με διαφορετικές μορφές συναισθηματικής αναπαράστασης. Όπως έχει ήδη επισημανθεί, η εργασία αυτή είναι ιδιαίτερα σχετική με την παρούσα ΔΕ, καθώς χρησιμοποιείται το ίδιο σύνολο δεδομένων και εξετάζεται η σχέση μεταξύ των συναισθηματικών διαστάσεων και των βασικών συναισθημάτων.

Το ζήτημα της σύνδεσης μεταξύ κατηγορικών και διαστασιακών αναπαραστάσεων παραμένει επίκαιρο. Ορισμένα σύνολα δεδομένων διαθέτουν κατηγορικές ετικέτες, άλλα συνεχείς διαστασιακές αξιολογήσεις, ενώ λίγα περιλαμβάνουν και τις δύο μορφές. Η ετερογένεια αυτή δυσκολεύει τη σύγκριση μοντέλων και τη γενίκευση μεταξύ διαφορετικών συνόλων δεδομένων (Kang & Herremans, 2025a). Σε πρόσφατη εργασία σε στάδιο προδημοσίευσης, οι Kang & Herremans (2025b) διερευνούν την αξιοποίηση διαστασιακών και κατηγορικών δεδομένων σε κοινό πλαίσιο μάθησης.

Η παρούσα ΔΕ αξιοποιεί τη συγκεκριμένη ερευνητική κατεύθυνση εξετάζοντας τη χαρτογράφηση από τον διαστασιακό χώρο Valence-Energy-Tension προς τον χώρο βασικών συναισθημάτων Happy-Sad-Anger-Fear-Tender. Η προσέγγιση αυτή συγκρίνεται στη συνέχεια με την άμεση πρόβλεψη βασικών συναισθημάτων από τα αντικειμενικά ηχητικά χαρακτηριστικά.

3.5 Σημασιολογική Αναπαράσταση Μουσικού Συναισθήματος

Ένα βασικό ζήτημα στο MER είναι το χάσμα μεταξύ χαμηλού επιπέδου ηχητικών χαρακτηριστικών και υψηλού επιπέδου ανθρώπινων εννοιών, το οποίο συχνά αναφέρεται ως σημασιολογικό χάσμα (semantic gap). Στο πεδίο της μουσικής, το χάσμα αυτό αφορά τη δυσκολία σύνδεσης αντικειμενικών ακουστικών ιδιοτήτων, όπως η ένταση, το φάσμα ή ο ρυθμός, με έννοιες που χρησιμοποιούνται από τους ακροατές για να περιγράψουν τη μουσική, όπως διάθεση, συναίσθημα, ύφος ή είδος (Celma et al., 2006; Panda et al., 2023).

Οι Panda et al. (2023), σε επισκόπηση των αντικειμενικών χαρακτηριστικών που χρησιμοποιούνται στο MER, τονίζουν ότι η γεφύρωση του σημασιολογικού χάσματος απαιτεί χαρακτηριστικά ικανά να αποτυπώσουν συναισθηματικά σχετική πληροφορία. Στην έρευνά τους εξετάζονται χαρακτηριστικά διαφορετικών επιπέδων, από χαμηλού επιπέδου ηχητικά χαρακτηριστικά, όπως φασματικά χαρακτηριστικά, μέχρι αντιληπτικά και υψηλότερου επιπέδου χαρακτηριστικά, όπως η τονικότητα, το είδος ή η χορευτικότητα. Παρόμοια κατεύθυνση ακολούθησαν οι Chowdhury et al. (2019), οι οποίοι πρότειναν την αξιοποίηση χαρακτηριστικών μεσαίου επιπέδου (mid-level perceptual features) με στόχο τη γεφύρωση του σημασιολογικού χάσματος μεταξύ χαμηλού και υψηλού επιπέδου συναισθηματικών εννοιών, καθώς και την ενίσχυση της ερμηνευσιμότητας των αποτελεσμάτων.

Μια ακόμα χρήσιμη προσέγγιση για την κατανόηση του σημασιολογικού χάσματος προτείνεται από τους Zhang et al. (2016), οι οποίοι διακρίνουν τα μουσικά χαρακτηριστικά σε τέσσερα επίπεδα: φυσικό, αντιληπτικό, μουσικό και σημασιολογικό επίπεδο. Το φυσικό επίπεδο περιλαμβάνει βασικές ακουστικές ιδιότητες, όπως το πλάτος και η συχνότητα, ενώ το αντιληπτικό επίπεδο περιλαμβάνει χαρακτηριστικά όπως η ένταση, το ηχόχρωμα και το τονικό ύψος. Το μουσικό επίπεδο αφορά μουσικολογικά χαρακτηριστικά, όπως η τονικότητα, η μελωδία και οι συγχορδίες, ενώ το σημασιολογικό επίπεδο περιλαμβάνει πιο αφηρημένες έννοιες, όπως το θέμα, το συναίσθημα, το είδος και τους στίχους. Σύμφωνα με αυτήν την προσέγγιση, όσο πιο υψηλό είναι το επίπεδο αναπαράστασης, τόσο πιο άμεσα μπορεί να συνδεθεί με το μουσικό συναίσθημα, αλλά τόσο δυσκολότερη γίνεται η αυτόματη εξαγωγή του.

Σε ευρύτερο επίπεδο, η σημασιολογική αναπαράσταση μουσικής πληροφορίας έχει εξεταστεί και μέσα από μουσικές οντολογίες και τεχνολογίες σημασιολογικού ιστού

(semantic web). Η Music Ontology, για παράδειγμα, επιχειρεί να παρέχει ένα εννοιολογικό πλαίσιο για την περιγραφή μουσικών πόρων, συνδέοντας πληροφορίες όπως καλλιτέχνες, έργα, ηχογραφήσεις και μεταδεδομένα (Raimond et al., 2007). Τέτοιες προσεγγίσεις δεν εστιάζουν απαραίτητα στη χρήση μοντέλων πρόβλεψης, ωστόσο αναδεικνύουν τη σημασία της οργάνωσης μουσικής πληροφορίας σε πιο κατανοητές και επαναχρησιμοποιήσιμες σημασιολογικές δομές.

Η παρούσα ΔΕ αξιοποιεί την ιδέα της σημασιολογικής αναπαράστασης σε βασικό επίπεδο. Συγκεκριμένα, οι προβλέψεις των μοντέλων οργανώνονται ως συναισθηματικά προφίλ, τα οποία μπορούν να υποστηρίξουν πιο κατανοητές λεκτικές περιγραφές του συναισθηματικού χαρακτήρα των μουσικών αποσπασμάτων.

3.6 Κενά της Βιβλιογραφίας και Σημασία της Διπλωματικής Εργασίας

Από τη βιβλιογραφική επισκόπηση προκύπτει ότι η πρόβλεψη μουσικού συναισθήματος αποτελεί ενεργό και πολυδιάστατο ερευνητικό πεδίο. Έχουν αναπτυχθεί διαφορετικές προσεγγίσεις για την πρόβλεψη συναισθηματικών αξιολογήσεων από ηχητικά χαρακτηριστικά, τόσο σε κατηγορικό όσο και σε διαστασιακό επίπεδο. Τα τελευταία χρόνια παρατηρείται αυξημένο ενδιαφέρον για πιο σύνθετα μοντέλα, πολυτροπικά δεδομένα και τεχνικές ερμηνευσιμότητας. Παρά την πρόοδο που έχει σημειωθεί, ορισμένα ζητήματα παραμένουν ανοιχτά.

Η σχέση μεταξύ αντικειμενικών ηχητικών χαρακτηριστικών και συναισθηματικών αξιολογήσεων παραμένει σύνθετη. Παρά την ύπαρξη πολλών χαρακτηριστικών για την περιγραφή διαφορετικών πτυχών του μουσικού σήματος, δεν είναι πάντα σαφές ποια από αυτά συμβάλλουν ουσιαστικά στην πρόβλεψη συγκεκριμένων συναισθηματικών μεταβλητών. Ως εκ τούτου, η επιλογή χαρακτηριστικών και η ερμηνεία της συμβολής τους εξακολουθούν να αποτελούν σημαντικές ανάγκες στο πεδίο.

Επιπλέον, στη βιβλιογραφία εξετάζονται τόσο διαστασιακές όσο και κατηγορικές μορφές συναισθηματικής αναπαράστασης, αλλά η σχέση μεταξύ των δύο χώρων παραμένει ακόμα υπό διερεύνηση. Η δυνατότητα χαρτογράφησης από συναισθηματικές διαστάσεις σε βασικά συναισθήματα μπορεί να προσφέρει ένα ενδιάμεσο επίπεδο ερμηνείας, ωστόσο χρειάζεται περαιτέρω διερεύνηση, ιδιαίτερα όταν οι διαστάσεις δεν είναι πραγματικές ανθρώπινες αξιολογήσεις αλλά προβλέψεις μοντέλων.

Αν και οι προσεγγίσεις βαθιάς μάθησης μπορούν να επιτύχουν υψηλές επιδόσεις σε μεγάλα σύνολα δεδομένων, παραμένει σημαντική η ανάπτυξη ερμηνεύσιμων μοντέλων. Σε αυτό το πλαίσιο, οι κλασικές προσεγγίσεις μηχανικής μάθησης που βασίζονται σε αντικειμενικά ηχητικά χαρακτηριστικά εξακολουθούν να είναι χρήσιμες, καθώς επιτρέπουν πιο άμεση σύνδεση μεταξύ μουσικών ιδιοτήτων και συναισθηματικών αξιολογήσεων.

Η παρούσα ΔΕ διερευνά την πρόβλεψη συναισθηματικών αξιολογήσεων από αντικειμενικά ηχητικά χαρακτηριστικά με χρήση μοντέλων παλινδρόμησης. Η συμβολή της έγκειται στη συγκριτική αξιολόγηση διαφορετικών συνόλων χαρακτηριστικών και μοντέλων, στην εφαρμογή διαδικασιών επιλογής χαρακτηριστικών και ερμηνευσιμότητας μέσω SHAP και ανάλυσης σταθερότητας, καθώς και στη διερεύνηση της σχέσης μεταξύ διαστασιακών και κατηγορικών συναισθηματικών χώρων μέσω χαρτογράφησης. Επιπλέον, η ΔΕ επιχειρεί να οργανώσει τις προβλέψεις σε συναισθηματικά προφίλ, με στόχο μια πιο κατανοητή σημασιολογική αναπαράσταση του συναισθηματικού χαρακτήρα των μουσικών αποσπασμάτων.

4. Εξαγωγή και Ανάλυση Ηχητικών Παραμέτρων

Στο Κεφάλαιο αυτό παρουσιάζονται τα δεδομένα στα οποία βασίστηκε η παρούσα ΔΕ, καθώς και η διαδικασία διαμόρφωσης των αντικειμενικών ηχητικών χαρακτηριστικών που χρησιμοποιήθηκαν στα πειράματα. Συγκεκριμένα, περιγράφονται το σύνολο δεδομένων, οι συναισθηματικές μεταβλητές-στόχοι, η εξαγωγή και κατηγοριοποίηση των χαρακτηριστικών, καθώς και η προεπεξεργασία και τελική οργάνωση των δεδομένων.

4.1 Περιγραφή Συνόλου Δεδομένων

Το σύνολο δεδομένων που χρησιμοποιήθηκε στην παρούσα ΔΕ προέρχεται από τη μελέτη των Eerola & Vuoskoski (2011), η οποία εξετάζει τη σύγκριση διαστασιακών και κατηγορικών μοντέλων συναισθήματος στη μουσική. Το σύνολο δεδομένων είναι διαθέσιμο μέσω του αποθετηρίου Open Science Framework (OSF) του Center for Open Science (COS) (<https://osf.io/p6vkg>).

Το σύνολο δεδομένων περιλαμβάνει 110 σύντομα μουσικά αποσπάσματα από κομμάτια κινηματογραφικής μουσικής. Τα αποσπάσματα αποτελούν αντιπροσωπευτικά παραδείγματα συγκεκριμένων συναισθηματικών κατηγοριών και συναισθηματικών διαστάσεων. Ειδικότερα, το σύνολο καλύπτει αφενός πέντε βασικές κατηγορίες συναισθήματος (Happy, Sad, Anger, Fear, Tender) και αφετέρου τρεις συναισθηματικές διαστάσεις (Valence, Energy, Tension) επιτρέποντας τη συστηματική διερεύνηση της σχέσης ανάμεσα στα δύο μοντέλα συναισθηματικής αναπαράστασης. Για κάθε απόσπασμα είναι διαθέσιμες μέσες ανθρώπινες αξιολογήσεις από 116 συμμετέχοντες σε κλίμακα 1-9 για τις παραπάνω μεταβλητές, καθώς και επιπλέον αξιολογήσεις όπως Beauty και Liking. Στην παρούσα ΔΕ ως στόχοι πρόβλεψης χρησιμοποιήθηκαν οι τρεις συναισθηματικές διαστάσεις και οι πέντε βασικές κατηγορίες συναισθήματος.

Στο αποθετήριο είναι επίσης διαθέσιμα μεταδεδομένα σχετικά με την προέλευση των αποσπασμάτων, όπως το μουσικό άλμπουμ, ο αριθμός του μουσικού κομματιού και το χρονικό σημείο από το οποίο εξήχθη κάθε απόσπασμα. Η διάρκεια των αποσπασμάτων κυμαίνεται από 11 έως 31 δευτερόλεπτα, με μέση διάρκεια περίπου 15 δευτερόλεπτα. Ο πλήρης πίνακας των μουσικών αποσπασμάτων του συνόλου δεδομένων, μαζί με τα βασικά μεταδεδομένα και τις αντίστοιχες συναισθηματικές αξιολογήσεις, παρατίθεται στο Παράρτημα Α.

Η επιλογή του συγκεκριμένου συνόλου δεδομένων θεωρείται ιδιαίτερα κατάλληλη για τους στόχους της διπλωματικής, καθώς επιτρέπει την παράλληλη μελέτη τόσο διαστασιακών όσο και κατηγορικών μοντέλων συναισθήματος με χρήση συστηματικά επιλεγμένων και αξιολογημένων μουσικών αποσπασμάτων.

4.2 Συναισθηματικές Αξιολογήσεις και Στόχοι Πρόβλεψης

Στο πλαίσιο της παρούσας ΔΕ χρησιμοποιήθηκαν ως μεταβλητές-στόχοι δύο διαφορετικές αλλά συμπληρωματικές μορφές συναισθηματικής αναπαράστασης της μουσικής: η διαστασιακή και η κατηγορική. Η αξιολόγηση του ίδιου συνόλου δεδομένων και με τις δύο προσεγγίσεις επιτρέπει τόσο την ανεξάρτητη πρόβλεψη των δύο ειδών συναισθηματικής περιγραφής όσο και τη διερεύνηση της μεταξύ τους σχέσης.

Η πρώτη ομάδα στόχων αφορά τις τρεις συναισθηματικές διαστάσεις Valence, Energy και Tension. Οι διαστάσεις αυτές χρησιμοποιούνται για τη συνεχή αναπαράσταση του συναισθηματικού περιεχομένου των μουσικών αποσπασμάτων. Η Valence περιγράφει το βαθμό θετικότητας ή αρνητικότητας του αντιλαμβανόμενου συναισθήματος, η Energy σχετίζεται με το επίπεδο διέγερσης ή ενεργητικότητας, ενώ η Tension αφορά τον βαθμό έντασης ή συναισθηματικής φόρτισης. Στην παρούσα ΔΕ, οι τρεις αυτές μεταβλητές αντιμετωπίστηκαν ως συνεχείς στόχοι παλινδρόμησης.

Η δεύτερη ομάδα στόχων αφορά πέντε βασικές κατηγορίες συναισθήματος: Happy, Sad, Anger, Fear και Tender. Οι αξιολογήσεις των συγκεκριμένων μεταβλητών εκφράζονται επίσης σε συνεχή κλίμακα 1-9. Επομένως, και σε αυτήν την περίπτωση οι μεταβλητές αντιμετωπίστηκαν ως συνεχείς μεταβλητές παλινδρόμησης, οι οποίες αποτυπώνουν τον βαθμό στον οποίο κάθε μουσικό απόσπασμα εκφράζει το αντίστοιχο συναίσθημα.

Με βάση τα παραπάνω, διαμορφώθηκαν δύο κύριες ομάδες πειραμάτων πρόβλεψης. Η πρώτη ομάδα αφορούσε την πρόβλεψη των τριών συναισθηματικών διαστάσεων από αντικειμενικά ηχητικά χαρακτηριστικά. Η δεύτερη ομάδα αφορούσε την πρόβλεψη των πέντε βασικών συναισθημάτων από τα ίδια χαρακτηριστικά. Η παράλληλη μελέτη των δύο ομάδων επέτρεψε τόσο τη σύγκριση της απόδοσης διαφορετικών μοντέλων πρόβλεψης όσο και τη μεταγενέστερη χαρτογράφηση μεταξύ των δύο συναισθηματικών χώρων.

4.3 Διαδικασία Εξαγωγής και Επιλογής Αντικειμενικών Χαρακτηριστικών

Η διαδικασία εξαγωγής αντικειμενικών ηχητικών χαρακτηριστικών είχε ως στόχο τη δημιουργία ενός συνόλου μεταβλητών ικανού να αποτυπώσει διαφορετικές πτυχές του μουσικού σήματος, όπως η ένταση και η δυναμική, η ρυθμική οργάνωση, η τονικότητα, το φασματικό περιεχόμενο και η ηχοχρωματική δομή.

Η εξαγωγή των χαρακτηριστικών υλοποιήθηκε σε Python με χρήση εργαλείων της βιβλιοθήκης Essentia (<https://essentia.upf.edu/>), καθώς και συμπληρωματική επεξεργασία για τη διαμόρφωση πρόσθετων χαρακτηριστικών. Η επιλογή της συγκεκριμένης βιβλιοθήκης βασίστηκε στην ευρεία χρήση της στο πεδίο της MIR και στη δυνατότητα εξαγωγής ηχητικών χαρακτηριστικών διαφορετικής φύσης, τα οποία κρίθηκαν απαραίτητα για την παρούσα ΔΕ.

Στο αρχικό στάδιο της έρευνας διαμορφώθηκε ένα βασικό σύνολο 48 αντικειμενικών χαρακτηριστικών. Πιο συγκεκριμένα, συμπεριλήφθηκαν:

- Χαρακτηριστικά έντασης και δυναμικής, όπως τα `loudness_global`, `dynamic_complexity`, `loudness_ebu`, `rms_mean`, `rms_std` και `rms_slope`, τα οποία αφορούν την ενεργειακή στάθμη και τη δυναμική του σήματος.
- Ρυθμικά χαρακτηριστικά, όπως τα `tempo_bpm`, `beat_confidence` και `beats_count`, τα οποία περιγράφουν την ταχύτητα, τη ρυθμική οργάνωση και τη μεταβλητότητα του μουσικού αποσπάσματος.
- Τονικά / Αρμονικά χαρακτηριστικά, όπως τα `key_strength`, `majorness`, `pitch_mean`, `pitch_std` και `pitch_confidence_mean`, τα οποία σχετίζονται με τη σταθερότητα του τονικού κέντρου, το τονικό ύψος και την αξιοπιστία της εκτίμησής του.
- Ηχοχρωματικά / Φασματικά χαρακτηριστικά, όπως τα `spectral_centroid_mean`, `spectral_centroid_std`, `spectral_flux_mean`, `spectral_flux_std`, `spectral_entropy_mean`, `spectral_entropy_std`, `fluctuation_mean`, καθώς και οι συντελεστές Mel (Mel-Frequency Cepstral Coefficients, MFCCs) μέσω των μέσων τιμών και των τυπικών αποκλίσεων των `mfcc_1` έως `mfcc_13`. Τα χαρακτηριστικά αυτά αφορούν το φασματικό βάρος, τη φασματική μεταβολή, τη φασματική πολυπλοκότητα και την ηχοχρωματική δομή του ήχου

Παράλληλα με το βασικό αυτό σύνολο, διερευνήθηκε και ένα εναλλακτικό σύνολο 21 αντικειμενικών χαρακτηριστικών, το οποίο διαμορφώθηκε με βάση τη σχετική βιβλιογραφία στο πεδίο της MER για την πρόβλεψη συναισθηματικών αξιολογήσεων, και ειδικότερα την έρευνα των Eerola et al. (2009), καθώς και το MIRtoolbox 1.7 User's Manual (Lartillot, 2017). Το σύνολο αυτό περιλάμβανε επιλεγμένα χαρακτηριστικά από το αρχικό σύνολο, καθώς και πέντε επιπλέον χαρακτηριστικά: `hrcp_entropy_mean`, `pitch_salience_mean`, `spectral_spread_mean`, `spectral_flatness_mean` και `zcr_mean`.

Σε επόμενο στάδιο, το αρχικό σύνολο των 48 χαρακτηριστικών εμπλουτίστηκε με τα πέντε αυτά επιπλέον χαρακτηριστικά, με αποτέλεσμα να προκύψει το τελικό διευρυμένο σύνολο 53 αντικειμενικών χαρακτηριστικών. Το συγκεκριμένο σύνολο χρησιμοποιήθηκε ως το βασικό αρχικό σύνολο υποψήφιων μεταβλητών στα τελικά πειράματα της ΔΕ.

Ως προς τη διαδικασία υπολογισμού, η εξαγωγή των αντικειμενικών χαρακτηριστικών πραγματοποιήθηκε είτε σε επίπεδο ολόκληρου μουσικού αποσπάσματος είτε σε επίπεδο διαδοχικών χρονικών πλαισίων με διαφορετικές χρονικές κλίμακες ανάλυσης. Στο πρώτο επίπεδο, το ηχητικό σήμα τεμαχίστηκε σε μικρά επικαλυπτόμενα frames με χρήση των παραμέτρων `frame_size = 2048` και `hop_size = 1024`, πάνω στα οποία εφαρμόστηκαν αλγόριθμοι όπως RMS, Centroid, Entropy, MFCC, PitchYinFFT και ZeroCrossingRate. Οι frame-level τιμές συνοψίστηκαν σε επίπεδο αποσπάσματος μέσω στατιστικών μέτρων, κυρίως μέσου όρου και τυπικής απόκλισης. Στο δεύτερο επίπεδο, για την εκτίμηση χαρακτηριστικών που απαιτούν μεγαλύτερο χρονικό ορίζοντα (π.χ. tempo και pulse clarity), το σήμα χωρίστηκε σε μεγαλύτερα επικαλυπτόμενα τμήματα διάρκειας 3 δευτερολέπτων με βήμα 1,5 δευτερολέπτου.

Σημειώνεται ότι για κάθε σύνολο χαρακτηριστικών πραγματοποιήθηκαν διαδικασίες αξιολόγησης σημασίας και επιλογής χαρακτηριστικών, με σκοπό την επιλογή των βέλτιστων υποσυνόλων για την εκπαίδευση των μοντέλων πρόβλεψης. Περισσότερες πληροφορίες σχετικά με αυτή τη διαδικασία παρουσιάζονται στο Κεφάλαιο 5.

Στον Πίνακα 4.1 παρουσιάζονται όλα τα αντικειμενικά χαρακτηριστικά που χρησιμοποιήθηκαν στην παρούσα ΔΕ, η αντιστοίχισή τους με τους αντίστοιχους αλγόριθμους της βιβλιοθήκης Essentia ή/και με τις διαδικασίες υπολογισμού που εφαρμόστηκαν, καθώς και σύντομη περιγραφή για το καθένα.

Πίνακας 4.1: Αντικειμενικά χαρακτηριστικά, τρόπος υπολογισμού και περιγραφή

Χαρακτηριστικά έντασης και δυναμικής		
Χαρακτηριστικό(ά)	Αλγόριθμος / Τρόπος υπολογισμού	Σύντομη περιγραφή
loudness_global	Αλγόριθμος Loudness	Ψυχοακουστικό μέτρο της συνολικής αντιλαμβανόμενης έντασης βασισμένο στον νόμο του Stevens
dynamic_complexity	Αλγόριθμος DynamicComplexity	Βαθμός δυναμικής μεταβλητότητας του σήματος
loudness_ebu	Αλγόριθμος DynamicComplexity	Εκτίμηση της συνολικής έντασης του αποσπάσματος σε λογαριθμική κλίμακα dB
rms_mean , rms_std , rms_slope	Αλγόριθμος RMS σε διαδοχικά frames και υπολογισμός μέσου όρου, τυπικής απόκλισης και γραμμικής κλίσης	Μέση ενεργειακή στάθμη του σήματος, μεταβλητότητα της ενεργειακής στάθμης, και τάση αύξησης ή μείωσης της ενέργειας κατά τη διάρκεια του αποσπάσματος
Ρυθμικά χαρακτηριστικά		
Χαρακτηριστικό(ά)	Αλγόριθμος / Τρόπος υπολογισμού	Σύντομη περιγραφή
tempo_bpm , tempo_std , beat_confidence	Αλγόριθμος RhythmExtractor2013 σε ολόκληρο το απόσπασμα και σε διαδοχικά τμήματα για τον υπολογισμό της τυπικής απόκλισης του tempo	Εκτιμώμενος ρυθμός σε bpm, μεταβλητότητα του tempo, και βαθμός αξιοπιστίας της ανίχνευσης beat
beats_count	Αλγόριθμος RhythmExtractor2013 και υπολογισμός του πλήθους των ανιχνευμένων beats	Συνολικό πλήθος beats που ανιχνεύθηκαν στο απόσπασμα
pulse_clarity , pulse_clarity_std	Αλγόριθμος PulseClarity σε ολόκληρο το απόσπασμα και σε διαδοχικά τμήματα για τον υπολογισμό τυπικής απόκλισης	Δείκτης σαφήνειας του ρυθμικού παλμού και μεταβλητότητάς του στο χρόνο
Τονικά / Αρμονικά χαρακτηριστικά		
Χαρακτηριστικό(ά)	Αλγόριθμος / Τρόπος υπολογισμού	Σύντομη περιγραφή
key_strength	Αλγόριθμος KeyExtractor	Δείκτης εκτίμησης της ισχύος της τονικότητας
majorness	Αλγόριθμος KeyExtractor και μετατροπή της μεταβλητής scale σε δυαδικό δείκτη (major=1, minor=0)	Δείκτης μείζονος ή ελάσσονος χαρακτήρα του αποσπάσματος
pitch_mean , pitch_std	Αλγόριθμος PitchYinFFT σε frames και υπολογισμός μέσου όρου και τυπικής απόκλισης	Μέση τιμή και μεταβλητότητα της εκτιμώμενης θεμελιώδους συχνότητας

pitch_confidence_mean	Αλγόριθμος PitchYinFFT και υπολογισμός μέσου όρου της μεταβλητής pitchConfidence	Μέση αξιοπιστία της εκτίμησης του pitch
hpcp_entropy_mean	Αλγόριθμοι SpectralPeaks και HPCP σε frames και υπολογισμός της εντροπίας του HPCP	Δείκτης κατανομής της τονικής ενέργειας στις pitch classes
pitch_salience_mean	Αλγόριθμος PitchSalience σε frames και υπολογισμός μέσου όρου	Εκτίμηση καθαρότητας προβολής του pitch
Ηχοχρωματικά / Φασματικά χαρακτηριστικά		
Χαρακτηριστικό(ά)	Αλγόριθμος / Τρόπος υπολογισμού	Σύντομη περιγραφή
spectral_centroid_mean , spectral_centroid_std	Αλγόριθμος Centroid σε frames και υπολογισμός μέσου όρου και τυπικής απόκλισης	Μέση θέση και μεταβλητότητα του κέντρου βάρους του φάσματος
spectral_flux_mean , spectral_flux_std	Υπολογισμός της φασματικής μεταβολής μεταξύ διαδοχικών frame-level φασμάτων και υπολογισμός μέσου όρου και τυπικής απόκλισης	Μέση μεταβολή του φασματικού περιεχομένου και μεταβλητότητά της
spectral_entropy_mean , spectral_entropy_std	Αλγόριθμος Entropy σε κανονικοποιημένα frame-level φάσματα και υπολογισμός μέσου όρου και τυπικής απόκλισης	Δείκτης φασματικής πολυπλοκότητας και μεταβλητότητάς της
spectral_spread_mean	Αλγόριθμοι CentralMoments και DistributionShape σε frames και υπολογισμός μέσου όρου	Δείκτης διασποράς του φάσματος γύρω από το spectral centroid
spectral_flatness_mean	Αλγόριθμος Flatness σε frames και υπολογισμός μέσου όρου	Δείκτης βαθμού τονικού ή θορυβώδους χαρακτήρα φάσματος
fluctuation_mean	Υπολογισμός της μέσης απόλυτης διαφοράς μεταξύ διαδοχικών frame-level φασμάτων	Δείκτης χρονικής φασματικής μεταβολής
zcr_mean	Αλγόριθμος ZeroCrossingRate σε frames και υπολογισμός μέσου όρου	Μέση συχνότητα αλλαγής πρόσημου του σήματος
mfcc_1_mean – mfcc_13_mean , mfcc_1_std – mfcc_13_std	Αλγόριθμος MFCC σε frames και υπολογισμός μέσου όρου και τυπικής απόκλισης για κάθε συντελεστή	Δείκτες της ηχοχρωματικής δομής και μεταβλητότητάς της

4.4 Προεπεξεργασία και Τελική Διαμόρφωση των Δεδομένων

Μετά την εξαγωγή των αντικειμενικών ηχητικών χαρακτηριστικών, ακολούθησε η διαδικασία προεπεξεργασίας και οργάνωσης των δεδομένων, η οποία περιλάμβανε τη συγχώνευση των εξαγόμενων χαρακτηριστικών με τις διαθέσιμες συναισθηματικές αξιολογήσεις και τα βασικά μεταδεδομένα των μουσικών αποσπασμάτων.

Σε αυτό το πλαίσιο, διαμορφώθηκε ένα ενοποιημένο σύνολο δεδομένων σε μορφή πίνακα, στο οποίο κάθε γραμμή αντιστοιχούσε σε ένα μουσικό απόσπασμα και κάθε στήλη σε αντικειμενικό χαρακτηριστικό, μεταδεδομένο ή μεταβλητή-στόχο. Τα μεταδεδομένα διατηρήθηκαν για λόγους αντιστοίχισης των αποσπασμάτων και δε χρησιμοποιήθηκαν στα μοντέλα.

Κατά τη διαδικασία ελέγχου του τελικού πίνακα διαπιστώθηκε ότι ορισμένα χαρακτηριστικά εμφάνιζαν ελλειπείς τιμές. Συγκεκριμένα, οι μεταβλητές `pulse_clarity` και `pulse_clarity_std` ήταν πλήρως κενές και δε χρησιμοποιήθηκαν στα τελικά μοντέλα. Γενικότερα, πριν από την εκπαίδευση των μοντέλων αφαιρούνταν τα χαρακτηριστικά που περιείχαν πλήρως ελλειπείς τιμές, ενώ λοιπές ελλειπείς τιμές αντιμετωπίζονταν στο στάδιο της μοντελοποίησης μέσω κατάλληλων διαδικασιών συμπλήρωσης, όπως περιγράφεται αναλυτικότερα στο Κεφάλαιο 5.

Με βάση τη συνολική διαδικασία διερεύνησης της καταλληλότητας των χαρακτηριστικών για την πρόβλεψη των συναισθηματικών αξιολογήσεων, προέκυψαν τρία σύνολα χαρακτηριστικών τα οποία χρησιμοποιήθηκαν στα πειράματα. Πρώτον, το αρχικό σύνολο των 48 χαρακτηριστικών. Δεύτερον, το εναλλακτικό σύνολο 21 χαρακτηριστικών, το οποίο χρησιμοποιήθηκε για συγκριτικά πειράματα. Τρίτον, το τελικό διευρυμένο σύνολο 53 χαρακτηριστικών, το οποίο προέκυψε από το αρχικό σύνολο με την προσθήκη πέντε επιπλέον χαρακτηριστικών. Το τελευταίο σύνολο αποτέλεσε τη βασική δεξαμενή υποψήφιων μεταβλητών για τα τελικά πειράματα της ΔΕ.

5. Υλοποίηση Μοντέλων και Πειραματική Διαδικασία

Στο παρόν Κεφάλαιο παρουσιάζεται η μεθοδολογική προσέγγιση που ακολουθήθηκε για την πρόβλεψη συναισθηματικών αξιολογήσεων από αντικειμενικά χαρακτηριστικά μουσικών αποσπασμάτων. Ειδικότερα, περιγράφονται η πειραματική διαδικασία, τα μοντέλα πρόβλεψη που αναπτύχθηκαν και τα μέτρα αξιολόγησης που χρησιμοποιήθηκαν, η διαδικασία αξιολόγησης σημασίας και επιλογής χαρακτηριστικών, καθώς και οι διαδικασίες χαρτογράφησης συναισθηματικού προφίλ.

5.1 Συνοπτική Περιγραφή της Πειραματικής Διαδικασίας

Η πειραματική διαδικασία της ΔΕ σχεδιάστηκε με στόχο τη διερεύνηση της δυνατότητας πρόβλεψης συναισθηματικών αξιολογήσεων από αντικειμενικά ηχητικά χαρακτηριστικά μουσικών αποσπασμάτων. Για τον σκοπό αυτό, διαμορφώθηκαν εναλλακτικά σύνολα χαρακτηριστικών, εκπαιδεύτηκαν και αξιολογήθηκαν διαφορετικά μοντέλα παλινδρόμησης, εφαρμόστηκαν διαδικασίες αξιολόγησης σημασίας και επιλογής χαρακτηριστικών, και, τέλος, πραγματοποιήθηκε σύγκριση άμεσων και έμμεσων προσεγγίσεων πρόβλεψης συναισθηματικών χώρων.

Σε πρώτη φάση, εξετάστηκαν μοντέλα πρόβλεψης με χρήση του βασικού συνόλου 48 χαρακτηριστικών, καθώς και των εναλλακτικών συνόλων των 21 και των 53 χαρακτηριστικών, τα οποία περιγράφονται στο Κεφάλαιο 4. Στο πλαίσιο αυτό μελετήθηκαν τόσο διαστασιακοί στόχοι (Valence, Energy, Tension) όσο και βασικά συναισθήματα (Happy, Sad, Anger, Fear, Tender) με σκοπό την παράλληλη αξιολόγηση των δύο μορφών συναισθηματικής αναπαράστασης.

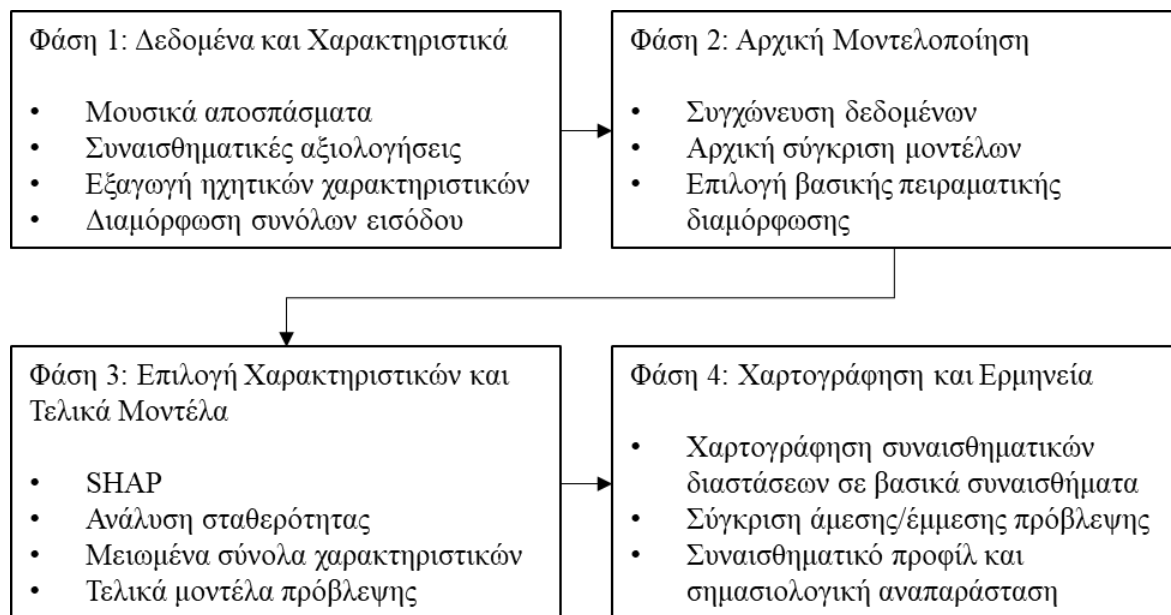
Στη συνέχεια εφαρμόστηκαν διαδικασίες αξιολόγησης σημασίας και επιλογής χαρακτηριστικών, με σκοπό τη μείωση του αριθμού των χαρακτηριστικών και τη διαμόρφωση αποδοτικότερων υποσυνόλων εισόδου για τα μοντέλα. Η διαδικασία αυτή πραγματοποιήθηκε τόσο σε κοινό επίπεδο για όλους τους στόχους όσο και σε επίπεδο επιμέρους στόχου, ώστε να εξεταστεί εάν η βέλτιστη επιλογή χαρακτηριστικών διαφοροποιείται ανάλογα με τη συναισθηματική μεταβλητή που προβλέπεται και εάν μπορεί να οδηγήσει σε βελτιωμένη απόδοση των μοντέλων πρόβλεψης.

Τα τελικά μοντέλα αξιολογήθηκαν μέσω διασταυρούμενης επικύρωσης (cross-validation), με χρήση κατάλληλων μετρικών παλινδρόμησης, ενώ πραγματοποιήθηκαν και στατιστικοί

έλεγχοι για τη σύγκριση διαφορετικών επιλογών μοντελοποίησης. Παράλληλα, χρησιμοποιήθηκαν τεχνικές ερμηνευσιμότητας για την αξιολόγηση της συμβολής των επιμέρους χαρακτηριστικών στις τελικές προβλέψεις.

Τέλος, εξετάστηκε η χαρτογράφηση μεταξύ συναισθηματικών χώρων και ειδικότερα η μετατροπή προβλεπόμενων συναισθηματικών διαστάσεων σε βασικά συναισθήματα. Με τον τρόπο αυτό κατέστη δυνατή η σύγκριση μεταξύ ενός άμεσου μοντέλου πρόβλεψης συναισθηματικού προφίλ και μιας έμμεσης προσέγγισης σημασιολογικής χαρτογράφησης.

Η συνολική ροή της πειραματικής διαδικασίας, από τη διαμόρφωση του συνόλου δεδομένων και την εξαγωγή των αντικειμενικών ηχητικών χαρακτηριστικών έως την εκπαίδευση των τελικών μοντέλων, τη χαρτογράφηση μεταξύ συναισθηματικών χώρων και τη σημασιολογική αναπαράσταση των αποτελεσμάτων, παρουσιάζεται συνοπτικά στην Εικόνα 5.1.



Εικόνα 5.1: Διάγραμμα ροής της πειραματικής διαδικασίας

5.2 Επιλογή Μοντέλων Πρόβλεψης και Παράμετροι

Για την πρόβλεψη των συναισθηματικών αξιολογήσεων από τα αντικειμενικά ηχητικά χαρακτηριστικά χρησιμοποιήθηκαν μοντέλα παλινδρόμησης βασισμένα σε δέντρα αποφάσεων και συναθροιστικές μεθόδους μάθησης (ensemble learning) (Breiman, 2001; Friedman, 2001; Friedman, 2002; Hastie et al., 2009). Ειδικότερα, στα πειράματα της ΔΕ αξιοποιήθηκαν οι αλγόριθμοι RandomForestRegressor και GradientBoostingRegressor της βιβλιοθήκης scikit-learn, καθώς κρίθηκαν κατάλληλοι για τη μοντελοποίηση πιθανών μη

γραμμικών σχέσεων μεταξύ των ηχητικών χαρακτηριστικών και των συναισθηματικών μεταβλητών-στόχων (Scikit-learn developers, n.d.a, n.d.b, n.d.c; Pedregosa et al., 2011).

Η επιλογή των συγκεκριμένων μοντέλων βασίστηκε στο γεγονός ότι μπορούν να διαχειριστούν αποτελεσματικά ετερογενή σύνολα χαρακτηριστικών, όπως αυτά που αξιοποιήθηκαν στη ΔΕ, χωρίς να απαιτούν αυστηρές υποθέσεις για τη μορφή της σχέσης εισόδου-εξόδου. Επιπλέον, μπορούν να εφαρμοστούν αποτελεσματικά σε σύνολα δεδομένων περιορισμένου μεγέθους καθώς και να συνδυαστούν με τεχνικές ερμηνευσιμότητας, όπως το SHAP, για την αποτίμηση της συμβολής των επιμέρους χαρακτηριστικών στις προβλέψεις (Hastie et al., 2009).

Στα πειράματα χρησιμοποιήθηκαν τόσο προσεγγίσεις πολλαπλής εξόδου (multi-output regression) όσο και προσεγγίσεις μονοδιάστατης πρόβλεψης ανά στόχο (single-output regression). Οι προσεγγίσεις πολλαπλής εξόδου αξιοποιήθηκαν στα πειράματα όπου προβλέπονταν ταυτόχρονα όλες οι συναισθηματικές διαστάσεις ή όλα τα βασικά συναισθήματα με χρήση ενός ενιαίου μοντέλου. Αντίθετα, στα πειράματα ανά στόχο, όπου εξετάστηκε η καταλληλότητα διαφορετικών υποσυνόλων χαρακτηριστικών για κάθε επιμέρους συναισθηματική μεταβλητή, χρησιμοποιήθηκαν ξεχωριστά μονοδιάστατα μοντέλα παλινδρόμησης. Για την υλοποίηση των multi-output πειραμάτων χρησιμοποιήθηκε ο αλγόριθμος MultiOutputRegressor της βιβλιοθήκης scikit-learn, ο οποίος επιτρέπει την ανεξάρτητη εκπαίδευση ενός παλινδρομητή για κάθε μεταβλητή-στόχο μέσα σε ενιαίο πειραματικό πλαίσιο (Scikit-learn developers, n.d.d).

Οι βασικές υπερπαράμετροι (hyperparameters) των μοντέλων διατηρήθηκαν σταθερές στα πειράματα, ώστε να είναι δυνατή η σύγκρισή τους και οι διαφορές στην απόδοσή τους να αποδίδονται κυρίως στη σύνθεση των χαρακτηριστικών. Για τον αλγόριθμο RandomForestRegressor χρησιμοποιήθηκαν $n_estimators = 300$, $min_samples_split = 4$, $min_samples_leaf = 2$ και $random_state = 42$. Ο αριθμός των δέντρων ($n_estimators$) καθορίζει πόσα επιμέρους δέντρα συμμετέχουν στην τελική πρόβλεψη, ενώ οι παράμετροι $min_samples_split$ και $min_samples_leaf$ αφορούν τη δομή των δέντρων και συμβάλλουν στον περιορισμό της υπερπροσαρμογής (overfitting). Η παράμετρος $random_state$ χρησιμοποιήθηκε για λόγους διασφάλισης της αναπαραγωγιμότητας των αποτελεσμάτων (scikit-learn developers, n.d.b).

Για τον αλγόριθμο GradientBoostingRegressor χρησιμοποιήθηκαν $n_estimators = 200$, $learning_rate = 0.05$, $max_depth = 3$, $subsample = 0.9$ και $random_state = 42$. Η παράμετρος $n_estimators$ καθορίζει τον αριθμό των διαδοχικών δέντρων που κατασκευάζονται, η $learning_rate$ ελέγχει τη συμβολή κάθε νέου δέντρου στη συνολική πρόβλεψη, ενώ η max_depth αφορά την πολυπλοκότητα κάθε επιμέρους δέντρου. Η χρήση $subsample = 0.9$ επιτρέπει την εκπαίδευση κάθε σταδίου σε τυχαίο υποσύνολο των δεδομένων, συμβάλλοντας στη σταθερότητα του μοντέλου και στη μείωση του κινδύνου υπερπροσαρμογής (scikit-learn developers, n.d.c).

Παράλληλα, χρησιμοποιήθηκε και ένα απλό μοντέλο βάσης (DummyRegressor), το οποίο προβλέπει για κάθε στόχο τη μέση τιμή του εκάστοτε συνόλου εκπαίδευσης (training set). Η χρήση του επέτρεψε την εκτίμηση του κατά πόσο τα κύρια μοντέλα αξιοποιούν ουσιαστικά πληροφορία από τα ηχητικά χαρακτηριστικά (scikit-learn developers, n.d.e).

5.3 Μέτρα Αξιολόγησης

Η αξιολόγηση των μοντέλων πραγματοποιήθηκε με διαφορετικά μέτρα, ανάλογα με τη φύση του εκάστοτε πειράματος. Τα κύρια πειράματα πρόβλεψης τόσο συναισθηματικών διαστάσεων όσο και βασικών συναισθημάτων αξιολογήθηκαν ως προβλήματα παλινδρόμησης, δεδομένου ότι οι μεταβλητές-στόχοι ήταν συνεχείς. Στα πειράματα χαρτογράφησης από συναισθηματικές διαστάσεις σε βασικά συναισθήματα χρησιμοποιήθηκαν επιπλέον μέτρα, προκειμένου να αξιολογηθεί τόσο η ακρίβεια του πλήρους συναισθηματικού προφίλ όσο και η ικανότητα ανάκτησης του κυρίαρχου συναισθήματος.

Για την αξιολόγηση των μοντέλων παλινδρόμησης χρησιμοποιήθηκαν κυρίως οι μετρικές R^2 και Μέσο Απόλυτο Σφάλμα (Mean Absolute Error, MAE) (Hastie et al., 2009; scikit-learn developers, n.d.f). Ο συντελεστής προσδιορισμού R^2 εκτιμά το ποσοστό της διακύμανσης της μεταβλητής-στόχου που εξηγείται από το μοντέλο. Τιμές του R^2 κοντά στο 1 υποδηλώνουν καλύτερη προσαρμογή, ενώ χαμηλές ή αρνητικές τιμές υποδηλώνουν χαμηλή προβλεπτική ικανότητα. Το MAE εκφράζει το μέσο απόλυτο σφάλμα μεταξύ πραγματικών και προβλεπόμενων τιμών και παρέχει ένα άμεσα ερμηνεύσιμο μέτρο σφάλματος στην ίδια κλίμακα με τις μεταβλητές-στόχους. Οι δύο αυτές μετρικές παρέχουν συμπληρωματική πληροφόρηση και επιτρέπουν την εκτίμηση της απόδοσης των μοντέλων

τόσο σε επίπεδο συνολικής εξηγούμενης διακύμανσης όσο και σε επίπεδο πραγματικής απόκλισης από τις ανθρώπινες αξιολογήσεις.

Στα πειράματα που αφορούσαν τη χαρτογράφηση από συναισθηματικές διαστάσεις σε βασικά συναισθήματα και τη σύγκρισή της με το άμεσο μοντέλο πρόβλεψης συναισθηματικού προφίλ, χρησιμοποιήθηκε επιπλέον το profile MAE, δηλαδή το μέσο απόλυτο σφάλμα πάνω στο σύνολο των τιμών του συναισθηματικού προφίλ (Happy, Sad, Anger, Fear, Tender). Για κάθε μουσικό απόσπασμα, οι πραγματικές και οι προβλεπόμενες τιμές κανονικοποιήθηκαν ώστε να προκύψουν προφίλ με άθροισμα ίσο με 1, και στη συνέχεια υπολογίστηκε το μέσο απόλυτο σφάλμα μεταξύ των αντίστοιχων προφίλ. Επιπλέον, για την αξιολόγηση της ομοιότητας μεταξύ πραγματικών και προβλεπόμενων συναισθηματικών προφίλ χρησιμοποιήθηκε η cosine similarity (Manning et al., 2008). Η μετρική αυτή αποτιμά το βαθμό ομοιότητας της σχετικής μορφής ή κατανομής δύο προφίλ, ανεξάρτητα από τις απόλυτες τιμές τους.

Τέλος, στα πειράματα χαρτογράφησης εξετάστηκε επίσης η απόδοση ως προς το κυρίαρχο συναίσθημα (dominant emotion). Στο πλαίσιο αυτό, από κάθε πραγματικό και προβλεπόμενο συναισθηματικό προφίλ επιλέχθηκε το συναίσθημα με τη μέγιστη τιμή και η αξιολόγηση αντιμετωπίστηκε ως πρόβλημα ταξινόμησης. Για τον σκοπό αυτό χρησιμοποιήθηκαν οι μετρικές accuracy, precision, recall και F1-score, καθώς και πίνακες σύγχυσης (confusion matrices) (Sokolova & Lapalme, 2009).

5.4 SHAP και Ανάλυση Σταθερότητας

Για την εκτίμηση της συνεισφοράς κάθε χαρακτηριστικού στην απόδοση των μοντέλων χρησιμοποιήθηκε η μέθοδος SHAP (Shapley Additive exPlanations) (Lundberg & Lee, 2017). Η προσέγγιση αυτή επιτρέπει την ποσοτικοποίηση της συμβολής κάθε χαρακτηριστικού στην τελική πρόβλεψη ενός μοντέλου με βάση τις τιμές Shapley από τη θεωρία παιγνίων (Shapley, 1953). Η SHAP χρησιμοποιήθηκε για την εξαγωγή τόσο συνολικών δεικτών σημασίας χαρακτηριστικών όσο και γραφικών παραστάσεων που αποτυπώνουν τη σχέση μεταξύ των τιμών των χαρακτηριστικών και της κατεύθυνσης επίδρασής τους στις προβλέψεις.

Η ανάλυση πραγματοποιήθηκε ανά μεταβλητή-στόχο, με χρήση μοντέλων GradientBoostingRegressor, καθώς τα συγκεκριμένα μοντέλα αποτέλεσαν τον βασικό κορμό των τελικών πειραμάτων. Για κάθε στόχο, μετά την εκπαίδευση του αντίστοιχου

μοντέλου, υπολογίστηκαν οι τιμές SHAP για όλα τα χαρακτηριστικά και εξήχθη η μέση απόλυτη τιμή SHAP (mean absolute SHAP value) ανά χαρακτηριστικό. Το μέτρο αυτό αποτελεί ένα δείκτη της συνολικής σημασίας κάθε χαρακτηριστικού. Επίσης, δημιουργήθηκαν γραφήματα τύπου bar plot και beeswarm plot. Τα bar plots χρησιμοποιήθηκαν για τη συνολική κατάταξη των χαρακτηριστικών ως προς τη μέση απόλυτη συμβολή τους, ενώ τα beeswarm plots χρησιμοποιήθηκαν για την οπτικοποίηση της κατανομής των τιμών SHAP κάθε χαρακτηριστικού, με στόχο να αποτυπωθεί ο τρόπος με τον οποίο οι τιμές του επηρεάζουν τις προβλέψεις του μοντέλου (Molnar, 2025).

Παράλληλα, εφαρμόστηκε ανάλυση σταθερότητας (stability analysis) με χρήση 5-fold cross-validation (Hastie et al., 2009). Σε κάθε fold εκπαιδευόταν εκ νέου ένα μοντέλο GradientBoostingRegressor και υπολογίζονταν οι τιμές SHAP για το αντίστοιχο fold. Για κάθε μεταβλητή-στόχο καταγραφόταν η κατάταξη των χαρακτηριστικών με βάση τη μέση απόλυτη τιμή SHAP και διατηρούνταν τα 10 σημαντικότερα χαρακτηριστικά κάθε fold. Η σταθερότητα των χαρακτηριστικών αποτιμήθηκε κυρίως μέσω της συχνότητας εμφάνισης ενός χαρακτηριστικού στα 10 σημαντικότερα χαρακτηριστικών των διαφορετικών folds (count_top10), καθώς και μέσω της μέσης θέσης κατάταξής του (mean rank).

Τα χαρακτηριστικά που παρουσίαζαν ταυτόχρονα υψηλή τιμή SHAP και σταθερή παρουσία στις πρώτες θέσεις σε κάθε fold κρίθηκαν περισσότερο αξιόπιστα για τη διαμόρφωση μειωμένων συνόλων χαρακτηριστικών. Η προσέγγιση αυτή εφαρμόστηκε στα σύνολα χαρακτηριστικών που περιγράφονται στο Κεφάλαιο 4.

5.5 Κοινή και ανά Στόχο Επιλογή Αντικειμενικών Χαρακτηριστικών

Με βάση τα αποτελέσματα της ανάλυσης SHAP και της ανάλυσης σταθερότητας εφαρμόστηκαν δύο εναλλακτικές στρατηγικές επιλογής χαρακτηριστικών: η κοινή επιλογή χαρακτηριστικών για όλους τους στόχους και η επιλογή χαρακτηριστικών ανά στόχο. Στόχος της διαδικασίας ήταν να εξεταστεί αν τα μοντέλα αποδίδουν καλύτερα με χρήση ενός ενιαίου συνόλου χαρακτηριστικών εισόδου για όλες τις μεταβλητές-στόχους ή αν απαιτείται διαφορετικό υποσύνολο χαρακτηριστικών για κάθε στόχο.

Στην κοινή επιλογή χαρακτηριστικών, για κάθε στόχο προσδιορίστηκαν αρχικά τα 10 σημαντικότερα χαρακτηριστικά με βάση τη μέση απόλυτη τιμή SHAP, καθώς και τα χαρακτηριστικά που παρουσίαζαν επαρκή σταθερότητα μέσω της ανάλυσης σταθερότητας με 5-fold cross-validation. Σταθερά χαρακτηριστικά θεωρήθηκαν αυτά που εμφανίζονταν

με συχνότητα τουλάχιστον ίση με ένα καθορισμένο κατώφλι σταθερότητας (threshold) (π.χ. $N \geq 3$ ή $N \geq 4$) στα 10 σημαντικότερα χαρακτηριστικά των διαφορετικών folds. Για κάθε στόχο κρατήθηκε η τομή των δύο αυτών συνόλων, δηλαδή των χαρακτηριστικών που ήταν ταυτόχρονα σημαντικά και σταθερά. Το τελικό κοινό μειωμένο σύνολο προέκυψε από την ένωση των επιλεγμένων χαρακτηριστικών όλων των στόχων και χρησιμοποιήθηκε εκ νέου σε μοντέλα GradientBoostingRegressor πολλαπλής εξόδου.

Όσον αφορά την επιλογή χαρακτηριστικών ανά στόχο, εφαρμόστηκε η ίδια λογική ξεχωριστά για κάθε μεταβλητή-στόχο, χωρίς ενοποίηση των επιμέρους συνόλων. Για κάθε στόχο διατηρήθηκαν μόνο τα χαρακτηριστικά που ανήκαν ταυτόχρονα στα σημαντικότερα βάσει SHAP και στα επαρκώς σταθερά χαρακτηριστικά της αντίστοιχης ανάλυσης σταθερότητας. Τα υποσύνολα χαρακτηριστικών που προέκυψαν χρησιμοποιήθηκαν σε μονοδιάστατα μοντέλα GradientBoostingRegressor ανά στόχο.

Τέλος, πραγματοποιήθηκε συγκριτική αξιολόγηση των μοντέλων, με στόχο την επικύρωση των αποδοτικότερων στρατηγικών επιλογής χαρακτηριστικών για την πρόβλεψη συναισθηματικών διαστάσεων και βασικών συναισθημάτων.

5.6 Τελικά Μοντέλα Πρόβλεψης για Διαστάσεις και Βασικά Συναισθήματα

Μετά την ολοκλήρωση της διερευνητικής φάσης και της διαδικασίας επιλογής χαρακτηριστικών, η τελική φάση μοντελοποίησης της ΔΕ επικεντρώθηκε σε μοντέλα βασισμένα στο διευρυμένο σύνολο των 53 αντικειμενικών χαρακτηριστικών. Η επιλογή αυτή κρίθηκε κατάλληλη, καθώς το συγκεκριμένο σύνολο αποτελούσε το πληρέστερο σύνολο περιγραφής των μουσικών αποσπασμάτων και τα αντίστοιχα μοντέλα παρουσίαζαν βελτιωμένη απόδοση.

Ως βασικός αλγόριθμος στα τελικά πειράματα χρησιμοποιήθηκε ο GradientBoostingRegressor, διότι στη διερευνητική φάση αποδείχθηκε η πιο αποδοτική επιλογή για την πρόβλεψη των συναισθηματικών μεταβλητών-στόχων σε όλες τις περιπτώσεις. Για τα τελικά μοντέλα χρησιμοποιήθηκε κοινό πλαίσιο αξιολόγησης και κοινή λογική χειρισμού των ελλিপών τιμών, με στόχο οι διαφορές στην απόδοση να αποδίδονται κυρίως στη σύνθεση των χαρακτηριστικών. Με αυτό το σκεπτικό, δεν πραγματοποιήθηκε εκτεταμένη διαδικασία υπερπαραμετρικής βελτιστοποίησης των μοντέλων η οποία ενδεχομένως να οδηγούσε σε βελτιωμένη απόδοση.

Για την πρόβλεψη των συναισθηματικών διαστάσεων (Valence, Energy, Tension) εξετάστηκαν τρεις βασικές διαμορφώσεις. Η πρώτη αντιστοιχούσε στο πλήρες διευρυμένο σύνολο των 53 χαρακτηριστικών (FULL_EXTENDED_53F). Η δεύτερη αντιστοιχούσε σε ένα κοινό μειωμένο σύνολο χαρακτηριστικών, το οποίο προέκυψε από τη διαδικασία SHAP και ανάλυσης σταθερότητας, και εφαρμόστηκε ενιαία σε όλους τους στόχους (REDUCED_EXTENDED_COMMON_N $\geq$ 4). Η τρίτη αντιστοιχούσε σε επιμέρους μειωμένα σύνολα χαρακτηριστικών ανά στόχο (REDUCED_EXTENDED_PER_TARGET_N $\geq$ 3), επιτρέποντας τη χρήση διαφορετικού υποσυνόλου εισόδου για κάθε συναισθηματική διάσταση.

Για την πρόβλεψη των βασικών συναισθημάτων (Happy, Sad, Anger, Fear, Tender) εξετάστηκαν αντίστοιχα τρεις διαμορφώσεις: το πλήρες διευρυμένο σύνολο των 53 χαρακτηριστικών, το κοινό μειωμένο σύνολο χαρακτηριστικών (REDUCED_EXTENDED_COMMON_N $\geq$ 4) και τα εξειδικευμένα μειωμένα υποσύνολα χαρακτηριστικών ανά στόχο (REDUCED_EXTENDED_PER_TARGET_N $\geq$ 3).

Με τον τρόπο αυτόν, προέκυψαν τα αποδοτικότερα μοντέλα πρόβλεψης συναισθηματικών διαστάσεων και βασικών συναισθημάτων, για τα οποία πραγματοποιήθηκε στατιστικός έλεγχος για την εκτίμηση της σημαντικότητας της βελτίωσής τους έπειτα από τη μείωση των χαρακτηριστικών/εισόδων, ενώ χρησιμοποιήθηκαν επίσης στη διαδικασία χαρτογράφησης συναισθηματικών διαστάσεων.

5.7 Στατιστικός Έλεγχος Βελτίωσης των Τελικών Μοντέλων

Για την εξέταση της σημαντικότητας της διαφοράς απόδοσης των τελικών μοντέλων πραγματοποιήθηκε στατιστικός έλεγχος σημαντικότητας. Συγκεκριμένα, εφαρμόστηκε το Wilcoxon signed-rank test, το οποίο είναι κατάλληλο για σύγκριση δύο συζευγμένων συνόλων παρατηρήσεων χωρίς την απαίτηση κανονικότητας (Wilcoxon, 1945). Στην παρούσα ΔΕ, τα δύο συζευγμένα σύνολα αντιστοιχούσαν στα απόλυτα σφάλματα πρόβλεψης που παρήγαγαν δύο διαφορετικά μοντέλα για τα ίδια μουσικά αποσπάσματα.

Για κάθε σύγκριση μοντέλων υπολογίστηκαν αρχικά out-of-fold προβλέψεις μέσω 5-fold cross-validation. Στη συνέχεια, για κάθε μεταβλητή-στόχο υπολογίστηκε το απόλυτο σφάλμα πρόβλεψης ανά δείγμα, και εφαρμόστηκε ο έλεγχος Wilcoxon μεταξύ των δύο αντίστοιχων συνόλων σφαλμάτων. Παράλληλα, υπολογίστηκε και ένα συνολικό μέτρο

σφάλματος ανά δείγμα, ως μέσος όρος των απόλυτων σφαλμάτων όλων των στόχων για κάθε δείγμα, με στόχο την εξέταση της συνολικής διαφοροποίησης των μοντέλων.

Ο στατιστικός έλεγχος εφαρμόστηκε στα τελικά μοντέλα της ΔΕ. Για τις συναισθηματικές διαστάσεις εξετάστηκε η σύγκριση μεταξύ του πλήρους διευρυμένου μοντέλου FULL_EXTENDED_53F και του κοινού μειωμένου μοντέλου REDUCED_EXTENDED_COMMON_N $\geq$ 4, καθώς και η σύγκριση μεταξύ του κοινού μειωμένου μοντέλου και του ανά στόχο μειωμένου μοντέλου REDUCED_EXTENDED_PER_TARGET_N $\geq$ 3. Για τα βασικά συναισθήματα εξετάστηκε η σύγκριση μεταξύ του FULL_EXTENDED_53F και του REDUCED_EXTENDED_COMMON_N $\geq$ 4.

Για κάθε σύγκριση καταγράφηκαν η τιμή του στατιστικού ελέγχου (W) και η αντίστοιχη τιμή p (p-value). Μια διαφορά θεωρήθηκε στατιστικά σημαντική όταν η τιμή p ήταν μικρότερη από 0.05 ($p < 0.05$).

5.8 Χαρτογράφηση Συναισθηματικών Διαστάσεων σε Βασικά Συναισθήματα

Στην παρούσα ΔΕ εξετάστηκε επίσης η δυνατότητα χαρτογράφησης μεταξύ των δύο μορφών συναισθηματικής αναπαράστασης. Πιο συγκεκριμένα, διερευνήθηκε κατά πόσο οι συναισθηματικές διαστάσεις Valence, Energy και Tension μπορούν να αξιοποιηθούν για την πρόβλεψη του προφίλ βασικών συναισθημάτων (Happy, Sad, Anger, Fear, Tender).

Αρχικά, πραγματοποιήθηκε ένα πείραμα χαρτογράφησης από τις πραγματικές συναισθηματικές διαστάσεις προς τα πραγματικά βασικά συναισθήματα (true dimensions $\rightarrow$ true emotions). Ως είσοδοι χρησιμοποιήθηκαν οι μεταβλητές Valence, Energy και Tension, ενώ ως έξοδοι οι πέντε μεταβλητές βασικών συναισθημάτων. Το πείραμα αντιμετωπίστηκε ως πρόβλημα παλινδρόμησης πολλαπλής εξόδου και αξιολογήθηκε με τα ίδια μέτρα παλινδρόμησης που χρησιμοποιήθηκαν και στα υπόλοιπα μοντέλα της ΔΕ. Στόχος του συγκεκριμένου πειράματος ήταν να εκτιμηθεί σε ποιο βαθμό ο διαστασιακός συναισθηματικός χώρος παρέχει επαρκή πληροφορία για την πρόβλεψη του χώρου των βασικών συναισθημάτων, ακολουθώντας παρόμοια λογική με εκείνη των Eerola et al. (2009).

Στη συνέχεια, εξετάστηκε η δυνατότητα χαρτογράφησης από προβλεπόμενες συναισθηματικές διαστάσεις σε βασικά συναισθήματα (predicted dimensions $\rightarrow$ mapped

emotions). Στην περίπτωση αυτή, χρησιμοποιήθηκαν οι προβλέψεις των τελικών μοντέλων για τις τρεις συναισθηματικές διαστάσεις και επιχειρήθηκε η μετατροπή τους σε πλήρες συναισθηματικό προφίλ βασικών συναισθημάτων. Συγκεκριμένα, οι τρεις διαστάσεις προβλέπονταν με ξεχωριστά μοντέλα GradientBoostingRegressor και χρήση 5-fold cross-validation, χρησιμοποιώντας τα ανά στόχο μειωμένα υποσύνολα χαρακτηριστικών που είχαν προκύψει από τη διαδικασία SHAP και ανάλυσης σταθερότητας.

Για τον σκοπό αυτό, στο εκάστοτε σύνολο εκπαίδευσης υπολογίστηκαν κέντρα βάρους (centroids) στον χώρο Valence-Energy-Tension για κάθε κυρίαρχο βασικό συναίσθημα, το οποίο ορίστηκε ως το συναίσθημα με τη μέγιστη τιμή στις ανθρώπινες αξιολογήσεις του αντίστοιχου αποσπάσματος. Έτσι, κάθε συναίσθημα αντιστοιχίστηκε σε ένα χαρακτηριστικό σημείο στον τρισδιάστατο συναισθηματικό χώρο.

Για κάθε προβλεπόμενο σημείο στον χώρο των διαστάσεων υπολογίστηκε η απόστασή του από τα πέντε κέντρα βάρους. Η απόδοση του κυρίαρχου συναισθήματος προέκυψε από το κοντινότερο κέντρο (nearest centroid), ενώ για την παραγωγή ολόκληρου συναισθηματικού προφίλ οι αποστάσεις μετατρέπονταν σε πιθανότητες μέσω συνάρτησης softmax, ώστε να προκύπτει μια κανονικοποιημένη κατανομή των πέντε βασικών συναισθημάτων (Hastie et al., 2009).

Δύο εναλλακτικοί τρόποι μέτρησης της απόστασης από τα centroids εξετάστηκαν: η Euclidean και η Mahalanobis απόσταση (Hastie et al., 2009). Η δεύτερη χρησιμοποιείται για τη μέτρηση αποστάσεων λαμβάνοντας υπόψη τη συνδιακύμανση των μεταβλητών (Mahalanobis, 1936). Επιπλέον, η παράμετρος alpha της softmax ρυθμίστηκε εμπειρικά μέσα από εσωτερική διαδικασία αξιολόγησης στο εκάστοτε σύνολο εκπαίδευσης, με στόχο τη βέλτιστη αναπαράσταση του συναισθηματικού προφίλ. Η τελική αξιολόγηση της χαρτογράφησης πραγματοποιήθηκε σε επίπεδο κυρίαρχου συναισθήματος, πλήρους συναισθηματικού προφίλ και ομοιότητας μεταξύ πραγματικών και προβλεπόμενων προφίλ, με χρήση των μετρικών που περιγράφονται στην Ενότητα 5.3.

Τέλος, τα αποτελέσματα της έμμεσης πρόβλεψης συναισθηματικού προφίλ συγκρίθηκαν με τα αντίστοιχα της άμεσης πρόβλεψης βασικών συναισθημάτων από τα ηχητικά χαρακτηριστικά, με στόχο την αξιολόγηση της χρησιμότητας του διαστασιακού χώρου ως ενδιάμεσο επίπεδο αναπαράστασης.

5.9 Άμεση Πρόβλεψη Συναισθηματικού Προφίλ και Σύγκριση με Χαρτογράφηση

Στο τελικό στάδιο της πειραματικής διαδικασίας πραγματοποιήθηκε συγκριτική αξιολόγηση των δύο εναλλακτικών προσεγγίσεων για την πρόβλεψη του προφίλ βασικών συναισθημάτων. Η πρώτη προσέγγιση αφορούσε την άμεση πρόβλεψη των πέντε βασικών συναισθημάτων από τα αντικειμενικά ηχητικά χαρακτηριστικά μέσω των τελικών μοντέλων παλινδρόμησης. Η δεύτερη προσέγγιση ήταν η έμμεση πρόβλεψη, κατά την οποία αρχικά προβλέπονταν οι συναισθηματικές διαστάσεις Valence, Energy και Tension από τα ηχητικά χαρακτηριστικά και στη συνέχεια οι προβλεπόμενες διαστάσεις χαρτογραφούνταν σε προφίλ βασικών συναισθημάτων μέσω της διαδικασίας που περιγράφεται στην προηγούμενη ενότητα.

Για την άμεση πρόβλεψη χρησιμοποιήθηκε το τελικό κοινό μειωμένο μοντέλο βασικών συναισθημάτων. Οι προβλέψεις για κάθε συναισθήμα παράχθηκαν μέσω 5-fold cross-validation και στη συνέχεια συνδυάστηκαν σε πλήρες συναισθηματικό προφίλ για κάθε απόσπασμα, με κανονικοποίηση των προβλεπόμενων τιμών ώστε να προκύπτει συγκρίσιμο προφίλ με άθροισμα ίσο με 1. Για την έμμεση πρόβλεψη χρησιμοποιήθηκαν οι προβλεπόμενες συναισθηματικές διαστάσεις από τα τελικά μοντέλα διαστάσεων και η διαδικασία χαρτογράφησης από τον διαστασιακό χώρο στον χώρο των βασικών συναισθημάτων.

Η σύγκριση των δύο προσεγγίσεων πραγματοποιήθηκε με τα ίδια μέτρα αξιολόγησης προφίλ και ειδικότερα την ακρίβεια ως προς το κυρίαρχο συναισθήμα (dominant emotion accuracy), το profile MAE και την cosine similarity. Με τον τρόπο αυτόν αξιολογήθηκε εάν η έμμεση πρόβλεψη μέσω του διαστασιακού χώρου μπορεί να προσεγγίσει ανταγωνιστικά την άμεση πρόβλεψη, καθώς και αν προσφέρει ένα περισσότερο ερμηνεύσιμο ενδιάμεσο πλαίσιο σύνδεσης μεταξύ αντικειμενικών ηχητικών χαρακτηριστικών και βασικών συναισθημάτων.

Στο παρόν Κεφάλαιο παρουσιάστηκε η μεθοδολογική προσέγγιση της πειραματικής διαδικασίας της ΔΕ, από την επιλογή και αξιολόγηση των μοντέλων πρόβλεψης έως τη διαδικασία επιλογής χαρακτηριστικών και τη χαρτογράφηση μεταξύ συναισθηματικών χώρων. Τα αποτελέσματα της εφαρμογής των παραπάνω μεθόδων παρουσιάζονται αναλυτικά στο επόμενο Κεφάλαιο.

6. Πειραματικά Αποτελέσματα

Στο παρόν Κεφάλαιο παρουσιάζονται τα πειραματικά αποτελέσματα της ΔΕ, με βάση τη μεθοδολογική διαδικασία που περιγράφεται στο προηγούμενο Κεφάλαιο. Αρχικά, εξετάζεται η απόδοση διαφορετικών μοντέλων και συνόλων χαρακτηριστικών, με στόχο την επιλογή των αποδοτικότερων για τον σκοπό της έρευνας. Στη συνέχεια παρουσιάζονται τα αποτελέσματα των τελικών μοντέλων για την πρόβλεψη συναισθηματικών διαστάσεων και βασικών συναισθημάτων, καθώς και η επίδραση της επιλογής χαρακτηριστικών στην απόδοσή τους.

Επιπλέον, αναλύονται τα σημαντικότερα χαρακτηριστικά όπως προέκυψαν από τις μεθόδους ερμηνευσιμότητας, εξετάζεται η στατιστική σημαντικότητα των διαφορών μεταξύ των τελικών μοντέλων και παρουσιάζονται τα αποτελέσματα της χαρτογράφησης μεταξύ διαστασιακών και κατηγορικών μορφών συναισθηματικής αναπαράστασης. Τέλος, συγκρίνεται η άμεση πρόβλεψη βασικών συναισθημάτων με την έμμεση προσέγγιση μέσω του διαστασιακού χώρου.

6.1 Αρχική Σύγκριση Μοντέλων και Συνόλων Χαρακτηριστικών

Στο αρχικό στάδιο της πειραματικής διαδικασίας εξετάστηκε η απόδοση διαφορετικών μοντέλων και συνόλων αντικειμενικών χαρακτηριστικών. Η σύγκριση πραγματοποιήθηκε τόσο για τις συναισθηματικές διαστάσεις (Valence, Energy, Tension) όσο και για τα βασικά συναισθήματα (Happy, Sad, Anger, Fear, Tender), χρησιμοποιώντας ως μετρικές το R^2 και το MAE.

Αρχικά, για το βασικό σύνολο των 48 χαρακτηριστικών συγκρίθηκαν οι αλγόριθμοι RandomForestRegressor και GradientBoostingRegressor. Όπως φαίνεται στον Πίνακα 6.1, ο GradientBoostingRegressor παρουσίασε καλύτερη συνολική απόδοση τόσο για τις συναισθηματικές διαστάσεις όσο και για τα βασικά συναισθήματα. Συγκεκριμένα, παρατηρείται υψηλότερο μέσο R^2 και χαμηλότερο μέσο MAE και στις δύο περιπτώσεις σε σχέση με τον RandomForestRegressor. Παρόμοια αποτελέσματα παρατηρήθηκαν σε επιπλέον αξιολογήσεις μοντέλων και ως εκ τούτου ο GradientBoostingRegressor επιλέχθηκε ως ο βασικός αλγόριθμος για τα επόμενα στάδια της ΔΕ.

Πίνακας 6.1: Αρχική σύγκριση RandomForestRegressor και GradientBoostingRegressor στο σύνολο των 48 χαρακτηριστικών

Ομάδα στόχων	Μοντέλο	Χαρακτηριστικά (N)	Mean R ²	Mean MAE
Συναισθηματικές διαστάσεις	RandomForestRegressor	48	0.50	1.04
	GradientBoostingRegressor	48	0.55	0.97
Βασικά συναισθήματα	RandomForestRegressor	48	0.37	1.03
	GradientBoostingRegressor	48	0.42	0.96

Στη συνέχεια, εξετάστηκε η επίδραση της σύνθεσης του συνόλου χαρακτηριστικών στην απόδοση των μοντέλων. Για τον σκοπό αυτό συγκρίθηκαν το αρχικό σύνολο των 48 χαρακτηριστικών, το εναλλακτικό σύνολο των 21 χαρακτηριστικών και το τελικό διευρυμένο σύνολο των 53 χαρακτηριστικών. Η σύγκριση πραγματοποιήθηκε με χρήση του GradientBoostingRegressor και τα αποτελέσματα παρουσιάζονται στον Πίνακα 6.2.

Πίνακας 6.2: Αρχική σύγκριση συνόλων χαρακτηριστικών με GradientBoostingRegressor

Ομάδα στόχων	Σύνολο χαρακτηριστικών	Χαρακτηριστικά (N)	Mean R ²	Mean MAE
Συναισθηματικές διαστάσεις	Αρχικό σύνολο	48	0.55	0.97
	Εναλλακτικό σύνολο	21	0.54	0.97
	Διευρυμένο σύνολο	53	0.57	0.96
Βασικά συναισθήματα	Αρχικό σύνολο	48	0.42	0.96
	Εναλλακτικό σύνολο	21	0.42	0.93
	Διευρυμένο σύνολο	53	0.41	0.95

Για τις συναισθηματικές διαστάσεις, το διευρυμένο σύνολο των 53 χαρακτηριστικών οδήγησε σε ελαφρώς βελτιωμένη απόδοση σε σχέση με το αρχικό σύνολο των 48 χαρακτηριστικών. Αντίθετα, για τα βασικά συναισθήματα, το εναλλακτικό σύνολο των 21 χαρακτηριστικών παρουσίασε ανταγωνιστική απόδοση σε σχέση με τα άλλα δύο σύνολα, με χαμηλότερο μέσο MAE.

Παρά την ανταγωνιστική συμπεριφορά του μικρότερου συνόλου χαρακτηριστικών, το διευρυμένο σύνολο των 53 χαρακτηριστικών επιλέχθηκε ως βασική δεξαμενή υποψήφιων μεταβλητών για τα τελικά πειράματα, δεδομένου ότι παρείχε τον πληρέστερο διαθέσιμο χώρο χαρακτηριστικών και επέτρεπε τη μετέπειτα εφαρμογή διαδικασιών επιλογής χαρακτηριστικών μέσω SHAP και ανάλυσης σταθερότητας. Παρότι διαδικασίες μείωσης χαρακτηριστικών εφαρμόστηκαν και στα τρία σύνολα χαρακτηριστικών, τα μειωμένα

μοντέλα που βασίστηκαν στο διευρυμένο σύνολο παρουσίασαν την καλύτερη συνολική απόδοση. Για τον λόγο αυτό, οι επόμενες Ενότητες επικεντρώνονται στα μοντέλα που προέκυψαν από το διευρυμένο σύνολο χαρακτηριστικών.

6.2 Τελικά Μοντέλα για Συναισθηματικές Διαστάσεις

Μετά την αρχική σύγκριση μοντέλων και συνόλων χαρακτηριστικών, αξιολογήθηκαν τα τελικά μοντέλα πρόβλεψης των συναισθηματικών διαστάσεων Valence, Energy και Tension, με βάση το διευρυμένο σύνολο των 53 αντικειμενικών χαρακτηριστικών. Συγκεκριμένα, εξετάστηκαν τρεις διαμορφώσεις: το πλήρες διευρυμένο σύνολο χαρακτηριστικών (FULL_EXTENDED_53F), το κοινό μειωμένο σύνολο χαρακτηριστικών για όλους τους στόχους (REDUCED_EXTENDED_COMMON_N $\geq$ 4) και τα μειωμένα σύνολα χαρακτηριστικών ανά στόχο (REDUCED_EXTENDED_PER_TARGET_N $\geq$ 3).

Η αξιολόγηση πραγματοποιήθηκε με βάση τις μετρικές R^2 και MAE, τόσο ανά διάσταση όσο και συνολικά μέσω του μέσου όρου των τριών στόχων. Τα αποτελέσματα παρουσιάζονται στον Πίνακα 6.3.

Πίνακας 6.3: Απόδοση τελικών μοντέλων για την πρόβλεψη συναισθηματικών διαστάσεων

Μοντέλο	Valence		Energy		Tension		Mean R^2	Mean MAE
	R^2	MAE	R^2	MAE	R^2	MAE		
FULL_EXTENDED_53F	0.45	1.04	0.73	0.75	0.52	1.09	0.57	0.96
REDUCED_EXTENDED_COMMON_N $\geq$ 4	0.50	0.98	0.74	0.72	0.58	1.01	0.60	0.90
REDUCED_EXTENDED_PER_TARGET_N $\geq$ 3	0.54	0.92	0.75	0.74	0.60	0.96	0.63	0.87

Τα αποτελέσματα δείχνουν ότι η μείωση χαρακτηριστικών οδήγησε σε βελτίωση της απόδοσης σε σχέση με το πλήρες διευρυμένο σύνολο. Το κοινό μειωμένο σύνολο χαρακτηριστικών παρουσίασε υψηλότερο μέσο R^2 και χαμηλότερο μέσο MAE σε σχέση με το FULL_EXTENDED_53F, γεγονός που υποδηλώνει ότι η αφαίρεση λιγότερο σταθερά σημαντικών χαρακτηριστικών συνέβαλε θετικά στη γενίκευση του μοντέλου.

Η καλύτερη συνολική απόδοση επιτεύχθηκε από το REDUCED_EXTENDED_PER_TARGET_N $\geq$ 3, το οποίο αξιοποίησε διαφορετικό υποσύνολο χαρακτηριστικών για κάθε συναισθηματική διάσταση. Το μοντέλο αυτό πέτυχε

το υψηλότερο μέσο R^2 (0.63) και το χαμηλότερο μέσο MAE (0.87). Εμφανής ήταν επίσης η βελτίωση και στις τρεις διαστάσεις, με ιδιαίτερα υψηλή επίδοση στην πρόβλεψη του Energy ($R^2 = 0.75$) και βελτιωμένη απόδοση για τα Valence και Tension σε σχέση με τις υπόλοιπες διαμορφώσεις.

Συνολικά, τα αποτελέσματα υποδεικνύουν ότι η πρόβλεψη των συναισθηματικών διαστάσεων επωφελείται από την ανά στόχο επιλογή χαρακτηριστικών, η οποία οδηγεί σε πιο αποδοτικά μοντέλα πρόβλεψης. Για τον λόγο αυτό, η συγκεκριμένη διαμόρφωση αξιοποιήθηκε και στη μετέπειτα διαδικασία χαρτογράφησης από συναισθηματικές διαστάσεις σε βασικά συναισθήματα.

6.3 Τελικά Μοντέλα Πρόβλεψης Βασικών Συναισθημάτων

Με αντίστοιχη διαδικασία αξιολογήθηκαν και τα τελικά μοντέλα πρόβλεψης των πέντε βασικών συναισθημάτων: Happy, Sad, Anger, Fear, Tender. Ομοίως, η αξιολόγηση βασίστηκε στο διευρυμένο σύνολο των 53 χαρακτηριστικών και στις μειωμένες εκδοχές του όπως προέκυψαν από τη διαδικασία SHAP και ανάλυσης σταθερότητας.

Στην περίπτωση των βασικών συναισθημάτων εξετάστηκαν τρεις βασικές διαμορφώσεις: το πλήρες διευρυμένο σύνολο χαρακτηριστικών (FULL_EXTENDED_53F), το κοινό μειωμένο σύνολο χαρακτηριστικών (REDUCED_EXTENDED_COMMON_N $\geq$ 4) και τα μειωμένα σύνολα χαρακτηριστικών ανά στόχο (REDUCED_EXTENDED_PER_TARGET_N $\geq$ 3). Τα αποτελέσματα παρουσιάζονται στον Πίνακα 6.4.

Σημειώνεται ότι τα μειωμένα σύνολα χαρακτηριστικών για τα βασικά συναισθήματα προέκυψαν από ξεχωριστή διαδικασία SHAP και ανάλυσης σταθερότητας σε σχέση με τα αντίστοιχα σύνολα των συναισθηματικών διαστάσεων. Επομένως, οι κοινές ονομασίες δηλώνουν την ίδια στρατηγική επιλογής χαρακτηριστικών, αλλά όχι τα ίδια χαρακτηριστικά μεταξύ των δύο ομάδων στόχων.

Τα αποτελέσματα δείχνουν ότι η μείωση χαρακτηριστικών βελτίωσε την απόδοση των μοντέλων σε σχέση με το πλήρες διευρυμένο σύνολο. Το κοινό μειωμένο σύνολο REDUCED_EXTENDED_COMMON_N $\geq$ 4 παρουσίασε την καλύτερη συνολική απόδοση ως προς το μέσο R^2 (0.49) και ίδια επίδοση με την ανά στόχο διαμόρφωση ως προς το μέσο MAE (0.88). Σε σύγκριση με το πλήρες σύνολο, η βελτίωση είναι εμφανής τόσο στο R^2 όσο και στο MAE.

Πίνακας 6.4: Απόδοση τελικών μοντέλων για την πρόβλεψη βασικών συναισθημάτων

Μοντέλο	Happy		Sad		Anger		Fear		Tender		Mean R ²	Mean MAE
	R ²	MAE	R ²	MAE	R ²	MAE	R ²	MAE	R ²	MAE		
FULL_EX TENDED_ 53F	0.42	1.12	0.48	0.86	0.42	0.75	0.42	1.10	0.34	0.91	0.41	0.95
REDUCE D_EXTEN DED_CO MMON_N >=4	0.53	0.97	0.53	0.85	0.51	0.67	0.43	1.06	0.45	0.85	0.49	0.88
REDUCE D_EXTEN DED_PER _TARGET _N>=3	0.49	1.03	0.51	0.85	0.48	0.67	0.45	1.00	0.42	0.83	0.47	0.88

Με βάση το κοινό μειωμένο μοντέλο, οι υψηλότερες τιμές R² παρατηρήθηκαν για τα Happy, Sad και Anger, ενώ το Fear παρέμεινε δυσκολότερο ως προς το σφάλμα πρόβλεψης, παρουσιάζοντας το υψηλότερο MAE. Το Tender εμφάνισε χαμηλότερες τιμές R² σε σχέση με τα περισσότερα άλλα συναισθήματα, παρόλο που η μείωση χαρακτηριστικών οδήγησε σε σαφή βελτίωση σε σχέση με το πλήρες σύνολο.

Η ανά στόχο επιλογή χαρακτηριστικών παρουσίασε επίσης ανταγωνιστική απόδοση, ιδιαίτερα ως προς το μέσο MAE, αλλά δεν ξεπέρασε το κοινό μειωμένο σύνολο ως προς την εξηγούμενη διακύμανση. Για τον λόγο αυτό, το κοινό μειωμένο μοντέλο βασικών συναισθημάτων χρησιμοποιήθηκε για την άμεση πρόβλεψη συναισθηματικού προφίλ στις επόμενες αναλύσεις.

Σε σχέση με τα αποτελέσματα των Eerola et al. (2009), τα τελικά μοντέλα της παρούσας ΔΕ παρουσίασαν συγκρίσιμη εικόνα απόδοσης με ορισμένες διαφοροποιήσεις. Για τις συναισθηματικές διαστάσεις, το ανά στόχο μειωμένο μοντέλο της παρούσας ΔΕ πέτυχε R² = 0.54 για το Valence, R² = 0.75 για το Energy και R² = 0.60 για το Tension. Οι αντίστοιχες τιμές της Πολλαπλής Γραμμικής Παλινδρόμησης (Multiple Linear Regression, MLR) των Eerola et al. (2009) ήταν 0.64, 0.75 και 0.67, ενώ οι καλύτερες τιμές με Μερική Παλινδρόμηση Ελαχίστων Τετραγώνων (Partial Least Squares Regression, PLS) και Box-Cox transformations ήταν υψηλότερες. Για τα βασικά συναισθήματα, το κοινό μειωμένο μοντέλο της παρούσας ΔΕ παρουσίασε R² = 0.53 για το Happy, 0.53 για το Sad, 0.51 για το Anger, 0.43 για το Fear και 0.45 για το Tender. Οι τιμές αυτές είναι ανταγωνιστικές σε σχέση με την MLR των Eerola et al. για τους περισσότερους στόχους, ιδιαίτερα για τα

Happy, Sad, Anger και Tender, αλλά χαμηλότερες από τις καλύτερες PLS εκδοχές με transformations. Η σύγκριση αυτή παραμένει ενδεικτική, καθώς στην παρούσα ΔΕ χρησιμοποιήθηκε διαφορετικό σύνολο αντικειμενικών χαρακτηριστικών, διαφορετική βιβλιοθήκη εξαγωγής χαρακτηριστικών, μη γραμμικά μοντέλα παλινδρόμησης και διαφορετική διαδικασία επιλογής χαρακτηριστικών.

Σημειώνεται ότι αν και οι PLS προσεγγίσεις παρουσίασαν υψηλότερες τιμές R^2 στην έρευνα των Eerola et al. (2009), το MIRtoolbox υλοποιεί την εκδοχή MLR με τα πέντε σημαντικότερα προβλεπτικά χαρακτηριστικά ανά στόχο (Lartillot, 2017). Η επιλογή αυτή πιθανώς συνδέεται με την πιο άμεση ερμηνευσιμότητα των επιμέρους χαρακτηριστικών. Επίσης, τα Box-Cox transformations δεν ενσωματώθηκαν στη συγκεκριμένη έκδοση του miremotion μέχρι να επιβεβαιωθούν οι αντίστοιχες παράμετροι μετασχηματισμού σε μεγαλύτερο μουσικό δείγμα. Στην επόμενη Ενότητα εξετάζεται επίσης η σχέση των χαρακτηριστικών που αξιοποιούνται στο miremotion με αυτά που προέκυψαν από τη διαδικασία επιλογής χαρακτηριστικών της παρούσας ΔΕ.

6.4 Αποτελέσματα Επιλογής Χαρακτηριστικών και Ερμηνευσιμότητας

Η βελτίωση που παρατηρήθηκε στα μειωμένα μοντέλα συνδέεται με τη διαδικασία επιλογής χαρακτηριστικών που περιγράφεται στο Κεφάλαιο 5. Η επιλογή βασίστηκε στον συνδυασμό της μέσης απόλυτης τιμής SHAP και της σταθερότητας εμφάνισης των χαρακτηριστικών στα σημαντικότερα χαρακτηριστικά των διαφορετικών folds. Με τον τρόπο αυτόν διατηρήθηκαν χαρακτηριστικά που ήταν ταυτόχρονα σημαντικά για τις προβλέψεις και σχετικά σταθερά ως προς τη διαδικασία της διασταυρούμενης επικύρωσης (cross-validation).

Στον Πίνακα 6.5 παρουσιάζονται τα κοινά μειωμένα σύνολα χαρακτηριστικών που χρησιμοποιήθηκαν στα τελικά μοντέλα συναισθηματικών διαστάσεων και βασικών συναισθημάτων. Όπως παρατηρείται, αρκετά χαρακτηριστικά εμφανίζονται και στα δύο σύνολα, γεγονός που υποδηλώνει ότι ορισμένες ηχητικές ιδιότητες είναι χρήσιμες και για τις δύο μορφές συναισθηματικής αναπαράστασης.

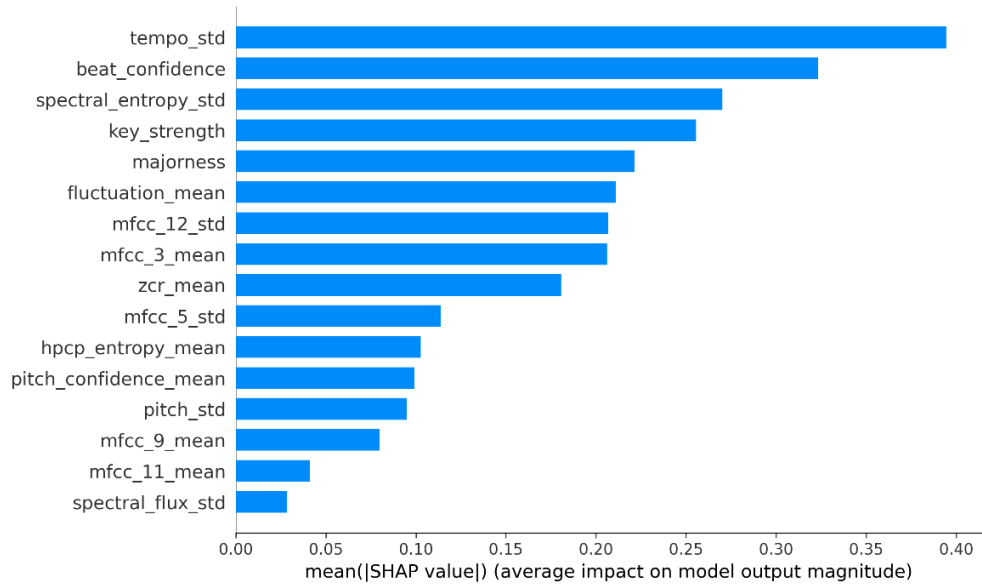
Πίνακας 6.5: Κοινά μειωμένα σύνολα χαρακτηριστικών για τα τελικά μοντέλα

Ομάδα στόχων	Χαρακτηριστικά (N)	Επιλεγμένα χαρακτηριστικά
Συναισθηματικές διαστάσεις	14	beat_confidence , fluctuation_mean , hpcp_entropy_mean , key_strength , majoriness , mfcc_11_mean , mfcc_12_std , mfcc_5_std , mfcc_9_mean , pitch_confidence_mean , spectral_entropy_mean , spectral_entropy_std , spectral_flux_mean , spectral_flux_std
Βασικά συναισθήματα	16	beat_confidence , fluctuation_mean , hpcp_entropy_mean , key_strength , majoriness , mfcc_11_mean , mfcc_12_std , mfcc_3_mean , mfcc_5_std , mfcc_9_mean , pitch_confidence_mean , pitch_std , spectral_entropy_std , spectral_flux_std , tempo_std , zcr_mean

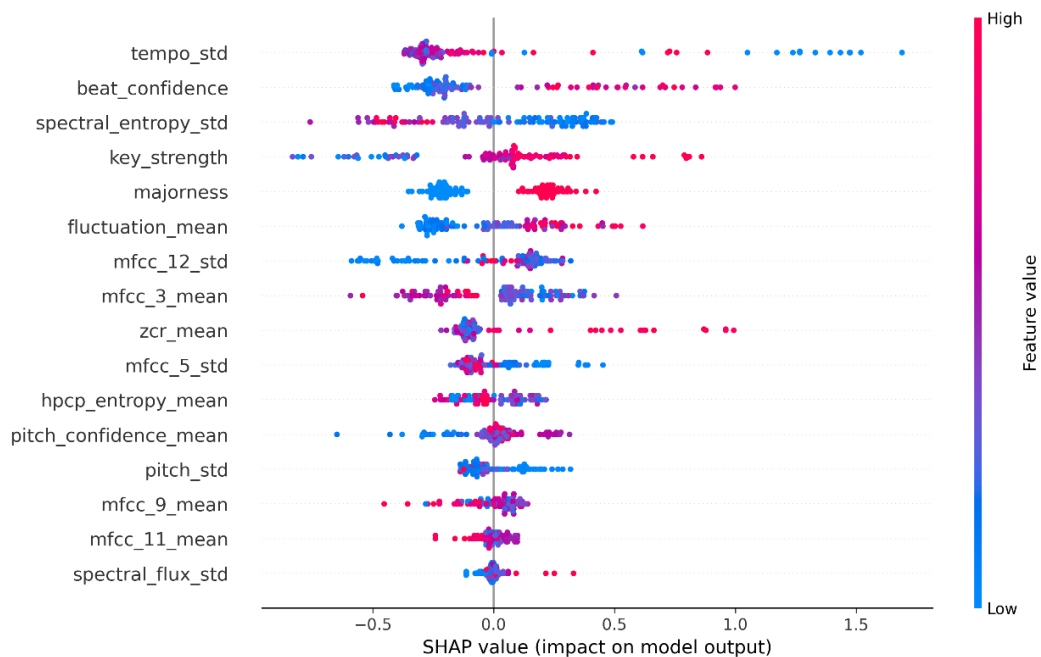
Παρατηρείται επίσης ότι στα κοινά μειωμένα σύνολα κυριαρχούν κυρίως τονικά/αρμονικά και ηχοχρωματικά/φασματικά χαρακτηριστικά. Χαρακτηριστικά όπως τα fluctuation_mean, hpcp_entropy_mean, key_strength, majoriness, pitch_confidence_mean, spectral_entropy_std, spectral_flux_std και αρκετοί συντελεστές MFCC εμφανίζονται και στις δύο ομάδες στόχων. Αυτό δείχνει ότι η τονική οργάνωση, το φασματικό περιεχόμενο και η ηχοχρωματική δομή του μουσικού αποσπάσματος είχαν σταθερά σημαντική συμβολή στην πρόβλεψη των συναισθηματικών αξιολογήσεων. Αξιοσημείωτο είναι το γεγονός ότι σε κανένα σύνολο δεν περιλαμβάνεται κάποιο χαρακτηριστικό έντασης και δυναμικής.

Παράλληλα, τα βασικά συναισθήματα περιλαμβάνουν επιπλέον χαρακτηριστικά όπως τα tempo_std, zcr_mean και pitch_std, τα οποία δεν εμφανίζονται στο αντίστοιχο κοινό μειωμένο σύνολο των διαστάσεων. Αυτό υποδηλώνει ότι ενδεχομένως η πρόβλεψη των συγκεκριμένων μεταβλητών/στόχων επηρεάζεται περισσότερο από επιπλέον πτυχές της ρυθμικής μεταβλητότητας, της ηχοχρωματικής υφής και της μεταβλητότητας του τονικού ύψους.

Ενδεικτικά, στις Εικόνες 6.1 και 6.2 παρουσιάζονται δύο γραφήματα που προέκυψαν από την ανάλυση SHAP. Συγκεκριμένα, η Εικόνα 6.1 αποτυπώνει τη συνολική κατάταξη των χαρακτηριστικών για την πρόβλεψη του Happy με χρήση του κοινού μειωμένου συνόλου, ενώ η Εικόνα 6.2 παρουσιάζει την κατανομή των τιμών SHAP για τα ίδια χαρακτηριστικά, αποτυπώνοντας τόσο τη σημασία όσο και την κατεύθυνση της επίδρασής τους στις προβλέψεις.



Εικόνα 6.1: SHAP bar plot για την πρόβλεψη του συναισθήματος Happy



Εικόνα 6.2: SHAP beeswarm plot για την πρόβλεψη του συναισθήματος Happy

Στον Πίνακα 6.6 παρουσιάζονται τα ανά στόχο μειωμένα σύνολα χαρακτηριστικών. Τα αποτελέσματα αυτά είναι χρήσιμα για την ερμηνεία των διαφορών μεταξύ των επιμέρους στόχων.

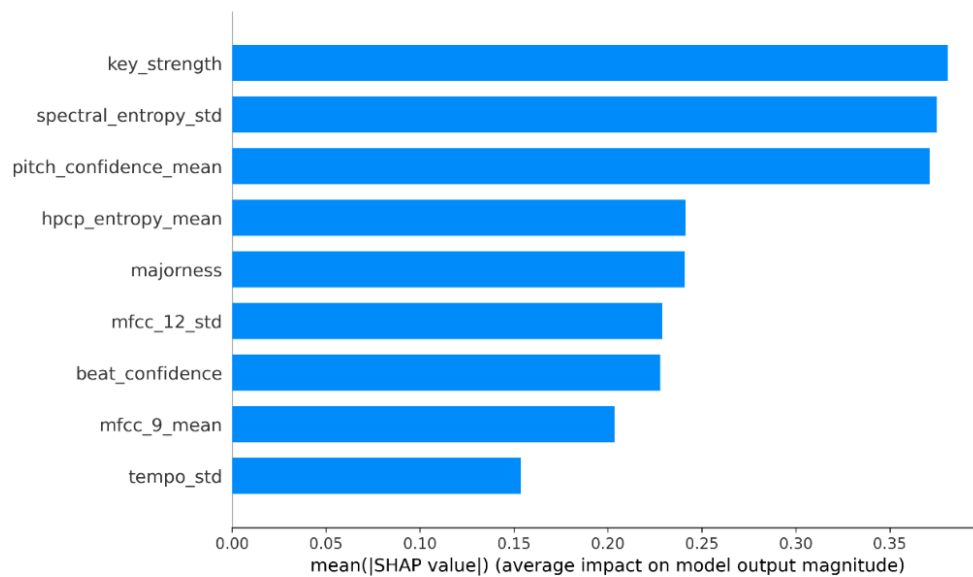
Πίνακας 6.6: Ανά στόχο μειωμένα σύνολα χαρακτηριστικών για τα τελικά μοντέλα

Στόχος	Χαρακτηριστικά (N)	Επιλεγμένα χαρακτηριστικά
Valence	9	beat_confidence , hpcp_entropy_mean , key_strength , majoriness , mfcc_12_std , mfcc_9_mean , pitch_confidence_mean , spectral_entropy_std , tempo_std
Energy	10	beat_confidence , fluctuation_mean , hpcp_entropy_mean , mfcc_11_mean , mfcc_1_std , mfcc_8_mean , spectral_entropy_mean , spectral_flux_mean , spectral_flux_std , tempo_bpm
Tension	7	hpcp_entropy_mean , key_strength , mfcc_2_mean , mfcc_5_std , mfcc_9_mean , pitch_confidence_mean , spectral_flux_std
Happy	8	beat_confidence , key_strength , majoriness , mfcc_12_std , mfcc_3_mean , spectral_entropy_std , tempo_std , zcr_mean
Sad	9	beat_confidence , fluctuation_mean , hpcp_entropy_mean , majoriness , mfcc_11_mean , pitch_confidence_mean , pitch_std , spectral_flux_mean , spectral_flux_std
Anger	6	hpcp_entropy_mean , key_strength , mfcc_12_std , pitch_confidence_mean , pitch_salience_mean , spectral_flux_std
Fear	7	hpcp_entropy_mean , key_strength , mfcc_12_std , mfcc_5_std , mfcc_6_mean , mfcc_9_mean , pitch_confidence_mean
Tender	8	hpcp_entropy_mean , mfcc_5_std , mfcc_9_mean , pitch_confidence_mean , rms_std , spectral_centroid_mean , spectral_flux_mean , tempo_bpm

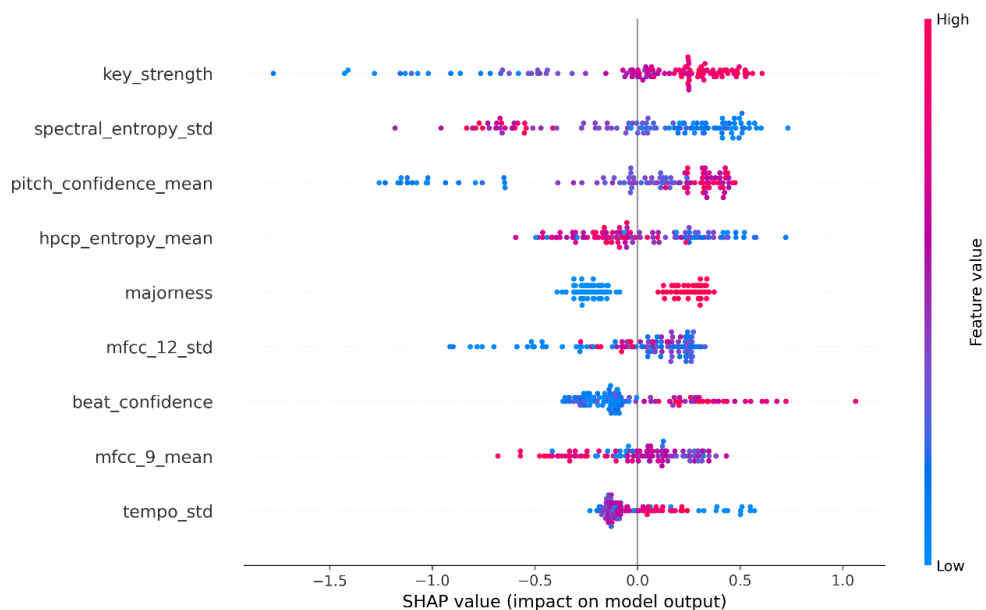
Η ανά στόχο επιλογή χαρακτηριστικών ανέδειξε ορισμένες διαφοροποιήσεις μεταξύ των συναισθηματικών μεταβλητών. Για παράδειγμα, το Energy συνδέθηκε με χαρακτηριστικά που περιγράφουν τη φασματική και ρυθμική μεταβολή, όπως τα fluctuation_mean, spectral_flux_mean, spectral_flux_std και tempo_bpm. Αντίθετα, το Valence και το Happy περιλαμβάνουν χαρακτηριστικά που σχετίζονται με την τονικότητα και τον μείζονα/ελάσσονα χαρακτήρα, όπως τα key_strength και majoriness. Το Tension, το Anger και το Fear συνδέθηκαν με χαρακτηριστικά που αφορούν τη φασματική μεταβλητότητα, την τονική οργάνωση και την ασάφεια του τονικού περιεχομένου. Το Tender περιλαμβάνει επίσης έναν συνδυασμό χαρακτηριστικών έντασης, ρυθμικών, τονικών/αρμονικών και

ηχοχρωματικών/φασματικών χαρακτηριστικών. Αξιοσημείωτο είναι και σε αυτήν την περίπτωση ότι κανένα άλλο σύνολο εκτός από αυτό του Tender δεν περιλαμβάνει χαρακτηριστικό έντασης και δυναμικής.

Στις Εικόνες 6.3 και 6.4 παρουσιάζονται επίσης ενδεικτικά τα γραφήματα της ανάλυσης SHAP σχετικά με την πρόβλεψη του Valence με χρήση του μειωμένου συνόλου χαρακτηριστικών ανά στόχο.



Εικόνα 6.3: SHAP bar plot για την πρόβλεψη της διάστασης Valence



Εικόνα 6.4: SHAP beeswarm plot για την πρόβλεψη της διάστασης Valence

Συγκρίνοντας τα χαρακτηριστικά που αναδείχθηκαν από τη διαδικασία SHAP με τα χαρακτηριστικά που αξιοποιούνται στο μοντέλο miremotion του MIRtoolbox, παρατηρείται μερική σύγκλιση ως προς τις βασικές κατηγορίες ηχητικής πληροφορίας. Σύμφωνα με το MIRtoolbox User's Manual, τα μοντέλα πρόβλεψης συναισθηματικών διαστάσεων και κατηγοριών βασίζονται σε επιλεγμένα χαρακτηριστικά που σχετίζονται με τη δυναμική του σήματος, τη φασματική κατανομή, τη ρυθμική/χρονική μεταβολή και την τονική οργάνωση, όπως RMS, fluctuation, spectral centroid, spectral spread, spectral entropy, key clarity και mode (Lartillot, 2017). Αντίστοιχα, από την παρούσα ανάλυση αναδείχθηκαν συχνά χαρακτηριστικά όπως τα fluctuation_mean, key_strength, majorness, spectral_entropy_std, spectral_flux_std, hpcp_entropy_mean και διάφοροι συντελεστές MFCC.

Η σύγκλιση αυτή είναι ιδιαίτερα εμφανής στα τονικά/αρμονικά χαρακτηριστικά, καθώς τα key_strength και majorness που επιλέχθηκαν στα τελικά μοντέλα αντιστοιχούν εννοιολογικά στα key clarity και mode που χρησιμοποιούνται στο MIRtoolbox. Ομοίως, η συχνή παρουσία των fluctuation_mean, spectral_entropy και spectral_flux δείχνει ότι η χρονική και φασματική μεταβολή του μουσικού σήματος αποτελεί σημαντική πηγή πληροφορίας και στις δύο προσεγγίσεις. Ωστόσο, παρατηρούνται και ορισμένες διαφοροποιήσεις. Χαρακτηριστικά έντασης και δυναμικής, όπως το RMS, εμφανίζονται στα μοντέλα του MIRtoolbox, αλλά είχαν περιορισμένη παρουσία στα τελικά επιλεγμένα σύνολα της παρούσας ΔΕ. Αντίθετα, οι συντελεστές MFCC και το hpcp_entropy_mean αναδείχθηκαν συχνά από την ανάλυση SHAP, υποδηλώνοντας ότι η ηχοχρωματική δομή και η τονική αβεβαιότητα συνέβαλαν περισσότερο στα συγκεκριμένα μοντέλα πρόβλεψης.

Σημειώνεται ότι πρόκειται για δύο προσεγγίσεις σε διαφορετικό υπολογιστικό περιβάλλον, με διαφορετική υλοποίηση χαρακτηριστικών και διαφορετικά μοντέλα παλινδρόμησης. Ως εκ τούτου, δεν είναι δυνατή η άμεση αντιστοίχιση των χαρακτηριστικών. Ωστόσο, η μερική σύγκλιση των επιλεγμένων χαρακτηριστικών με αυτά που αξιοποιούνται στο MIRtoolbox ενισχύει την ερμηνευσιμότητα των αποτελεσμάτων, καθώς δείχνει ότι τα τελικά μοντέλα αξιοποίησαν ηχητικές ιδιότητες που έχουν ήδη συνδεθεί βιβλιογραφικά με τη συναισθηματική έκφραση στη μουσική.

Γενικά, η ανάλυση ερμηνευσιμότητας υποδεικνύει ότι η συναισθηματική πληροφορία δεν αποτυπώνεται σε μία μόνο κατηγορία ηχητικών χαρακτηριστικών. Αντίθετα, τα τελικά μοντέλα αξιοποίησαν έναν συνδυασμό χαρακτηριστικών με κυρίαρχα τα τονικά/αρμονικά

και ηχοχρωματικά/φασματικά χαρακτηριστικά. Η σταθερή παρουσία χαρακτηριστικών όπως το `hrcp_entropy_mean`, το `key_strength`, το `pitch_confidence_mean` και οι συντελεστές MFCC υποδηλώνει ότι η τονική οργάνωση και η ηχοχρωματική δομή αποτέλεσαν βασικούς συντελεστές πρόβλεψης των συναισθηματικών αξιολογήσεων.

6.5 Στατιστική Σύγκριση Τελικών Μοντέλων

Για την περαιτέρω αξιολόγηση των διαφορών απόδοσης μεταξύ των τελικών μοντέλων πραγματοποιήθηκε έλεγχος σημαντικότητας με χρήση του Wilcoxon signed-rank test. Ο έλεγχος εφαρμόστηκε στα απόλυτα σφάλματα πρόβλεψης που προέκυψαν για τα ίδια μουσικά αποσπάσματα από δύο διαφορετικά μοντέλα. Στόχος ήταν η εξέταση της στατιστικής σημαντικότητας της μείωσης του σφάλματος στα μειωμένα μοντέλα.

Για κάθε σύγκριση υπολογίστηκε το απόλυτο σφάλμα ανά δείγμα και ανά στόχο, ενώ υπολογίστηκε και ένα συνολικό μέτρο σφάλματος ανά δείγμα, ως μέσος όρος των απόλυτων σφαλμάτων όλων των στόχων.

Στην περίπτωση των συναισθηματικών διαστάσεων πραγματοποιήθηκαν δύο συγκρίσεις. Η πρώτη αφορούσε το πλήρες διευρυμένο μοντέλο `FULL_EXTENDED_53F` σε σχέση με το κοινό μειωμένο μοντέλο `REDUCED_EXTENDED_COMMON_N>=4`. Η δεύτερη αφορούσε τη σύγκριση του κοινού μειωμένου μοντέλου με το ανά στόχο μειωμένο μοντέλο `REDUCED_EXTENDED_PER_TARGET_N>=3`. Στον Πίνακα 6.7 παρουσιάζονται τα συνολικά αποτελέσματα για τα τελικά μοντέλα συναισθηματικών διαστάσεων.

Παρατηρείται ότι το κοινό μειωμένο σύνολο οδήγησε σε στατιστικά σημαντική μείωση του συνολικού σφάλματος ($p = 0.0354$). Ωστόσο, οι αντίστοιχες επιμέρους συγκρίσεις ανά στόχο δεν ήταν στατιστικά σημαντικές για το Valence ($p = 0.1363$), το Energy ($p = 0.1850$) και το Tension ($p = 0.1294$). Επομένως, η βελτίωση του κοινού μειωμένου μοντέλου αποτυπώνεται κυρίως στο συνολικό επίπεδο πρόβλεψης των διαστάσεων.

Η σύγκριση μεταξύ του κοινού μειωμένου μοντέλου και του ανά στόχο μειωμένου μοντέλου δεν έδειξε στατιστικά σημαντική διαφορά ως προς το συνολικό σφάλμα ($p = 0.2577$), παρά το γεγονός ότι το ανά στόχο μειωμένο μοντέλο παρουσίασε χαμηλότερο συνολικό MAE. Αντίστοιχα, οι επιμέρους συγκρίσεις ανά στόχο δεν ήταν στατιστικά σημαντικές για το Valence ($p = 0.1485$), το Energy ($p = 0.4029$) και το Tension ($p = 0.2430$).

Πίνακας 6.7: Στατιστική σύγκριση τελικών μοντέλων συναισθηματικών διαστάσεων με Wilcoxon signed-rank test

Σύγκριση	MAE	W	p-value
FULL_EXTENDED_53F			
vs	0.96 → 0.90	2347.0	0.0354
REDUCED_EXTENDED_COMMON_N>=4			
REDUCED_EXTENDED_COMMON_N>=4			
vs	0.90 → 0.87	2673.0	0.2577
REDUCED_EXTENDED_PER_TARGET_N>=3			

Στον Πίνακα 6.8 παρουσιάζονται τα συνολικά αποτελέσματα για τα τελικά μοντέλα βασικών συναισθημάτων. Στην περίπτωση αυτήν πραγματοποιήθηκε μια σύγκριση μεταξύ του πλήρους διευρυμένου μοντέλου FULL_EXTENDED_53F και του κοινού μειωμένου μοντέλου REDUCED_EXTENDED_COMMON_N>=4. Παρατηρείται ότι το κοινό μειωμένο μοντέλο οδήγησε σε χαμηλότερο συνολικό MAE και παρουσίασε στατιστικά σημαντική μείωση του συνολικού σφάλματος ($p = 0.0033$). Σε επίπεδο επιμέρους στόχων, στατιστικά σημαντική βελτίωση παρατηρήθηκε για τα Happy ($p = 0.0041$) και Anger ($p = 0.0211$). Αντίθετα, οι διαφορές δεν ήταν στατιστικά σημαντικές για τα Sad ($p = 0.9726$), Fear ($p = 0.2503$) και Tender ($p = 0.2164$).

Πίνακας 6.8: Στατιστική σύγκριση τελικών μοντέλων βασικών συναισθημάτων με Wilcoxon signed-rank test

Σύγκριση	MAE	W	p-value
FULL_EXTENDED_53F			
vs	0.95 → 0.88	2068.0	0.0033
REDUCED_EXTENDED_COMMON_N>=4			

Συνολικά, η μείωση των χαρακτηριστικών οδήγησε σε απλούστερα μοντέλα και η χρήση ενός κοινού μειωμένου συνόλου χαρακτηριστικών έναντι του πλήρους συνόλου ανέδειξε στατιστικά σημαντική βελτίωση του συνολικού σφάλματος πρόβλεψης τόσο για τις συναισθηματικές διαστάσεις όσο και για τα βασικά συναισθήματα. Ωστόσο, η ανά στόχο επιλογή χαρακτηριστικών για τις συναισθηματικές διαστάσεις, αν και παρουσίασε χαμηλότερο MAE, δεν εμφάνισε στατιστικά σημαντική βελτίωση σε σχέση με την κοινή επιλογή χαρακτηριστικών.

6.6 Χαρτογράφηση μεταξύ Συναισθηματικών Χώρων

Στο επόμενο στάδιο εξετάστηκε η δυνατότητα χαρτογράφησης μεταξύ των δύο μορφών συναισθηματικής αναπαράστασης, δηλαδή από τον διαστασιακό χώρο Valence-Energy-Tension προς τον χώρο των βασικών συναισθημάτων Happy, Sad, Anger, Fear και Tender. Στόχος της ανάλυσης ήταν η διερεύνηση της δυνατότητας των συναισθηματικών διαστάσεων να λειτουργήσουν ως ενδιάμεσο επίπεδο αναπαράστασης για την εκτίμηση του πλήρους συναισθηματικού προφίλ.

Αρχικά, πραγματοποιήθηκε χαρτογράφηση από τις πραγματικές τιμές των συναισθηματικών διαστάσεων προς τις πραγματικές τιμές των βασικών συναισθημάτων. Σε αυτήν την περίπτωση, εφόσον δεν υπάρχει σφάλμα από την πρόβλεψη των διαστάσεων, εξετάστηκε ο βαθμός στον οποίο οι δύο συναισθηματικοί χώροι συνδέονται μεταξύ τους. Τα αποτελέσματα παρουσιάζονται στον Πίνακα 6.9 μαζί με τα αποτελέσματα των Eerola et al. (2009) που εξέτασαν επίσης αυτήν τη σχέση με χρήση διαφορετικών μοντέλων (Ridge Regression).

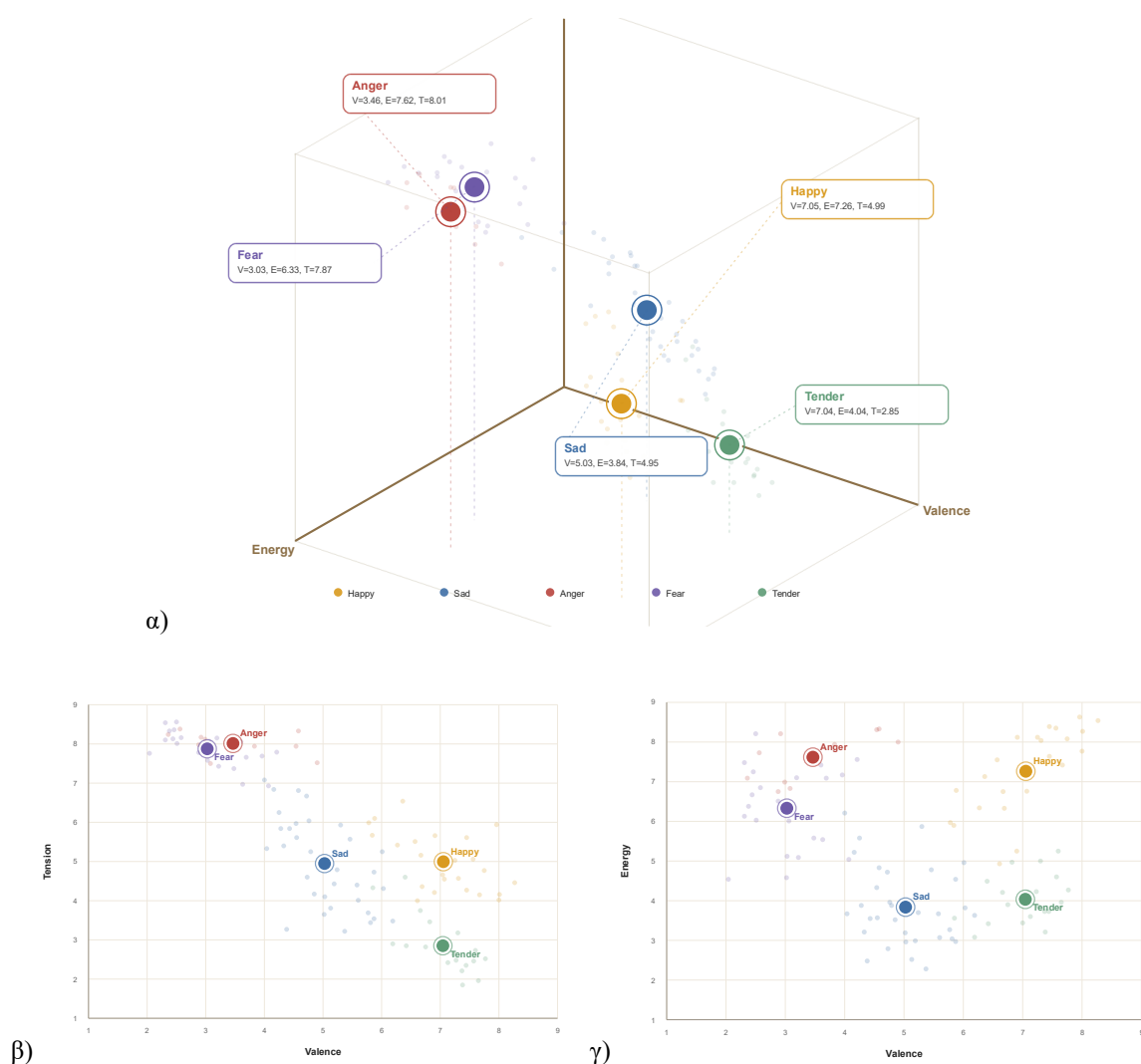
Πίνακας 6.9: Χαρτογράφηση από πραγματικές συναισθηματικές διαστάσεις σε πραγματικά βασικά συναισθήματα

Μοντέλο	Mean R^2	Mean MAE
GradientBoostingRegressor	0.84	0.42
Ridge Regression (Eerola et al., 2009)	0.76	-

Τα αποτελέσματα δείχνουν ότι οι πραγματικές συναισθηματικές διαστάσεις περιέχουν σημαντική πληροφορία για την πρόβλεψη των βασικών συναισθημάτων. Ο GradientBoostingRegressor παρουσίασε συνολική απόδοση με μέσο $R^2 = 0.84$ και μέσο $MAE = 0.42$. Η τιμή R^2 είναι υψηλότερη από την αντίστοιχη τιμή $R^2 = 0.76$ που αναφέρεται από τους Eerola et al. (2009) για γραμμικό μοντέλο Ridge Regression, γεγονός που υποδηλώνει ότι η μη γραμμική προσέγγιση που εφαρμόστηκε στην παρούσα ΔΕ αποτύπωσε αποτελεσματικά τη σύνδεση μεταξύ των δύο συναισθηματικών χώρων.

Στη συνέχεια εξετάστηκε η περίπτωση της χαρτογράφησης από προβλεπόμενες συναισθηματικές διαστάσεις σε βασικά συναισθήματα. Στην περίπτωση αυτή, ως είσοδος χρησιμοποιήθηκαν οι προβλέψεις του ανά στόχο μειωμένου μοντέλου REDUCED_EXTENDED_PER_TARGET_N $\geq$ 3 για τα Valence, Energy και Tension, και

στη συνέχεια εφαρμόστηκε η διαδικασία χαρτογράφησης μέσω κέντρων βάρους στον διαστασιακό χώρο. Η οπτική αναπαράσταση των κέντρων βάρους στον διαστασιακό χώρο βοηθά στην ερμηνεία της διαδικασίας χαρτογράφησης. Όπως φαίνεται στην Εικόνα 6.5, ορισμένα συναισθήματα καταλαμβάνουν πιο διακριτές περιοχές του χώρου, ενώ άλλα εμφανίζουν μεγαλύτερη επικάλυψη. Η εγγύτητα μεταξύ Sad και Tender, καθώς και η σχετική γειτνίαση αρνητικών συναισθημάτων όπως Anger και Fear, μπορεί να συμβάλλει στην ερμηνεία των συγχύσεων που παρατηρήθηκαν στον πίνακα σύγχυσης της χαρτογράφησης.



Εικόνα 6.5: Κέντρα βάρους των βασικών συναισθημάτων στον διαστασιακό χώρο: α) τρισδιάστατη αναπαράσταση Valence-Energy-Tension, β) προβολή στο επίπεδο Valence-Tension και γ) προβολή στο επίπεδο Valence-Energy.

Εξετάστηκαν δύο μέθοδοι υπολογισμού απόστασης από τα κέντρα βάρους, η Euclidean και η Mahalanobis απόσταση. Τα αποτελέσματα παρουσιάζονται στον Πίνακα 6.10.

Πίνακας 6.10: Απόδοση χαρτογράφησης από προβλεπόμενες διαστάσεις σε βασικά συναισθήματα

Μέθοδος απόστασης	Dominant emotion accuracy	Profile MAE	Cosine similarity	Mean alpha
Euclidean	0.60	0.069	0.924	0.25
Mahalanobis	0.51	0.074	0.906	0.75

Η μέθοδος με Euclidean απόσταση παρουσίασε καλύτερη συνολική απόδοση σε σχέση με τη Mahalanobis απόσταση. Συγκεκριμένα, πέτυχε υψηλότερη ακρίβεια ως προς την ανάκτηση του κυρίαρχου συναισθήματος (0.60 έναντι 0.51), χαμηλότερο profile MAE (0.069 έναντι 0.074) και υψηλότερη μέση cosine similarity (0.924 έναντι 0.906).

Σε επίπεδο κυρίαρχου συναισθήματος, το συναίσθημα με τη μέγιστη τιμή στο πραγματικό και στο προβλεπόμενο συναισθηματικό προφίλ θεωρήθηκε ως η αντίστοιχη κλάση, και υπολογίστηκαν επίσης οι μετρικές accuracy, precision, recall, και F1-score. Τα αποτελέσματα παρουσιάζονται στον Πίνακα 6.11. Παρατηρείται ότι η Euclidean μέθοδος παρουσιάζει υψηλότερες τιμές σε όλες τις μετρικές. Ειδικότερα, σε επίπεδο στόχου, η Euclidean χαρτογράφηση εμφάνισε καλύτερη ανάκτηση (recall) για τα Anger (0.80), Sad (0.70) και Happy (0.63), ενώ χαμηλότερη απόδοση παρατηρήθηκε για τα Fear (0.48) και Tender (0.45). Το αποτέλεσμα αυτό δείχνει ότι ορισμένες συναισθηματικές κατηγορίες αντιστοιχίζονται πιο καθαρά σε περιοχές του διαστασιακού χώρου, ενώ άλλες εμφανίζουν μεγαλύτερη επικάλυψη.

Πίνακας 6.11: Απόδοση χαρτογράφησης ως προς το κυρίαρχο συναίσθημα

Μέθοδος απόστασης	Accuracy	Macro precision	Macro recall	Macro F1-score
Euclidean	0.60	0.59	0.61	0.59
Mahalanobis	0.51	0.50	0.51	0.49

Το γεγονός αυτό επιβεβαιώνεται επίσης από τον Πίνακα 6.12 όπου παρουσιάζεται ο πίνακας σύγχυσης (confusion matrix) για τη χαρτογράφηση με Euclidean απόσταση. Παρατηρείται ότι τα Anger, Happy και Sad ανακτώνται με σχετικά καλύτερη συνέπεια, ενώ περισσότερες συγχύσεις εμφανίζονται μεταξύ Fear, Sad και Tender. Ιδιαίτερα, αρκετά αποσπάσματα με πραγματικό κυρίαρχο συναίσθημα Tender ταξινομήθηκαν ως Sad, γεγονός που υποδηλώνει

επικάλυψη των δύο συναισθηματικών κατηγοριών στον διαστασιακό χώρο Valence-Energy-Tension.

Η παρατήρηση αυτή είναι συμβατή με τα ευρήματα των Eerola & Vuoskoski (2011), από τη μελέτη των οποίων προέρχεται το σύνολο δεδομένων της παρούσας ΔΕ. Συγκεκριμένα, αναφέρεται ότι τα πιο χαρακτηριστικά παραδείγματα βασικών συναισθημάτων αναγνωρίζονταν με συνέπεια από τους ακροατές, ενώ τα λιγότερο χαρακτηριστικά παραδείγματα εμφάνιζαν συγχύσεις με γειτονικές συναισθηματικές κατηγορίες. Ειδικότερα, παρατηρήθηκαν συγχύσεις μεταξύ Sad και Tender, μεταξύ Tender, Happy και Sad, καθώς και μεταξύ Anger, Fear και Sad. Επομένως, οι συγχύσεις που εμφανίζονται στην παρούσα χαρτογράφηση, ιδιαίτερα η τάση αποσπασμάτων με κυρίαρχο Tender να ταξινομούνται ως Sad και η δυσκολότερη ανάκτηση του Fear, μπορούν να ερμηνευθούν εν μέρει ως απόρροια της εγγενούς επικάλυψης ορισμένων συναισθηματικών κατηγοριών στο συγκεκριμένο μουσικό υλικό.

Πίνακας 6.12: Πίνακας σύγκρισης για τη χαρτογράφηση με Euclidean απόσταση

Πραγματικό / Προβλεπόμενο	Happy	Sad	Anger	Fear	Tender
Happy	15	2	3	3	1
Sad	1	23	0	3	6
Anger	0	0	8	2	0
Fear	3	5	4	11	0
Tender	3	8	0	0	9

Τα αποτελέσματα της χαρτογράφησης δείχνουν ότι ο διαστασιακός χώρος μπορεί να αξιοποιηθεί ως ενδιάμεσο επίπεδο για την εκτίμηση των βασικών συναισθημάτων. Ωστόσο, η απόδοση μειώνεται όταν η χαρτογράφηση βασίζεται σε προβλεπόμενες και όχι σε πραγματικές συναισθηματικές διαστάσεις, καθώς το σφάλμα των μοντέλων πρόβλεψης των διαστάσεων μεταφέρεται στο επόμενο στάδιο. Για τον λόγο αυτό, στην επόμενη Ενότητα συγκρίνεται η έμμεση αυτή προσέγγιση με την άμεση πρόβλεψη των βασικών συναισθημάτων από τα αντικειμενικά ηχητικά χαρακτηριστικά.

6.7 Σύγκριση Άμεσης και Έμμεσης Πρόβλεψης Βασικών Συναισθημάτων

Στο τελικό στάδιο της ανάλυσης συγκρίθηκαν δύο εναλλακτικές προσεγγίσεις για την πρόβλεψη του συναισθηματικού προφίλ βασικών συναισθημάτων. Η πρώτη προσέγγιση

αφορούσε την άμεση πρόβλεψη των βασικών συναισθημάτων από τα αντικειμενικά ηχητικά χαρακτηριστικά, με χρήση του τελικού κοινού μειωμένου μοντέλου REDUCED_EXTENDED_COMMON_N $\geq$ 4. Η δεύτερη προσέγγιση αφορούσε την έμμεση πρόβλεψη, κατά την οποία αρχικά προβλέπονταν οι συναισθηματικές διαστάσεις Valence, Energy και Tension και στη συνέχεια οι προβλεπόμενες διαστάσεις χαρτογραφούνταν σε προφίλ βασικών συναισθημάτων μέσω της Euclidean απόστασης που παρουσίασε την καλύτερη απόδοση στην προηγούμενη Ενότητα.

Η σύγκριση πραγματοποιήθηκε με βάση την ακρίβεια ανάκτησης του κυρίαρχου συναισθήματος, το profile MAE και τη μέση cosine similarity μεταξύ πραγματικού και προβλεπόμενου συναισθηματικού προφίλ. Τα αποτελέσματα παρουσιάζονται στον Πίνακα 6.13.

Πίνακας 6.13: Σύγκριση άμεσης και έμμεσης πρόβλεψης βασικών συναισθημάτων

Προσέγγιση	Dominant emotion accuracy	Profile MAE	Cosine similarity
Άμεση πρόβλεψη βασικών συναισθημάτων	0.67	0.065	0.928
Έμμεση πρόβλεψη μέσω συναισθηματικών διαστάσεων	0.60	0.069	0.924

Τα αποτελέσματα δείχνουν ότι η άμεση πρόβλεψη των βασικών συναισθημάτων παρουσίασε καλύτερη συνολική απόδοση. Συγκεκριμένα, πέτυχε υψηλότερη ακρίβεια ως προς το κυρίαρχο συναίσθημα (0.67 έναντι 0.60), χαμηλότερο profile MAE (0.065 έναντι 0.069) και ελαφρώς υψηλότερη μέση cosine similarity (0.928 έναντι 0.924). Ωστόσο, η διαφορά μεταξύ των δύο προσεγγίσεων ήταν σχετικά μικρή, ιδιαίτερα ως προς τις μετρικές profile MAE και cosine similarity που αξιολογούν το πλήρες συναισθηματικό προφίλ. Το αποτέλεσμα αυτό δείχνει ότι η έμμεση μέθοδος μπορεί να αναπαράγει σε σημαντικό βαθμό τη συνολική μορφή του συναισθηματικού προφίλ, παρά το γεγονός ότι επηρεάζεται από τη μεταφορά σφάλματος από το στάδιο πρόβλεψης των διαστάσεων.

Σε επίπεδο επιμέρους βασικών συναισθημάτων, η άμεση πρόβλεψη παρουσίασε χαμηλότερο profile MAE για τα Anger, Happy και Sad, ενώ η έμμεση πρόβλεψη παρουσίασε ελαφρώς χαμηλότερο σφάλμα για τα Fear και Tender. Συγκεκριμένα, για την άμεση πρόβλεψη οι τιμές profile MAE ήταν 0.079 για το Happy, 0.063 για το Sad, 0.045 για το Anger, 0.077 για το Fear και 0.058 για το Tender. Για την έμμεση προσέγγιση οι

αντίστοιχες τιμές ήταν 0.094, 0.070, 0.052, 0.075 και 0.053. Το εύρημα αυτό υποδηλώνει ότι η υπεροχή της άμεσης πρόβλεψης δεν είναι ομοιόμορφη για όλα τα συναισθήματα.

Συνολικά, η άμεση πρόβλεψη βασικών συναισθημάτων αποδείχτηκε η αποδοτικότερη προσέγγιση για την τελική εκτίμηση του συναισθηματικού προφίλ. Η έμμεση προσέγγιση μέσω των συναισθηματικών διαστάσεων αποδείχθηκε επίσης χρήσιμη, καθώς προσφέρει ένα ενδιάμεσο επίπεδο σύνδεσης μεταξύ αντικειμενικών ηχητικών χαρακτηριστικών και βασικών συναισθημάτων.

6.8 Σύνοψη και Ερμηνεία Αποτελεσμάτων

Τα αποτελέσματα της πειραματικής διαδικασίας έδειξαν ότι τα αντικειμενικά ηχητικά χαρακτηριστικά μπορούν να αξιοποιηθούν για την πρόβλεψη συναισθηματικών αξιολογήσεων μουσικών αποσπασμάτων, τόσο σε επίπεδο συναισθηματικών διαστάσεων όσο και σε επίπεδο βασικών συναισθημάτων. Ο GradientBoostingRegressor παρουσίασε καλύτερη απόδοση σε σχέση με τον RandomForestRegressor και αποτέλεσε τον βασικό αλγόριθμο των τελικών πειραμάτων.

Η χρήση του διευρυμένου συνόλου των 53 χαρακτηριστικών, σε συνδυασμό με τη διαδικασία επιλογής χαρακτηριστικών μέσω SHAP και ανάλυσης σταθερότητας, οδήγησε σε βελτιωμένη απόδοση των τελικών μοντέλων. Η μείωση των χαρακτηριστικών οδήγησε σε απλοποίηση των μοντέλων και συνέβαλε στη βελτίωση της προβλεπτικής τους ικανότητας. Το γεγονός αυτό επιβεβαιώθηκε επίσης από τους στατιστικούς ελέγχους, όπου η χρήση κοινού μειωμένου συνόλου χαρακτηριστικών οδήγησε σε στατιστικά σημαντική μείωση του συνολικού σφάλματος τόσο για τις συναισθηματικές διαστάσεις όσο και για τα βασικά συναισθήματα.

Η ανάλυση ερμηνευσιμότητας ανέδειξε τη σημασία κυρίως τονικών/αρμονικών και ηχοχρωματικών/φασματικών χαρακτηριστικών. Χαρακτηριστικά όπως τα `hpcp_entropy_mean`, `key_strength`, `majorness`, `pitch_confidence_mean`, `spectral_flux_std` και διάφοροι συντελεστές MFCC εμφανίστηκαν συχνά στα τελικά μειωμένα σύνολα, υποδηλώνοντας ότι η τονική οργάνωση, η φασματική μεταβλητότητα και η ηχοχρωματική δομή του μουσικού αποσπάσματος συνδέονται σημαντικά με τις συναισθηματικές αξιολογήσεις. Αντίθετα, τα χαρακτηριστικά έντασης και δυναμικής είχαν περιορισμένη παρουσία στα τελικά επιλεγμένα σύνολα.

Τα αποτελέσματα έδειξαν επίσης ότι οι συναισθηματικές διαστάσεις και τα βασικά συναισθήματα συνδέονται μεταξύ τους σε σημαντικό βαθμό. Η χαρτογράφηση από πραγματικές διαστάσεις σε βασικά συναισθήματα παρουσίασε υψηλή απόδοση, ενώ η χαρτογράφηση από προβλεπόμενες διαστάσεις εμφάνισε επίσης ικανοποιητική απόδοση. Η άμεση πρόβλεψη βασικών συναισθημάτων από τα ηχητικά χαρακτηριστικά αποδείχθηκε συνολικά αποδοτικότερη από την έμμεση προσέγγιση μέσω διαστάσεων. Ωστόσο, η έμμεση μέθοδος παρουσίασε ανταγωνιστική απόδοση ως προς την πρόβλεψη του πλήρους συναισθηματικού προφίλ.

Συμπερασματικά, τα αποτελέσματα υποστηρίζουν ότι η πρόβλεψη μουσικού συναισθήματος μπορεί να ωφεληθεί από τον συνδυασμό αντικειμενικών ηχητικών χαρακτηριστικών, μη γραμμικών μοντέλων παλινδρόμησης και τεχνικών επιλογής χαρακτηριστικών. Παράλληλα, η συνδυαστική εξέταση διαστασιακών και κατηγορικών μορφών συναισθηματική αναπαράστασης επιτρέπει μια πιο ολοκληρωμένη προσέγγιση της σχέσης μεταξύ μουσικού σήματος και ανθρώπινης συναισθηματικής αξιολόγησης.

7. Σημασιολογική Ερμηνεία και Αναπαράσταση Συναισθηματικού Προφίλ Μουσικών Αποσπασμάτων

Μετά την ολοκλήρωση της πειραματικής διαδικασίας και την αξιολόγηση των μοντέλων πρόβλεψης, το παρόν Κεφάλαιο εστιάζει στη σημασιολογική ερμηνεία και αναπαράσταση των αποτελεσμάτων. Συγκεκριμένα, εξετάζεται πως οι προβλεπόμενες συναισθηματικές τιμές μπορούν να παρουσιαστούν με έναν πιο κατανοητό και ερμηνεύσιμο τρόπο.

7.1 Στόχος της Σημασιολογικής Ερμηνείας

Στόχος της σημασιολογικής ερμηνείας είναι η μετάβαση από τα αριθμητικά αποτελέσματα των μοντέλων σε ένα πιο άμεσα αντιληπτό συναισθηματικό προφίλ κάθε μουσικού αποσπάσματος. Για παράδειγμα, αντί ένα απόσπασμα να περιγράφεται μόνο από επιμέρους τιμές για τα Happy, Sad, Anger, Fear και Tender, μπορεί να αποδοθεί ως ένα συνολικό προφίλ στο οποίο αναδεικνύονται το κυρίαρχο συναίσθημα, οι δευτερεύουσες συναισθηματικές τάσεις και ο βαθμός συναισθηματικής ασάφειας.

Η προσέγγιση αυτή επιτρέπει την αξιοποίηση των προβλέψεων όχι μόνο ως αριθμητικών αποτελεσμάτων, αλλά και ως περιγραφών που μπορούν να υποστηρίξουν διαδικασίες αναζήτησης, ταξινόμησης και οργάνωσης μουσικών αποσπασμάτων με βάση το συναισθηματικό τους περιεχόμενο. Επίσης, η αναπαράσταση του πλήρους συναισθηματικού προφίλ επιτρέπει την αποτύπωση περιπτώσεων όπου ένα μουσικό απόσπασμα εκφράζει περισσότερες από μία συναισθηματικές ποιότητες.

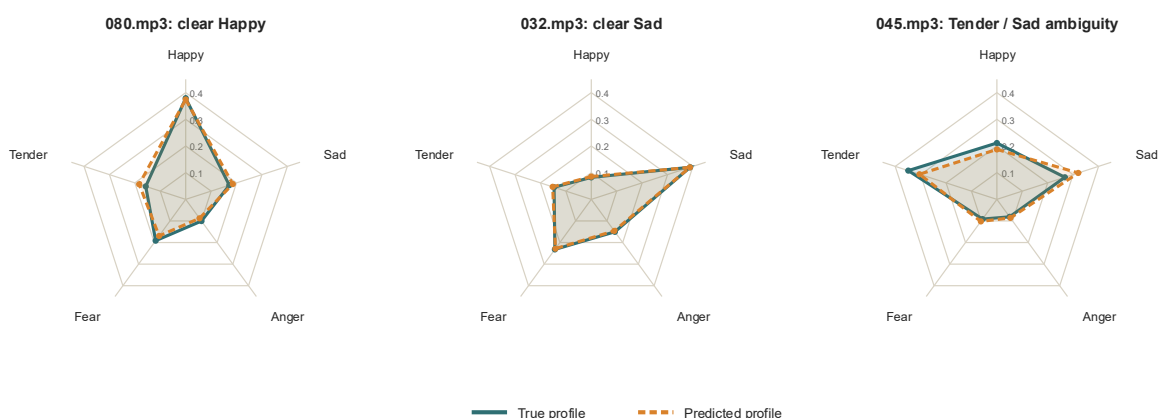
Στο πλαίσιο της παρούσας ΔΕ, η σημασιολογική ερμηνεία αποσκοπεί σε έναν άμεσα κατανοητό και αντιληπτό τρόπο παρουσίασης των αποτελεσμάτων που προέκυψαν από τα τελικά μοντέλα.

7.2 Αναπαράσταση Συναισθηματικού Προφίλ

Η αναπαράσταση του συναισθηματικού προφίλ βασίστηκε στις προβλεπόμενες τιμές των πέντε βασικών συναισθημάτων: Happy, Sad, Anger, Fear και Tender. Για κάθε μουσικό απόσπασμα, οι τιμές αυτές κανονικοποιήθηκαν ώστε να προκύπτει ένα συγκρίσιμο προφίλ με άθροισμα ίσο με 1. Με αυτόν τον τρόπο, κάθε τιμή μπορεί να ερμηνευθεί ως σχετική συμμετοχή του αντίστοιχου συναισθήματος στο συνολικό συναισθηματικό προφίλ του αποσπάσματος.

Η χρήση πλήρους συναισθηματικού προφίλ επιτρέπει την αποτύπωση τόσο του κυρίαρχου συναισθήματος όσο και των δευτερευουσών συναισθηματικών τάσεων. Το κυρίαρχο συναίσθημα ορίζεται ως το συναίσθημα με τη μεγαλύτερη τιμή στο προφίλ, ενώ οι υπόλοιπες τιμές αποτυπώνουν τον βαθμό στον οποίο συμμετέχουν και άλλες συναισθηματικές ποιότητες. Για παράδειγμα, ένα απόσπασμα μπορεί να έχει κυρίαρχο συναίσθημα Sad, αλλά ταυτόχρονα να παρουσιάζει υψηλή τιμή Tender, υποδηλώνοντας έναν πιο τρυφερό και μελαγχολικό χαρακτήρα.

Για την οπτική παρουσίαση των συναισθηματικών προφίλ μπορούν να χρησιμοποιηθούν γραφήματα ραντάρ ή ράβδων. Οι μορφές αυτές επιτρέπουν τη γρήγορη σύγκριση μεταξύ των συναισθημάτων και την αναγνώριση τόσο της κυρίαρχης όσο και των δευτερευουσών συναισθηματικών τάσεων. Στην Εικόνα 7.1 παρουσιάζονται ενδεικτικά πραγματικά και προβλεπόμενα συναισθηματικά προφίλ μουσικών αποσπασμάτων σε γραφήματα ραντάρ. Με συνεχόμενη γραμμή παρουσιάζεται το πραγματικό προφίλ και με διακεκομμένη γραμμή το προβλεπόμενο προφίλ από το τελικό μοντέλο άμεσης πρόβλεψης βασικών συναισθημάτων. Τα παραδείγματα αφορούν, από αριστερά προς τα δεξιά, ένα καθαρό παράδειγμα Happy (080.mp3), με πολύ κοντινό πραγματικό και προβλεπόμενο προφίλ, ένα καθαρό παράδειγμα Sad (032.mp3), επίσης με πολύ καλή αντιστοίχιση, και ένα παράδειγμα ασάφειας μεταξύ Tender και Sad (045.mp3), όπου το κυρίαρχο συναίσθημα δεν προβλέφθηκε σωστά, αλλά η συνολική μορφή του προβλεπόμενου προφίλ παραμένει κοντά στο πραγματικό.



Εικόνα 7.1: Ενδεικτικά συναισθηματικά προφίλ μουσικών αποσπασμάτων.

7.3 Σημασιολογική Ερμηνεία μέσω Χαρτογράφησης Διαστάσεων

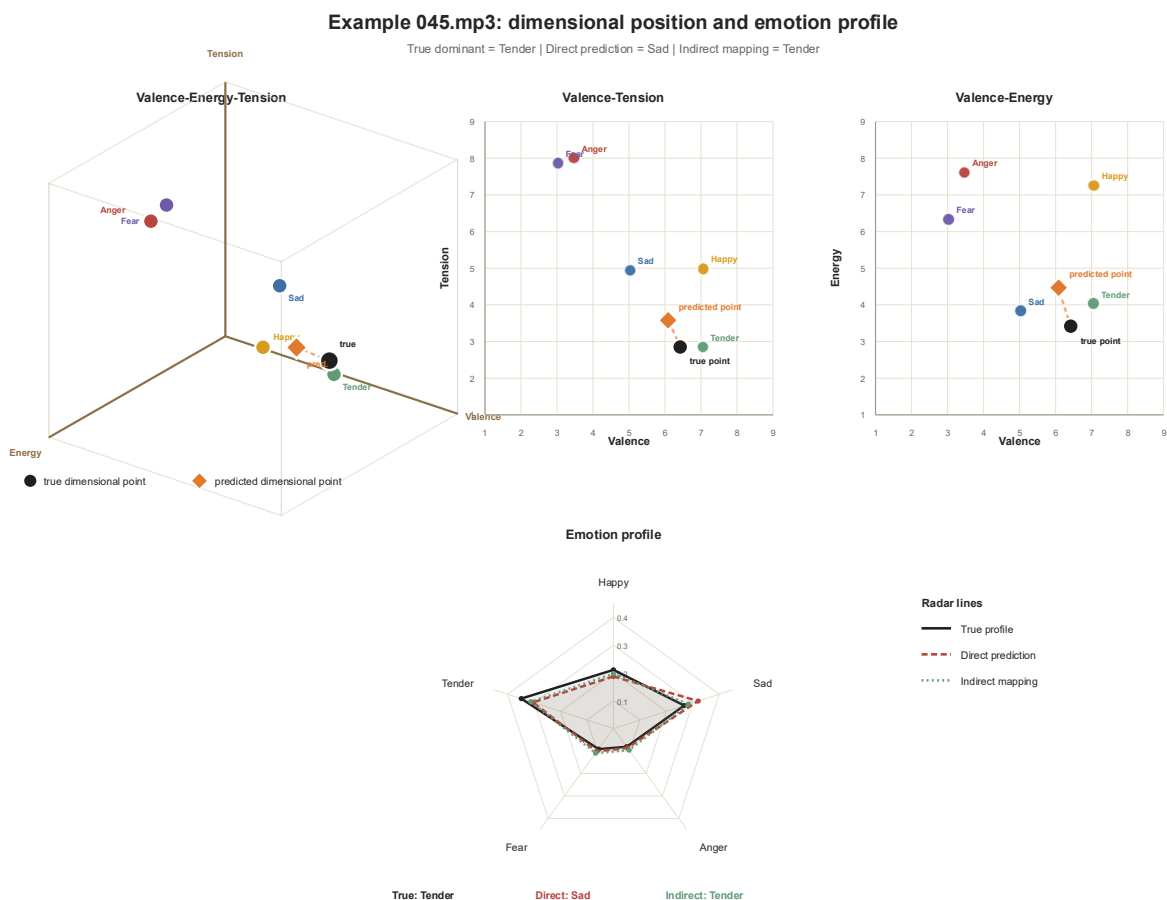
Η διαδικασία χαρτογράφησης από τις συναισθηματικές διαστάσεις προς τα βασικά συναισθήματα μπορεί επίσης να αξιοποιηθεί ως εργαλείο σημασιολογικής ερμηνείας, συμπληρωματικά με την άμεση αναπαράσταση του συναισθηματικού προφίλ. Η θέση ενός μουσικού αποσπάσματος στον χώρο Valence-Energy-Tension παρέχει πληροφορία για τον γενικό συναισθηματικό του χαρακτήρα, ενώ η αντιστοίχισή του με τα κέντρα βάρους των βασικών συναισθημάτων επιτρέπει τη μετάφραση αυτής της θέσης σε ένα πιο άμεσα ερμηνεύσιμο προφίλ βασικών συναισθημάτων.

Για παράδειγμα, αποσπάσματα με χαμηλή Valence και υψηλά Energy και Tension βρίσκονται πιο κοντά στα κέντρα βάρους των Anger και Fear, ενώ αποσπάσματα με υψηλή Valence και χαμηλότερη Tension προσεγγίζουν περισσότερο τα Happy και Tender. Επίσης, η γειτνίαση των Sad και Tender στον διαστασιακό χώρο μπορεί να εξηγήσει περιπτώσεις όπου ένα μουσικό απόσπασμα έχει σύνθετο συναισθηματικό προφίλ.

Στην Εικόνα 7.2 παρουσιάζεται ενδεικτικά η περίπτωση του μουσικού αποσπάσματος 045.mp3. Το απόσπασμα αυτό έχει πραγματικό κυρίαρχο συναίσθημα Tender. Στον διαστασιακό χώρο, η πραγματική και η προβλεπόμενη θέση του βρίσκονται κοντά στην περιοχή των Tender και Sad. Η άμεση πρόβλεψη του τελικού μοντέλου απέδωσε ως κυρίαρχο συναίσθημα το Sad, ενώ η έμμεση χαρτογράφηση μέσω των προβλεπόμενων διαστάσεων απέδωσε ως κυρίαρχο συναίσθημα το Tender.

Το αντίστοιχο γράφημα ραντάρ δείχνει ότι και οι δύο προβλέψεις αποτυπώνουν σε σημαντικό βαθμό τη συνολική μορφή του πραγματικού συναισθηματικού προφίλ, παρά τη διαφοροποίηση ως προς το κυρίαρχο συναίσθημα. Και στις δύο περιπτώσεις διατηρούνται οι υψηλές τιμές των Sad και Tender, ενώ οι υπόλοιπες τιμές παραμένουν χαμηλότερες. Το παράδειγμα αυτό αναδεικνύει τη χρησιμότητα της αναπαράστασης προφίλ για τη σημασιολογική ερμηνεία του συναισθηματικού περιεχομένου.

Με τον τρόπο αυτό, η χαρτογράφηση μπορεί να χρησιμοποιηθεί για την ερμηνεία της σχέσης μεταξύ των διαστασιακών και κατηγορικών μορφών συναισθηματικής αναπαράστασης, καθώς και για την παραγωγή προβλέψεων, λειτουργώντας συμπληρωματικά με την άμεση πρόβλεψη.



Εικόνα 7.2: Ενδεικτική σημασιολογική αναπαράσταση μουσικού αποσπάσματος μέσω της θέσης του στον διαστασιακό χώρο και της σύγκρισης πραγματικού, άμεσα προβλεπόμενου και έμμεσα χαρτογραφημένου συναισθηματικού προφίλ

7.4 Σημασιολογικές Περιγραφές Μουσικών Αποσπασμάτων

Η αναπαράσταση του συναισθηματικού προφίλ μπορεί να αξιοποιηθεί για τη δημιουργία σύντομων σημασιολογικών περιγραφών των μουσικών αποσπασμάτων. Τα αριθμητικά αποτελέσματα των μοντέλων ενδεχομένως να μην είναι πάντα άμεσα ερμηνεύσιμα από έναν τελικό χρήστη. Αντίθετα, μια λεκτική περιγραφή του προφίλ μπορεί να αποδώσει πιο άμεσα τον συναισθηματικό χαρακτήρα ενός αποσπάσματος.

Οι περιγραφές αυτές μπορούν να βασιστούν σε τρία συμπληρωματικά στοιχεία: το κυρίαρχο συναίσθημα, την κατανομή των τιμών στο συναισθηματικό προφίλ και τη θέση του αποσπάσματος στον διαστασιακό χώρο Valence-Energy-Tension. Για παράδειγμα, ένα απόσπασμα με υψηλές τιμές Happy και Valence και χαμηλές τιμές Sad και Fear μπορεί να περιγραφεί ως «κυρίως χαρούμενο και θετικό». Αντίστοιχα, ένα προφίλ στο οποίο συνυπάρχουν υψηλές τιμές Sad και Tender μπορεί να περιγραφεί ως «μελαγχολικό με

τρυφερή συναισθηματική χροιά». Σε πιο σύνθετες περιπτώσεις κατανομής του προφίλ, η περιγραφή μπορεί να αποδώσει τον μικτό χαρακτήρα του αποσπάσματος.

Στον Πίνακα 7.1 παρουσιάζονται ενδεικτικά παραδείγματα σημασιολογικών περιγραφών με βάση τα προβλεπόμενα συναισθηματικά προφίλ επιλεγμένων μουσικών αποσπασμάτων.

Πίνακας 7.1: Ενδεικτικές σημασιολογικές περιγραφές μουσικών αποσπασμάτων

Απόσπασμα	Βασικά στοιχεία προφίλ	Διαστασιακή πληροφορία	Ενδεικτική σημασιολογική περιγραφή
080.mp3	Κυρίαρχο Happy, με χαμηλότερες δευτερεύουσες τιμές	Μέτρια-υψηλή Valence και σχετικά υψηλή Energy	Χαρούμενο και θετικό απόσπασμα με ήπιες δευτερεύουσες συναισθηματικές αποχρώσεις
032.mp3	Κυρίαρχο Sad, με μέτρια τιμή Fear	Μέτρια Tension και χαμηλότερη Valence	Μελαγχολικό απόσπασμα με στοιχεία έντασης και φόβου
045.mp3	Συνύπαρξη Tender και Sad	Υψηλή Valence, μέτρια Energy και χαμηλότερη Tension	Συναισθηματικό προφίλ με συνύπαρξη τρυφερότητας και μελαγχολίας
001.mp3	Υψηλές τιμές Anger και Fear	Χαμηλή Valence, υψηλή Energy και υψηλή Tension	Έντονο αρνητικό απόσπασμα με στοιχεία θυμού και φόβου

Η συμπληρωματική παρουσίαση των αριθμητικών προβλέψεων, της αναπαράστασης στους δύο συναισθηματικούς χώρους και της σημασιολογικής περιγραφής επιτρέπει μια συνεκτική ερμηνεία του συναισθηματικού χαρακτήρα ενός μουσικού αποσπάσματος και μπορεί να αξιοποιηθεί σε εφαρμογές ανάκτησης μουσικής πληροφορίας, συναισθηματικής αναζήτησης ή οργάνωσης μουσικών συλλογών.

7.5 Σύνοψη και Περιορισμοί της Σημασιολογικής Αναπαράστασης

Στο παρόν Κεφάλαιο παρουσιάστηκε ένας τρόπος σημασιολογικής ερμηνείας των αποτελεσμάτων των μοντέλων πρόβλεψης, με στόχο τη μετάβαση από τις αριθμητικές τιμές σε πιο εύκολα αντιληπτές και κατανοητές περιγραφές του συναισθηματικού χαρακτήρα των μουσικών αποσπασμάτων. Η αναπαράσταση του πλήρους συναισθηματικού προφίλ επιτρέπει την αποτύπωση τόσο του κυρίαρχου συναισθήματος όσο και των δευτερευουσών συναισθηματικών τάσεων, ενώ η χαρτογράφηση μεταξύ διαστασιακού και κατηγορικού χώρου προσφέρει ένα επιπλέον επίπεδο ερμηνείας.

Η σημασιολογική αναπαράσταση συνοδεύεται ωστόσο από ορισμένους περιορισμούς. Οι λεκτικές περιγραφές που προκύπτουν από τα συναισθηματικά προφίλ έχουν ερμηνευτικό

χαρακτήρα και δεν αποτελούν απόλυτες περιγραφές των αποσπασμάτων. Επίσης, εξαρτώνται από την ακρίβεια των μοντέλων πρόβλεψης και από τον βαθμό στον οποίο τα επιλεγμένα βασικά συναισθήματα επαρκούν για την περιγραφή του κάθε μουσικού αποσπάσματος.

Συνεπώς, η προτεινόμενη σημασιολογική αναπαράσταση μπορεί να θεωρηθεί συμπληρωματικό εργαλείο παρουσίασης και κατανόησης των αποτελεσμάτων. Με τον τρόπο αυτό, μπορεί να προσφέρει ένα χρήσιμο επίπεδο ανάλυσης και κατανόησης του συναισθηματικού χαρακτήρα μουσικών αποσπασμάτων, λειτουργώντας παράλληλα με την ανθρώπινη ακρόαση και αξιολόγηση.

8. Συζήτηση, Συμπεράσματα και Μελλοντική Έρευνα

8.1 Συζήτηση

Τα αποτελέσματα της παρούσας ΔΕ δείχνουν ότι τα αντικειμενικά ηχητικά χαρακτηριστικά μπορούν να αξιοποιηθούν για την πρόβλεψη συναισθηματικών αξιολογήσεων μουσικών αποσπασμάτων, τόσο σε επίπεδο συναισθηματικών διαστάσεων όσο και σε επίπεδο βασικών συναισθημάτων. Η συνολική απόδοση των μοντέλων υποδηλώνει ότι ιδιότητες του μουσικού σήματος, όπως η τονικότητα, η φασματική μεταβλητότητα και η ηχοχρωματική δομή, περιέχουν πληροφορία σχετική με τον τρόπο με τον οποίο οι ακροατές αποδίδουν συναίσθημα στη μουσική. Το εύρημα αυτό είναι συμβατό με τη βιβλιογραφία της MER, σύμφωνα με την οποία η πρόβλεψη μουσικού συναισθήματος μπορεί να βασιστεί σε συνδυασμό διαφορετικών ηχητικών χαρακτηριστικών.

Η σύγκριση των μοντέλων έδειξε ότι ο GradientBoostingRegressor παρουσίασε καλύτερη συνολική απόδοση σε σχέση με τον RandomForestRegressor και αποτέλεσε τη βασική επιλογή για τα τελικά πειράματα. Η βελτίωση αυτή μπορεί να αποδοθεί στην ικανότητα του GradientBoostingRegressor να συνδυάζει διαδοχικά απλά μοντέλα και να διορθώνει σταδιακά τα σφάλματα των προηγούμενων προβλέψεων. Παρότι δεν πραγματοποιήθηκε εκτεταμένη υπερπαραμετρική βελτιστοποίηση, τα αποτελέσματα δείχνουν ότι η συγκεκριμένη προσέγγιση ήταν κατάλληλη για το μέγεθος και τη φύση του διαθέσιμου συνόλου δεδομένων.

Ιδιαίτερο ενδιαφέρον παρουσιάζει η επίδραση της επιλογής χαρακτηριστικών. Τα μειωμένα σύνολα χαρακτηριστικών οδήγησαν σε βελτιωμένη απόδοση σε σχέση με το πλήρες διευρυμένο σύνολο, γεγονός που δείχνει ότι η προσθήκη περισσότερων χαρακτηριστικών δεν οδηγεί απαραίτητα σε καλύτερα μοντέλα. Αντίθετα, η παρουσία πολλών χαρακτηριστικών μπορεί να αυξήσει την πολυπλοκότητα και να επηρεάσει αρνητικά τη γενίκευση των μοντέλων. Η χρήση της μεθόδου SHAP σε συνδυασμό με ανάλυση σταθερότητας επέτρεψε τη διαμόρφωση μειωμένων συνόλων που ήταν όχι μόνο αποδοτικότερα αλλά και πιο ερμηνεύσιμα.

Η ανάλυση των επιλεγμένων χαρακτηριστικών ανέδειξε τη σημασία κυρίως τονικών/αρμονικών και ηχοχρωματικών/φασματικών χαρακτηριστικών. Χαρακτηριστικά όπως το `hpcp_entropy_mean`, το `key_strength`, το `majorness`, το `pitch_confidence_mean`, το `spectral_flux_std` και διάφοροι συντελεστές MFCC εμφανίστηκαν συχνά στα τελικά

μειωμένα σύνολα. Το εύρημα αυτό υποδηλώνει ότι η τονική οργάνωση, η φασματική μεταβλητότητα και η ηχοχρωματική δομή συνδέονται σημαντικά με τις συναισθηματικές αξιολογήσεις. Αντίθετα, τα χαρακτηριστικά έντασης και δυναμικής είχαν περιορισμένη παρουσία στα τελικά επιλεγμένα σύνολα, γεγονός που μπορεί να οφείλεται είτε στη μικρότερη προβλεπτική τους συμβολή στο συγκεκριμένο σύνολο δεδομένων είτε σε πιθανή επικάλυψη της πληροφορίας τους με άλλα χαρακτηριστικά.

Σε σχέση με τις δύο μορφές συναισθηματικής αναπαράστασης, τα αποτελέσματα έδειξαν ότι τόσο οι συναισθηματικές διαστάσεις όσο και τα βασικά συναισθήματα μπορούν να προβλεφθούν σε ικανοποιητικό βαθμό από τα ηχητικά χαρακτηριστικά. Η χαρτογράφηση από πραγματικές συναισθηματικές διαστάσεις σε βασικά συναισθήματα παρουσίασε υψηλή απόδοση, ενισχύοντας την άποψη ότι οι δύο χώροι συνδέονται ουσιαστικά. Ωστόσο, όταν η χαρτογράφηση βασίστηκε σε προβλεπόμενες διαστάσεις, η απόδοση μειώθηκε, καθώς το σφάλμα από το πρώτο στάδιο πρόβλεψης μεταφέρθηκε στο τελικό συναισθηματικό προφίλ. Παρά τη μείωση αυτή, η έμμεση προσέγγιση μέσω των διαστάσεων παρουσίασε ανταγωνιστική συμπεριφορά ως προς την πρόβλεψη του πλήρους συναισθηματικού προφίλ.

Η σύγκριση άμεσης και έμμεσης πρόβλεψης έδειξε ότι η άμεση πρόβλεψη βασικών συναισθημάτων από τα ηχητικά χαρακτηριστικά ήταν συνολικά αποδοτικότερη. Ωστόσο, η έμμεση προσέγγιση μέσω των συναισθηματικών διαστάσεων παραμένει χρήσιμη, καθώς προσφέρει ένα ενδιάμεσο επίπεδο ερμηνείας και επιτρέπει τη σύνδεση μεταξύ διαστασιακής και κατηγορικής αναπαράστασης του συναισθήματος. Το εύρημα αυτό είναι σημαντικό, καθώς δείχνει ότι οι δύο μορφές αναπαράστασης μπορούν να θεωρηθούν συμπληρωματικές.

Παράλληλα, η σημασιολογική αναπαράσταση των αποτελεσμάτων μέσω συναισθηματικών προφίλ ανέδειξε τη χρησιμότητα μιας πιο ερμηνεύσιμης παρουσίασης των προβλέψεων. Αντί οι έξοδοι των μοντέλων να παρουσιάζονται μόνο ως αριθμητικές τιμές, μπορούν να οργανωθούν σε προφίλ που αποτυπώνουν το κυρίαρχο συναίσθημα, τις δευτερεύουσες συναισθηματικές τάσεις και πιθανές περιπτώσεις συναισθηματικής ασάφειας. Με τον τρόπο αυτό, η ΔΕ επιχειρεί να γεφυρώσει σε ένα βασικό επίπεδο το σημασιολογικό χάσμα μεταξύ υπολογιστικής πρόβλεψης και ανθρώπινης ερμηνείας του μουσικού συναισθήματος.

Ωστόσο, ορισμένοι περιορισμοί της έρευνας πρέπει να ληφθούν υπόψη. Το διαθέσιμο σύνολο δεδομένων είναι σχετικά μικρό και αφορά κινηματογραφική μουσική. Το γεγονός

αυτό περιορίζει τη δυνατότητα γενίκευσης των συμπερασμάτων σε άλλα είδη μουσικής. Επιπλέον, οι συναισθηματικές αξιολογήσεις προέρχονται από ανθρώπινες κρίσεις, οι οποίες χαρακτηρίζονται από υποκειμενικότητα και ενδεχόμενη ασάφεια, ιδιαίτερα σε περιπτώσεις μικτών συναισθηματικών προφίλ. Τέλος, η χρήση συγκεκριμένων μοντέλων και υπερπαραμέτρων, χωρίς εκτεταμένη υπερπαραμετρική βελτιστοποίηση, σημαίνει ότι τα αποτελέσματα αποτυπώνουν κυρίως τη συγκριτική αξία των χαρακτηριστικών και των προσεγγίσεων που εξετάστηκαν, και όχι απαραίτητα τη μέγιστη δυνατή απόδοση.

8.2 Συμπεράσματα

Η παρούσα ΔΕ είχε ως βασικό στόχο τη διερεύνηση της δυνατότητας πρόβλεψης συναισθηματικών αξιολογήσεων μουσικών αποσπασμάτων από αντικειμενικά ηχητικά χαρακτηριστικά. Με βάση τα αποτελέσματα των πειραμάτων, προκύπτει ότι τα ηχητικά χαρακτηριστικά μπορούν να αξιοποιηθούν για την πρόβλεψη τόσο συναισθηματικών διαστάσεων όσο και βασικών συναισθημάτων, επιβεβαιώνοντας τη χρησιμότητά τους στο πλαίσιο της MER.

Ως προς τα μοντέλα πρόβλεψης, ο GradientBoostingRegressor παρουσίασε την καλύτερη συνολική απόδοση και αποτέλεσε το βασικό μοντέλο των τελικών πειραμάτων. Η εφαρμογή επιλογής χαρακτηριστικών μέσω SHAP και ανάλυσης σταθερότητας οδήγησε σε μειωμένα σύνολα χαρακτηριστικών με βελτιωμένη απόδοση, γεγονός που δείχνει ότι η συστηματική επιλογή χαρακτηριστικών μπορεί να είναι πιο αποδοτική από τη χρήση ενός μεγαλύτερου συνόλου χαρακτηριστικών.

Σε επίπεδο ερμηνείας, τα χαρακτηριστικά που αναδείχθηκαν ως σημαντικότερα αφορούν κυρίως τονικά/αρμονικά και ηχοχρωματικά/φασματικά χαρακτηριστικά. Το εύρημα αυτό δείχνει ότι η συναισθηματική πληροφορία στη μουσική δεν περιορίζεται στην ένταση και τον ρυθμό, αλλά συνδέεται και με πτυχές της τονικής οργάνωσης, της φασματικής μεταβολής και της ηχοχρωματικής δομής.

Η διερεύνηση της σχέσης μεταξύ συναισθηματικών χώρων έδειξε ότι οι διαστασιακές και οι κατηγορικές αναπαραστάσεις μπορούν να λειτουργήσουν συμπληρωματικά. Η άμεση πρόβλεψη βασικών συναισθημάτων παρουσίασε καλύτερη συνολική απόδοση, ενώ η έμμεση προσέγγιση μέσω χαρτογράφησης από τις συναισθηματικές διαστάσεις προσέφερε ένα χρήσιμο ενδιάμεσο επίπεδο ερμηνείας.

Συνολικά, η ΔΕ υποστηρίζει ότι ο συνδυασμός αντικειμενικών ηχητικών χαρακτηριστικών, μοντέλων μηχανικής μάθησης, τεχνικών επιλογής χαρακτηριστικών και σημασιολογικής αναπαράστασης μπορεί να αποτελέσει ένα χρήσιμο πλαίσιο για την πρόβλεψη και ερμηνεία του μουσικού συναισθήματος.

8.3 Μελλοντική Έρευνα

Με βάση τα αποτελέσματα και τους περιορισμούς της ΔΕ, μπορούν να προταθούν ορισμένες κατευθύνσεις για μελλοντική έρευνα. Αρχικά, θα ήταν χρήσιμη η εφαρμογή της ίδιας μεθοδολογικής προσέγγισης σε μεγαλύτερα και πιο ετερογενή σύνολα δεδομένων, τα οποία θα περιλαμβάνουν διαφορετικά μουσικά είδη, πολιτισμικά πλαίσια και μορφές συναισθηματικών αξιολογήσεων. Με τον τρόπο αυτό, θα μπορούσε να εξεταστεί σε μεγαλύτερο βαθμό η δυνατότητα γενίκευσης των μοντέλων.

Επιπλέον, μελλοντικές εργασίες θα μπορούσαν να διερευνήσουν περισσότερα ή διαφορετικά σύνολα ηχητικών χαρακτηριστικών, καθώς και χαρακτηριστικά υψηλότερου επιπέδου, τα οποία ενδέχεται να αποτυπώνουν πιο άμεσα μουσικές και συναισθηματικές ποιότητες. Η ενσωμάτωση χαρακτηριστικών που σχετίζονται με τη μελωδική δομή, την υφή, τη μορφή ή την εκφραστικότητα της μουσικής θα μπορούσε να συμβάλει στη μείωση του σημασιολογικού χάσματος μεταξύ χαμηλού επιπέδου ηχητικών χαρακτηριστικών και υψηλού επιπέδου συναισθηματικών εννοιών.

Μια ακόμη κατεύθυνση αφορά τη διερεύνηση περισσότερων μοντέλων μηχανικής μάθησης και διαδικασιών βελτιστοποίησης. Μελλοντικές εργασίες θα μπορούσαν να εξετάσουν πιο εκτεταμένη υπερπαραμετρική βελτιστοποίηση, καθώς και εναλλακτικές μεθόδους όπως XGBoost ή νευρωνικά δίκτυα. Η σύγκριση αυτών των προσεγγίσεων θα μπορούσε να δείξει κατά πόσο είναι δυνατή η περαιτέρω βελτίωση της προβλεπτικής απόδοσης.

Ενδιαφέρον παρουσιάζει επίσης η περαιτέρω διερεύνηση της χαρτογράφησης μεταξύ διαστασιακών και κατηγορικών συναισθηματικών χώρων. Μελλοντική έρευνα θα μπορούσε να εξετάσει διαφορετικές μεθόδους χαρτογράφησης, πιο σύνθετα μοντέλα μετασχηματισμού ή ενοποιημένα πλαίσια μάθησης που αξιοποιούν ταυτόχρονα διαστασιακές και κατηγορικές αξιολογήσεις. Παράλληλα, θα ήταν χρήσιμη η αξιολόγηση της χαρτογράφησης σε διαφορετικά σύνολα δεδομένων, ώστε να διαπιστωθεί αν η σχέση μεταξύ των δύο συναισθηματικών χώρων παραμένει σταθερή σε διαφορετικά μουσικά πλαίσια.

Τέλος, η σημασιολογική αναπαράσταση των αποτελεσμάτων μπορεί να επεκταθεί περαιτέρω. Μελλοντικές εργασίες θα μπορούσαν να διερευνήσουν πιο συστηματικούς τρόπους παραγωγής λεκτικών περιγραφών συναισθηματικών προφίλ, καθώς και τη σύνδεσή τους με μουσικές οντολογίες ή εφαρμογές αναζήτησης και σύστασης μουσικής. Με αυτόν τον τρόπο, τα αποτελέσματα των μοντέλων πρόβλεψης θα μπορούσαν να αξιοποιηθούν σε πιο ολοκληρωμένα συστήματα συναισθηματικής αναζήτησης και οργάνωσης μουσικών συλλογών.

Βιβλιογραφικές Αναφορές

- Bogdanov, D., Wack, N., Gómez, E., Gulati, S., Herrera, P., Mayor, O., Roma, G., Salamon, J., Zapata, J. R., & Serra, X. (2013). Essentia: An audio analysis library for music information retrieval. *Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR 2013)*, 493–498. <https://doi.org/10.5281/zenodo.1415016>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Casey, M. A., Veltkamp, R., Goto, M., Leman, M., Rhodes, C., & Slaney, M. (2008). Content-based music information retrieval: Current directions and future challenges. *Proceedings of the IEEE*, 96(4), 668–696. <https://doi.org/10.1109/JPROC.2008.916370>
- Celma, Ò., Herrera, P., & Serra, X. (2006). Bridging the music semantic gap. *International Semantic Web Conference*.
- Chowdhury, S., Vall, A., Haunschmid, V., & Widmer, G. (2019). Towards explainable music emotion recognition: The route via mid-level features. *Proceedings of the 20th International Society for Music Information Retrieval Conference (ISMIR 2019)*, 237–243.
- Downie, J. S. (2003). Music information retrieval. *Annual Review of Information Science and Technology*, 37(1), 295–340. <https://doi.org/10.1002/aris.1440370108>
- Eerola, T. & Vuoskoski, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, 39(1), 18–49. <https://doi.org/10.1177/0305735610362821>
- Eerola, T., & Vuoskoski, J. K. (2013). A review of music and emotion studies: Approaches, emotion models, and stimuli. *Music Perception*, 30(3), 307–340. <https://doi.org/10.1525/MP.2012.30.3.307>
- Eerola, T., Lartillot, O., & Toivainen, P. (2009). Prediction of Multidimensional Emotional Ratings in Music from Audio Using Multivariate Regression Models. *Proceedings of the 10th International Society for Music Information Retrieval Conference*, 621–626. <https://doi.org/10.5281/zenodo.1416730>

- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3–4), 169–200. <https://doi.org/10.1080/02699939208411068>
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis*, 38(4), 367–378. [https://doi.org/10.1016/S0167-9473\(01\)00065-2](https://doi.org/10.1016/S0167-9473(01)00065-2)
- Gabrielsson, A. (2001). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae*, 5(1_suppl), 123–147. <https://doi.org/10.1177/10298649020050S105>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. <https://doi.org/10.1007/978-0-387-84858-7>
- Hu, X., Downie, J. S., Laurier, C., Bay, M., & Ehmann, A. F. (2008). The 2007 MIREX audio mood classification task: Lessons learned. In *Proceedings of the 9th International Society for Music Information Retrieval Conference (ISMIR 2008)*, 462–467.
- Huq, A., Bello, J. P., & Rowe, R. (2010). Automated music emotion recognition: A systematic evaluation. *Journal of New Music Research*, 39(3), 227–244. <https://doi.org/10.1080/09298215.2010.513733>
- Juslin, P. N., & Sloboda, J. A. (Eds) (2010). *Handbook of music and emotion: Theory, research, applications*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199230143.001.0001>
- Kang, J., & Herremans, D. (2025a). Are we there yet? a brief survey of music emotion prediction datasets, models and outstanding challenges. *IEEE Transactions on Affective Computing*. <http://arxiv.org/abs/2406.08809>
- Kang, J., & Herremans, D. (2025b). Towards unified music emotion recognition across dimensional and categorical models. *arXiv preprint* [arXiv:2502.03979](https://arxiv.org/abs/2502.03979)
- Kim, Y. E., Schmidt, E. M., Migneco, R., Morton, B. G., Richardson, P., Scott, J., Speck, J. A., & Turnbull, D. (2010). Music emotion recognition: A state of the art review. *Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR 2010)*, 255–266.

- Kraev, E., Koseoglu, B., Traverso, L., & Topiwalla, M. (2024). Shap-select: lightweight feature selection using SHAP values and regression. *arXiv preprint*. <http://arxiv.org/abs/2410.06815>
- Lartillot, O. (2017). *MIRtoolbox 1.7 user's manual*.
- Lartillot, O., & Toivianen, P. (2007). A Matlab toolbox for musical feature extraction from audio. *Proceedings of the 10th International Conference on Digital Audio Effects (DAFx-07)*, 237–244.
- Li, Z., Huang, J., Sun, Y., & Li, A. (2025). *Interpretable Framework for Continuous Music Emotion Recognition Based on Feature Selection and Machine Learning (preprint)*. <https://doi.org/10.21203/rs.3.rs-8200113/v1>
- Lundberg, S. M., & Lee, S. (2017). A unified approach to interpreting model predictions. In I. Guyon, U. von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in neural information processing systems 30*, 4765–4774, Curran Associates. https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf
- Mahalanobis, P. C. (1936). On the generalized distance in statistics. *Proceedings of the National Institute of Sciences of India*, 2(1), 49–55.
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press. <https://nlp.stanford.edu/IR-book/information-retrieval-book.html>
- Molnar, C. (2025). *Interpretable machine learning: A guide for making black box models explainable* (3rd ed.).
- Müller, M. (2015). *Fundamentals of music processing: Audio, analysis, algorithms, applications*. Springer. <https://doi.org/10.1007/978-3-319-21945-5>
- Orio, N. (2006). Music retrieval: A tutorial and review. *Foundations and Trends in Information Retrieval*, 1(1), 1–90. <https://doi.org/10.1561/1500000002>
- Panda, R., Malheiro, R., & Paiva, R. P. (2023). Audio Features for Music Emotion Recognition: A Survey. *IEEE Transactions on Affective Computing*, 14(1), 68–88. <https://doi.org/10.1109/TAFFC.2020.3032373>

- Panda, R., Rocha, B., & Paiva, R. P. (2015). Music emotion recognition with standard and melodic audio features. *Applied Artificial Intelligence*, 29(4), 313–334. <https://doi.org/10.1080/08839514.2015.1016389>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830. <http://jmlr.org/papers/v12/pedregosa11a.html>
- Raimond, Y., Abdallah, S. A., Sandler, M. B., & Giasson, F. (2007). The Music Ontology. *Proceedings of the 8th International Society for Music Information Retrieval Conference (ISMIR 2007)*.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161–1178. <https://doi.org/10.1037/h0077714>
- Schedl, M., Gómez, E., & Urbano, J. (2014). Music information retrieval: Recent developments and applications. *Foundations and Trends in Information Retrieval*, 8(2–3), 127–261. <https://doi.org/10.1561/15000000042>
- Scherer, K. R. (2004). Which Emotions Can be Induced by Music? What Are the Underlying Mechanisms? And How Can We Measure Them?. *Journal of New Music Research*, 33(3), 239–251. <https://doi.org/10.1080/0929821042000317822>
- Scikit-learn developers (n.d.a). *Scikit-learn: User Guide*. Retrieved May 16, 2026, from https://scikit-learn.org/stable/user_guide.html
- Scikit-learn developers (n.d.b). *RandomForestRegressor*. Scikit-learn: User Guide. Retrieved May 16, 2026, from <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html>
- Scikit-learn developers (n.d.c). *GradientBoostingRegressor*. Scikit-learn: User Guide. Retrieved May 16, 2026, from <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingRegressor.html>
- Scikit-learn developers (n.d.d). *MultiOutputRegressor*. Scikit-learn: User Guide. Retrieved May 16, 2026, from <https://scikit-learn.org/stable/modules/generated/sklearn.multioutput.MultiOutputRegressor.html>

- Scikit-learn developers (n.d.e). *DummyRegressor*. Scikit-learn: User Guide. Retrieved May 16, 2026, from <https://scikit-learn.org/stable/modules/generated/sklearn.dummy.DummyRegressor.html>
- Scikit-learn developers (n.d.f). *Metrics and scoring: Quantifying the quality of predictions*. Scikit-learn: User Guide. Retrieved May 16, 2026, from https://scikit-learn.org/stable/modules/model_evaluation.html
- Shapley, L. (1953). A Value for n-Person Games. In H. Kuhn & A. Tucker (Ed.), *Contributions to the Theory of Games*, 2 (pp. 307-318). Princeton: Princeton University Press. <https://doi.org/10.1515/9781400881970-018>
- Sloboda, J. A., & Juslin, P. N. (2010). At the interface between the inner and outer world: Psychological perspectives. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 73–97). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199230143.001.0001>
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427-437.
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6), 80-83. <https://doi.org/10.2307/3001968>
- Yang, Y. H., & Chen, H. H. (2012). Machine recognition of music emotion: A review. In *ACM Transactions on Intelligent Systems and Technology*, 3(3). <https://doi.org/10.1145/2168752.2168754>
- Zhang, J. L., Huang, X. L., Yang, L., & Nie, L. (2016). Bridge the semantic gap between pop music acoustic feature and emotion: Build an interpretable model. *Neurocomputing*, 208, 333–341. <https://doi.org/10.1016/j.neucom.2016.01.099>

Παράρτημα Α: Πίνακας Μουσικών Αποσπασμάτων και Συναισθηματικών Αξιολογήσεων

Πίνακας Α.1: Μουσικά αποσπάσματα και συναισθηματικές αξιολογήσεις

Number	Album Name	Track	Time	Valence	Energy	Tension	Happy	Sad	Anger	Fear	Tender
001	Lethal weapon 3	8	04:15-04:29	2.56	7.73	8.38	1.21	1.75	6.39	4.40	1.01
002	The Rainmaker	7	01:45-02:00	2.50	8.21	8.56	1.09	1.60	5.76	5.93	1.10
003	The Alien Trilogy	9	00:03-00:18	2.92	8.21	8.17	1.31	1.55	6.31	4.06	1.01
004	Cape Fear	1	02:15-02:30	2.88	6.75	7.98	1.19	2.10	5.42	4.51	1.06
005	The Fifth Element	19	00:00-00:20	3.08	6.83	7.50	1.07	2.96	6.13	4.24	1.07
006	Crouching Tiger Hidden Dragon	8	01:12-01:25	4.90	8.00	7.52	2.55	1.55	3.27	2.99	1.12
007	Batman Returns	2	00:18-00:33	3.83	7.94	7.94	1.49	2.07	4.64	4.60	1.04
008	Man of Galilee CD1	6	00:40-01:07	4.00	6.21	7.08	1.19	5.15	3.81	3.19	1.37
009	The Untouchables	8	01:38-01:53	4.54	8.31	7.94	1.91	1.37	4.13	4.12	1.00
010	Oliver Twist	15	02:05-02:25	2.44	6.67	8.13	1.04	2.21	4.42	5.91	1.07
011	Batman Returns	5	00:09-00:25	2.38	6.38	8.33	1.09	1.60	2.79	7.42	1.07
012	JFK	8	01:26-01:40	3.60	7.42	8.06	1.28	1.57	2.28	6.12	1.27
013	JFK	8	00:08-00:25	3.19	7.10	8.15	1.31	1.21	2.13	6.34	1.06
014	The Alien Trilogy	5	00:26-00:41	2.31	7.48	8.54	1.07	1.36	2.61	6.79	1.01
015	Hannibal	1	00:40-00:54	2.31	6.13	8.10	1.04	1.70	2.00	6.52	1.06
016	Running Scared	6	02:53-03:07	4.21	7.56	7.79	1.37	2.10	1.88	4.58	1.22
017	The Untouchables	8	01:38-01:53	3.02	4.58	7.75	1.10	1.75	1.75	6.25	1.16
018	The Fifth Element	17	00:00-00:19	3.96	7.17	7.69	1.21	1.45	4.42	4.81	1.10
019	Lethal weapon 3	7	00:00-00:16	4.58	4.83	6.81	1.49	3.69	1.34	2.93	2.12
020	Man of Galilee CD1	2	03:45-04:02	4.58	8.33	8.33	2.33	1.39	4.18	4.15	1.06
021	The Rainmaker	3	02:55-03:13	8.00	7.77	4.02	6.63	1.12	1.00	1.09	2.19

Number	Album Name	Track	Time	Valence	Energy	Tension	Happy	Sad	Anger	Fear	Tender
022	Batman	18	00:55-01:15	7.56	8.35	5.06	6.81	1.30	1.33	1.18	1.51
023	Shallow Grave	6	02:02-02:17	8.27	8.54	4.46	8.21	1.10	1.04	1.03	1.45
024	Man of Galilee CD1	2	03:02-03:18	7.96	8.63	5.94	6.94	1.09	1.27	1.12	1.27
025	Oliver Twist	1	00:17-00:34	7.67	7.42	4.15	6.28	1.58	1.09	1.12	2.22
026	The Omen	9	00:00-00:24	6.90	5.25	4.21	4.66	2.16	1.04	1.52	4.34
027	Oliver Twist	8	01:40-02:04	7.75	8.08	4.77	6.27	1.67	1.09	1.24	2.30
028	Grizzly Man	1	00:00-00:27	7.77	4.27	2.52	3.42	2.72	1.01	1.09	5.34
029	The Portrait of a Lady	3	00:23-00:45	7.13	3.60	2.42	3.63	2.90	1.00	1.21	5.31
030	Nostradamus	2	01:09-01:28	7.00	3.44	2.79	2.87	2.88	1.03	1.31	4.70
031	The English Patient	18	00:07-00:32	5.02	3.19	3.65	1.25	6.78	1.10	1.69	3.58
032	Running Scared	15	02:06-02:27	4.04	3.67	5.33	1.09	5.15	1.99	3.07	1.94
033	The Portrait of a Lady	9	00:00-00:22	4.38	2.48	3.27	1.12	7.18	1.07	1.78	3.03
034	Big Fish	15	00:55-01:11	4.85	3.52	4.17	1.40	6.12	1.13	1.39	3.15
035	Man of Galilee CD1	8	01:20-01:37	4.25	5.58	6.25	1.66	5.73	2.39	2.22	2.79
036	Angel Heart	4	00:08-00:28	5.13	2.52	3.81	1.34	5.21	1.10	2.19	2.43
037	Batman	5	01:08-01:22	4.79	3.88	5.25	1.72	4.18	1.16	3.27	3.04
038	Dracula	7	00:00-00:12	4.73	2.79	4.60	1.15	4.07	1.09	2.16	2.70
039	Shakespeare in Love	3	00:59-01:17	6.19	3.08	2.90	2.00	4.78	1.06	1.25	5.09
040	The English Patient	7	00:00-00:31	5.58	3.67	4.40	2.09	5.33	1.01	1.55	4.46
041	Shine	10	01:28-01:48	7.27	3.85	2.48	3.72	3.01	1.01	1.15	6.66
042	Pride Prejudice	1	00:10-00:26	7.65	3.96	1.96	3.70	2.49	1.00	1.18	6.06
043	Dances with Wolves	4	01:31-01:48	7.44	3.73	2.35	4.25	2.01	1.03	1.06	5.73
044	Pride Prejudice	12	00:01-00:15	7.38	3.21	1.85	2.85	2.87	1.01	1.10	5.48
045	Oliver Twist	8	00:14-00:30	6.42	3.42	2.85	2.61	3.31	1.00	1.13	4.31

Number	Album Name	Track	Time	Valence	Energy	Tension	Happy	Sad	Anger	Fear	Tender
046	Batman	9	00:00-00:19	5.81	3.04	3.44	1.57	4.81	1.04	1.60	4.40
047	Oliver Twist	8	01:15-01:32	7.23	4.23	2.63	3.75	3.52	1.03	1.21	5.39
048	Dracula	4	00:55-01:09	5.85	3.56	4.33	1.94	3.15	1.06	2.03	4.43
049	Juha	2	02:11-02:26	7.60	5.25	2.73	4.93	2.22	1.00	1.06	5.12
050	Oliver Twist	2	00:00-00:29	5.77	3.27	3.69	1.64	5.28	1.01	1.63	3.69
051	Juha	10	00:20-00:38	7.30	5.00	3.18	4.64	2.51	1.00	1.00	4.96
052	Blanc	12	00:51-01:06	6.82	4.97	3.46	3.34	2.81	1.00	1.09	4.62
053	Gladiator	17	00:14-00:27	7.07	6.76	4.55	4.79	1.85	1.00	1.21	3.60
054	Pride Prejudice	9	00:01-00:21	6.66	4.21	3.75	1.89	4.57	1.00	1.23	5.26
055	Dances with Wolves	10	00:28-00:46	7.04	7.27	4.66	5.60	1.17	1.38	1.40	1.96
056	Man of Galilee CD1	2	00:19-00:42	6.67	6.75	5.16	5.11	1.72	1.11	1.19	3.04
057	Shakespeare in Love	21	00:03-00:21	6.57	7.55	5.51	4.49	1.96	1.47	1.45	2.02
058	Vertigo OST	6	02:02-02:17	6.03	3.82	4.31	1.77	5.06	1.06	1.51	4.30
059	Vertigo OST	6	04:42-04:57	6.69	6.33	4.36	4.32	2.02	1.11	1.15	2.60
060	Outbreak	6	00:16-00:31	5.84	5.90	5.67	2.72	2.47	1.28	2.49	2.02
061	Juha	18	02:30-02:46	5.78	5.97	5.99	3.00	2.32	1.26	1.60	2.17
062	Shakespeare in Love	11	00:21-00:36	3.48	5.58	7.37	1.00	2.74	2.94	4.98	1.17
063	Batman	9	00:57-01:16	4.76	3.96	6.04	1.19	3.91	1.23	3.83	2.32
064	The Fifth Element	9	00:00-00:18	3.03	5.12	7.58	1.28	1.55	2.11	6.19	1.15
065	Big Fish	15	00:15-00:30	4.07	5.04	6.93	1.26	3.28	2.36	4.30	1.57
066	The English Patient	8	01:35-01:57	2.58	6.85	8.16	1.06	1.74	3.83	5.87	1.04
067	Lethal weapon 3	7	00:00-00:16	4.72	4.72	6.67	1.32	3.57	1.26	3.02	2.13
068	Road to Perdition	6	00:34-00:49	2.51	6.03	8.01	1.00	1.79	3.13	6.70	1.04
069	Hellraiser	5	00:00-00:15	2.36	7.09	8.24	1.00	1.83	6.38	5.34	1.02

Number	Album Name	Track	Time	Valence	Energy	Tension	Happy	Sad	Anger	Fear	Tender
070	Grizzly Man	16	01:05-01:32	2.04	4.54	7.76	1.04	1.68	2.17	6.17	1.06
071	The Untouchables	6	01:50-02:05	7.45	7.64	4.27	6.89	1.11	1.02	1.02	1.79
072	Man of Galilee CD1	2	03:02-03:18	7.45	8.39	5.61	7.34	1.09	1.19	1.17	1.11
073	Shine	5	02:00-02:16	6.91	8.12	5.66	6.36	1.17	1.04	1.09	2.13
074	Shine	15	01:00-01:19	6.36	7.13	6.54	4.87	1.38	1.32	1.62	2.11
075	Batman	18	00:55-01:15	7.31	8.04	4.57	7.09	1.04	1.21	1.17	1.40
076	Juha	2	00:07-00:18	6.61	4.93	4.00	3.45	2.26	1.00	1.04	2.49
077	Lethal weapon 3	4	01:40-02:00	6.27	6.34	5.42	4.53	1.55	1.04	1.26	1.83
078	Crouching Tiger Hidden Dragon	13	01:52-02:10	5.30	5.87	5.93	1.72	4.85	1.23	2.02	2.74
079	Batman	4	02:31-02:51	7.25	8.12	5.03	6.60	1.02	1.26	1.19	1.19
080	Oliver Twist	7	01:30-01:46	5.88	6.78	6.10	4.00	1.81	1.06	2.02	1.66
081	Juha	16	00:00-00:15	5.60	3.07	4.01	1.13	6.38	1.00	1.53	3.15
082	Big Fish	15	00:55-01:11	5.24	3.70	4.79	1.36	5.36	1.06	1.43	3.40
083	Big Fish	11	01:26-01:40	6.40	4.49	4.60	2.43	3.96	1.00	1.23	4.26
084	Blanc	18	00:00-00:16	5.19	2.99	4.43	1.04	6.17	1.04	2.02	3.02
085	Oliver Twist	6	00:51-01:07	4.54	4.33	5.97	1.23	3.66	1.23	3.47	1.68
086	Running Scared	15	02:06-02:27	4.28	3.88	5.84	1.02	5.94	1.96	3.13	1.57
087	Road to Perdition	16	00:17-00:32	4.55	3.57	5.61	1.00	5.60	1.60	3.66	1.81
088	Blanc	10	00:13-00:31	5.03	2.96	4.10	1.00	5.66	1.19	2.15	2.32
089	Blanc	16	00:00-00:15	4.33	3.21	5.40	1.04	4.94	1.09	3.23	2.13
090	Batman Returns	12	00:57-01:14	4.43	3.55	5.84	1.00	5.00	1.40	2.98	2.43
091	The Alien Trilogy	11	02:12-02:27	3.69	7.09	7.66	1.04	2.13	3.38	5.57	1.06
092	The Fifth Element	13	00:17-00:31	2.99	6.99	8.12	1.02	1.66	6.47	4.74	1.00
093	Babylon 5	3	02:47-03:00	2.46	7.25	8.36	1.06	1.51	5.51	6.51	1.00

Number	Album Name	Track	Time	Valence	Energy	Tension	Happy	Sad	Anger	Fear	Tender
094	Hellraiser	10	02:44-03:00	3.06	6.01	8.07	1.02	2.04	4.00	5.96	1.26
095	Oliver Twist	15	02:05-02:25	3.01	6.28	7.94	1.00	1.94	4.53	5.85	1.04
096	The Missing	3	02:45-03:06	3.63	5.54	6.97	1.09	3.47	2.74	4.11	1.60
097	Shallow Grave	4	01:04-01:19	2.88	6.51	7.78	1.53	1.19	2.47	5.74	1.02
098	Naked Lunch	7	01:01-01:20	5.46	4.78	5.57	1.79	3.19	1.30	2.19	2.89
099	Dracula	5	00:11-00:27	4.16	5.22	6.84	1.11	4.89	2.55	2.81	1.83
100	Cape Fear	2	01:25-01:40	3.22	5.09	7.43	1.00	2.15	2.11	5.28	1.19
101	Juha	2	02:11-02:26	7.57	4.60	2.46	3.94	2.28	1.00	1.00	4.94
102	Shakespeare in Love	6	00:00-00:19	6.01	4.96	5.25	2.40	3.81	1.04	1.26	3.30
103	The Fifth Element	12	00:00-00:17	5.87	4.54	4.73	2.06	3.64	1.17	1.36	2.74
104	Crouching Tiger Hidden Dragon	11	00:28-00:46	6.19	3.63	3.48	1.32	5.60	1.09	1.30	3.79
105	Pride Prejudice	4	00:10-00:29	8.01	8.27	4.16	7.85	1.19	1.00	1.00	1.45
106	Lethal weapon 3	10	01:59-02:17	7.37	3.73	2.21	3.13	2.45	1.00	1.00	6.57
107	The Godfather	5	01:12-01:28	7.09	4.21	2.75	4.26	2.43	1.00	1.09	5.89
108	Gladiator	4	00:48-01:06	5.87	2.97	3.54	1.17	5.13	1.02	1.45	4.09
109	Pride Prejudice	13	01:02-01:20	5.37	2.28	3.22	1.02	7.00	1.13	1.45	3.47
110	Big Fish	8	00:12-00:34	6.75	3.90	2.82	3.23	3.28	1.00	1.02	4.85

Υπεύθυνη Δήλωση Συγγραφέα:

Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν.1599/1986, η παρούσα εργασία αποτελεί αποκλειστικά προϊόν προσωπικής μου εργασίας, δεν προσβάλλει κάθε μορφής δικαιώματα διανοητικής ιδιοκτησίας, προσωπικότητας και προσωπικών δεδομένων τρίτων, δεν περιέχει έργα/εισφορές τρίτων για τα οποία απαιτείται άδεια των δημιουργών/δικαιούχων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και πληρούν τους κανόνες της επιστημονικής παράθεσης.