

ALZHEIMER'S DISEASE

AUDIO CLASSIFICATION

ROIDIS DIMITRIOS



ALZHEIMER'S DISEASE

Το **alzheimer** είναι μια **νευροεκφυλιστική νόσος** που επηρεάζει τη μνήμη, τη σκέψη και τη συμπεριφορά. Είναι η πιο κοινή μορφή **άνοιας** και επηρεάζει κυρίως ηλικιωμένα άτομα.

Κυριότερα συμπτώματα της ασθένειας είναι η **δυσκολία στην ομιλία** και η εύρεση λέξεων, η απώλεια μνήμης, οι αλλαγές στη συμπεριφορά καθώς και κινητικά προβλήματα.

Η διάγνωση γίνεται με **νευροψυχολογικά τεστ** (MMSE, MoCA), **απεικονιστικές εξετάσεις** (MRI, PET scan) και **βιοδείκτες** στο εγκεφαλονωτιαίο υγρό.

AUDIO BINARY CLASSIFICATION

Το συγκεκριμένο project επικεντρώνεται στην **ομιλία** ως ένδειξη πιθανού alzheimer.

Το dataset που χρησιμοποιήθηκε για το modelling αποτελείται από:

- **Audios** απο προσωπικές συνεδρίες πιθανών ασθενών με ειδικούς
- **Groundtruth** dataset το οποίο εκτός των labels-classes (***probableAD, Control***), περιέχει πληροφορίες σχετικά με ηλικία, φύλο, εκπαίδευση, και αποτελέσματα cognitive test (MMSE).

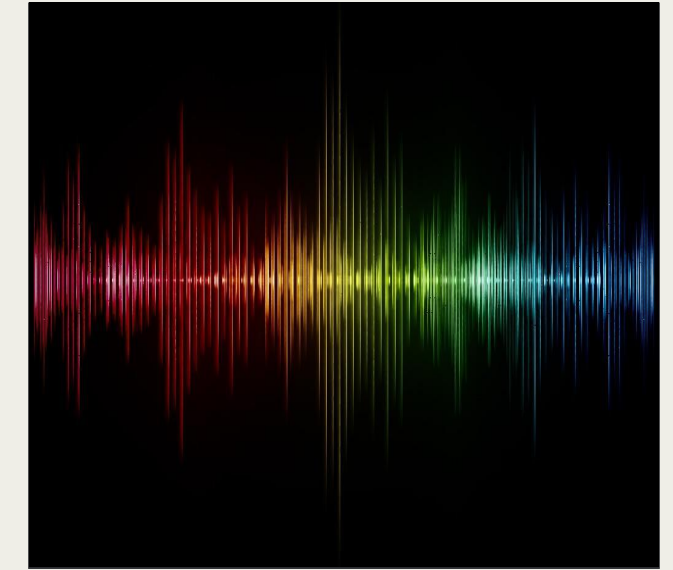
DATA EXPLORATION



237 mp3 files
διάρκειας απο 45sec
εως και 2 λεπτά
ισορροπημένα
ανάμεσα στις 2 κλάσεις
(52% Alzheimer, 48%
Control).



Metadata Analysis
Δεν υπήρχαν
σημαντικά metadata
σε όλα τα audio ώστε
να επικυρωθεί η
μοναδικότητα του
ασθενή.
Χρησιμοποιήθηκε η
βιβλιοθήκη mutagen.



Know Your Data
Listen To Your Audios
Χαρακτηριστικές
διαφορές ανάμεσα στα
audios των 2 κλάσεων
(ροή λόγου,
παρεμβολή και
καθοδήγηση γιατρού)
οι οποίες πιθανόν
αποτελούν feature.

FEATURE EXTRACTION



137

4

2

MFCC

Εξαγωγή 137 mel-frequency features ανά 15sec από τα audio με τη βιβλιοθήκη ***pyAudioAnalysis***

Groundtruth

Dataset με 4 features : mmse, gender, age, educ

Audio to Text to Features

Εξαγωγή text διαλόγων από τα audio με τη βιβλιοθήκη ***assemblyai***, tokenization του text σε λέξεις και εξαγωγή feature σχετικών με lexical richness και word variance με τη βιβλιοθήκη ***ntlk***

FURTHER ANALYSIS - DATA CLEANING

Τα νέα dataset που προέκυψαν από την εξαγωγή των features εξετάστηκαν περεταίρω.

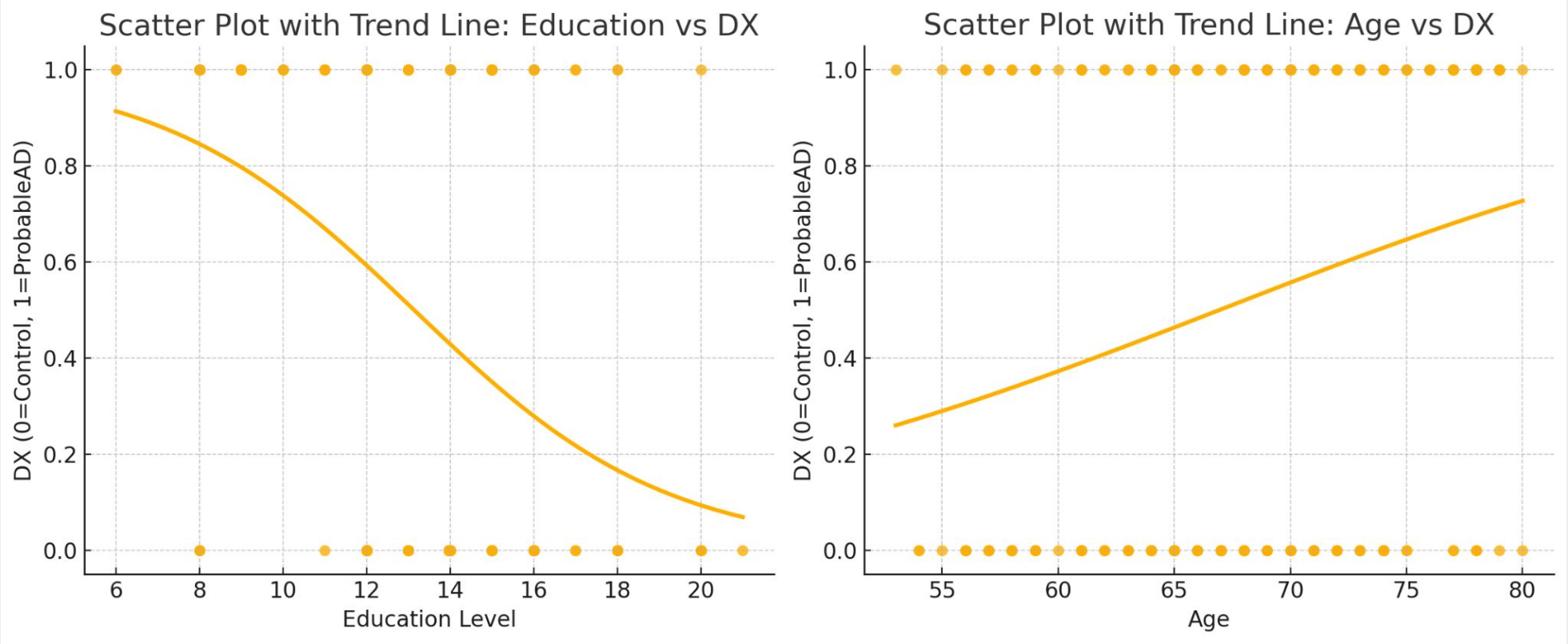
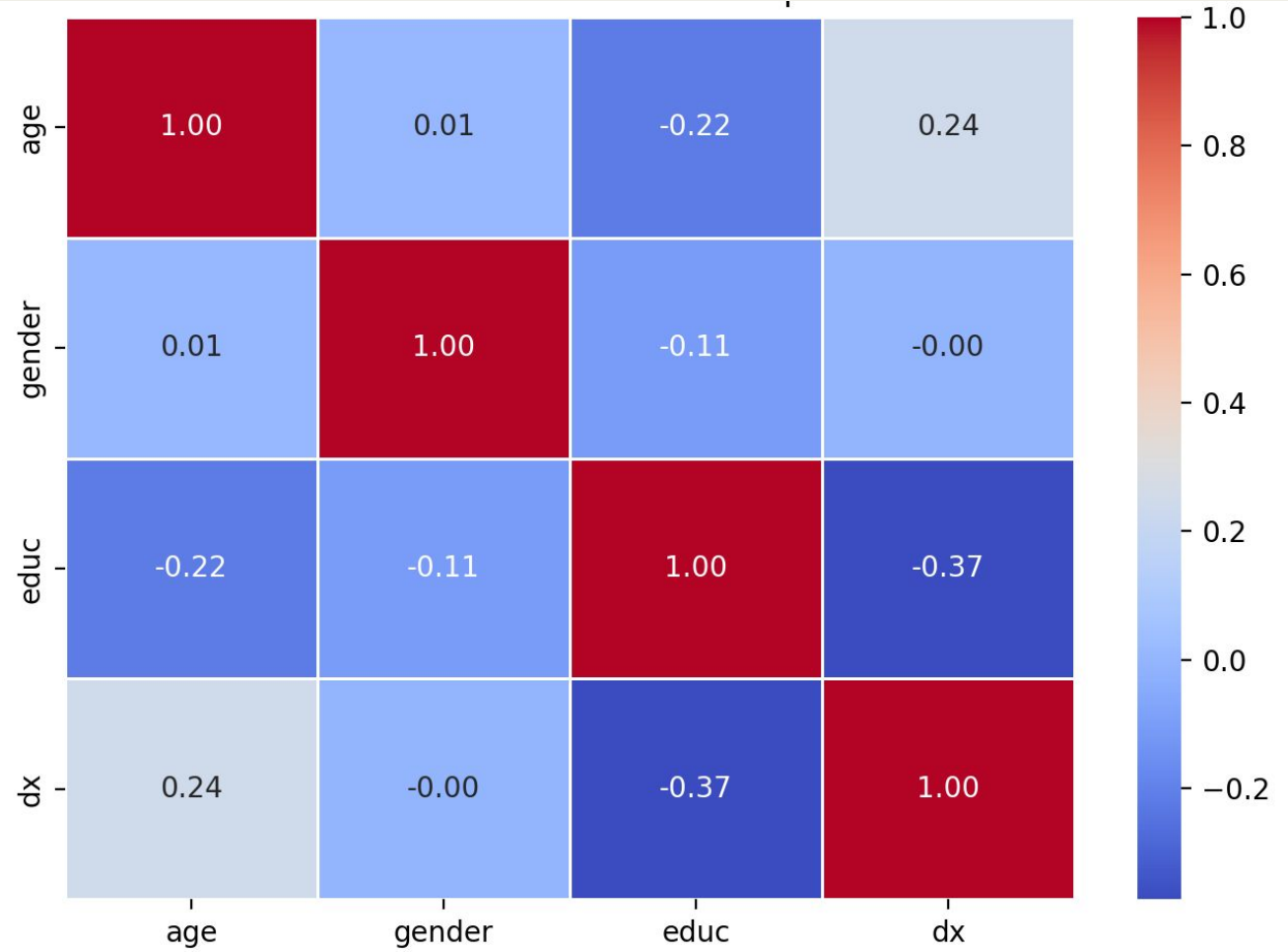
Συγκεκριμένα, εξήχθησαν **correlation matrices** και έγινε **linearity analysis** προσφέροντας έτσι μεγαλύτερο intuition πάνω στα data και σκέψεις για πιθανά μοντέλα. Το training ground truth dataset έδειξε συσχέτιση ανάμεσα στο age και το educ με την διάγνωση για Alzheimer η μη.

Προηγουμένως το ground truth είχε καθαριστεί από null values, το mmse έγινε drop ενώ τα categorical values του gender μετατράπηκαν σε 0 και 1.

Όσον αφορά τα mfcc features, πολύ λίγα εκ των 137 έδειξαν linear relationship με το ground truth. Αυτό επιβεβαιώνεται και από importance analysis χρησιμοποιώντας π.χ. το ***RandomForestClassifier*** της ***sklearn***.

Τέλος, τα 2 text features, έδειξαν ύπαρξη γραμμικότητας η οποία είναι πιο έντονη στο feature *Hapax Legomena*.

FURTHER ANALYSIS

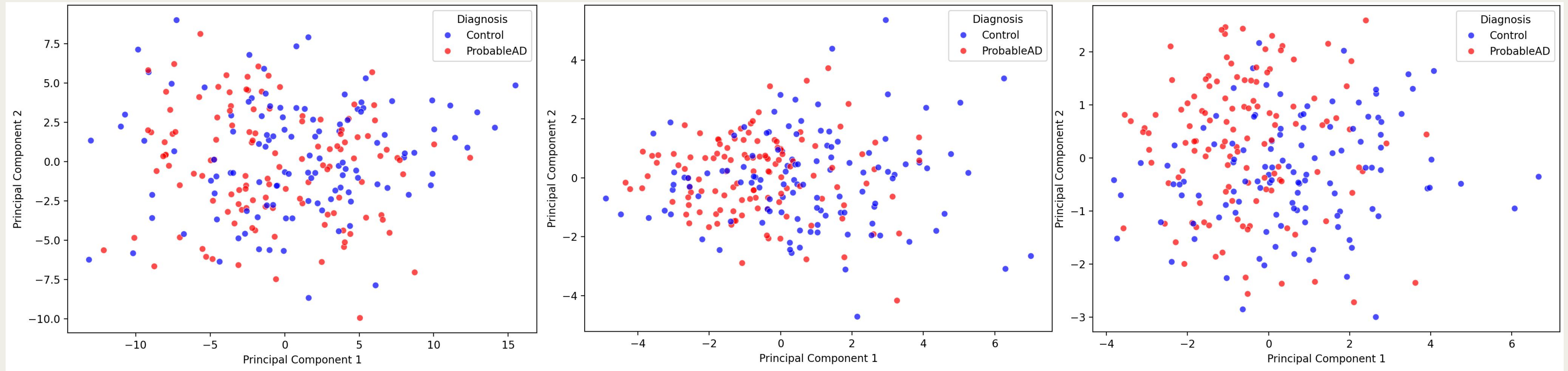


FINAL DATASETS - FUSIONS - EXTRACTIONS

- ***fuse***(MFCC + Text Features + Ground Truth) **(137 + 2 + 3 features)**
- MFCC **(137 features)**
- Text Features **(2 features)**
- Ground Truth **(3 features)**
- MFCC top 5 **(5 features)**
- MFCC top 10 **(10 features)**
- ***fuse***(MFCC top 5 + Text Features + Ground Truth) **(5 + 3 + 3 features)**

όλα τα dataframes έγιναν extract σε csv για εύκολη επαναχρησιμοποίηση στους classifiers.

VARIANCE CAPTURE AND SCATTER PLOTS



fuse(MFCC + Text Features + Ground
Truth) (**137 + 2 + 3 features**)

fuse(MFCC top 10 + Text
Features + Ground Truth) (**10 +
2 + 3 features**)

fuse(MFCC top 5 + Text
Features + Ground Truth) (**5 + 2
+ 3 features**)

CANDIDATE CLASSIFIERS TO UTILIZE

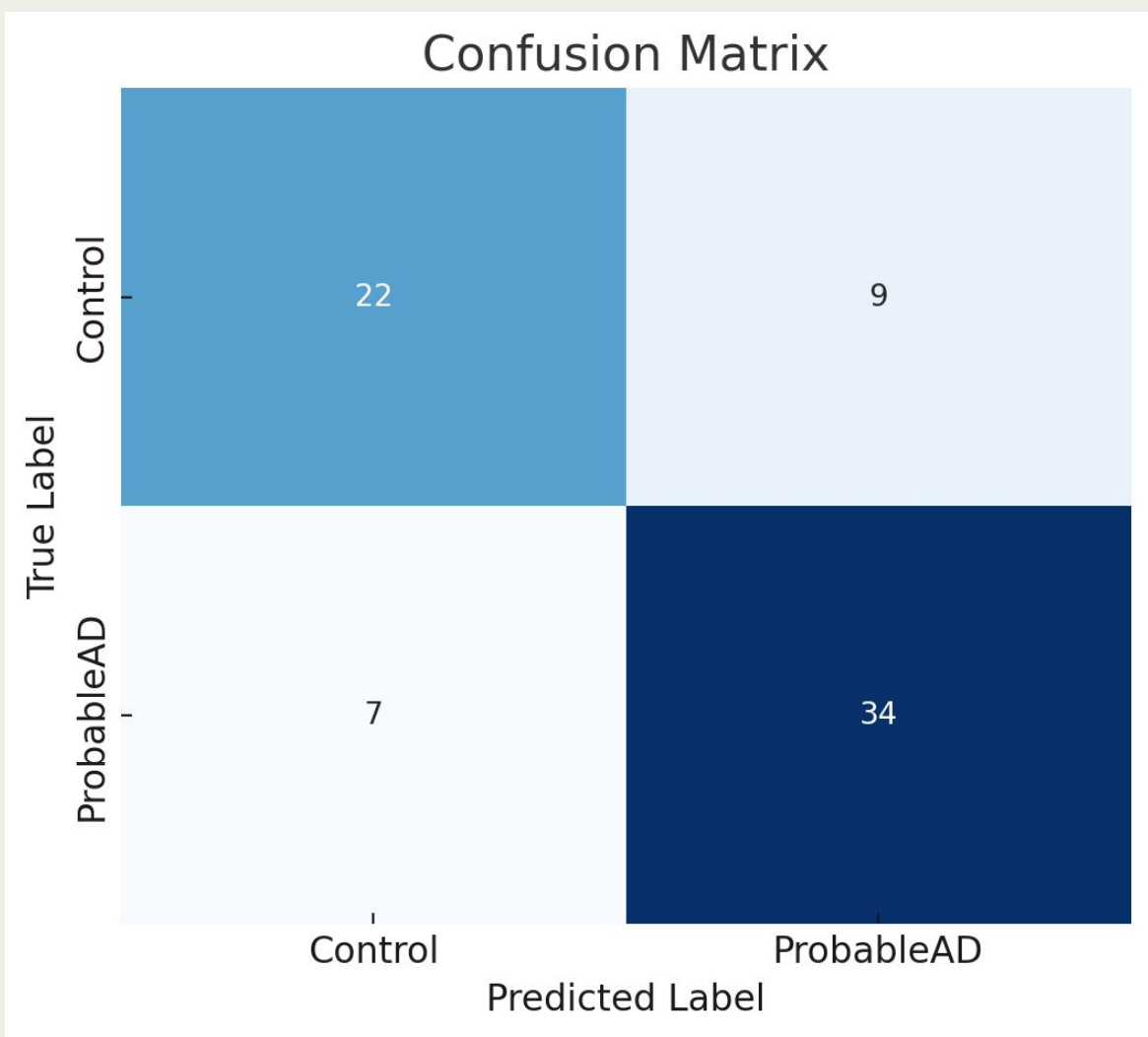
Τα dataset που προέκυψαν χαρακτηρίζονται από υψηλό **dimensionality** αλλά και κάποιο ισχυρό **correlation** μεταξύ ορισμένων feature με το ground truth. Αυτό κάνει point προς συγκεκριμένες κατευθύνσεις :

- **Support Vector Machines** για να βρεθεί το optimal hyperplane σε high dimensional data.
- **Logistic Regression, LDA** λόγω της ύπαρξης γραμμικότητας σε κάποια features και κυρίως ως ένα baseline για αρχικά test.
- **Ensemble Voting Classifier** για τα 3 διαφορετικά dataset (mfcc, text features και ground truth) ώστε να γίνουν utilise οι διαφορετικότητες τους και τα predictions να προκύπτουν από majority vote (2/3). πχ. *SVM* στα mfcc, *LDA* στα text features και *Logistic Regression* στο ground truth.
- Διάφοροι άλλοι classifiers για λόγους ευρύτερου τεστ.

Κάθε dataset υπόκειται σε normalization ή standardization πριν το modelling, ανάλογα με τις ιδιαιτερότητες του ίδιου του dataset αλλά και του classifier.

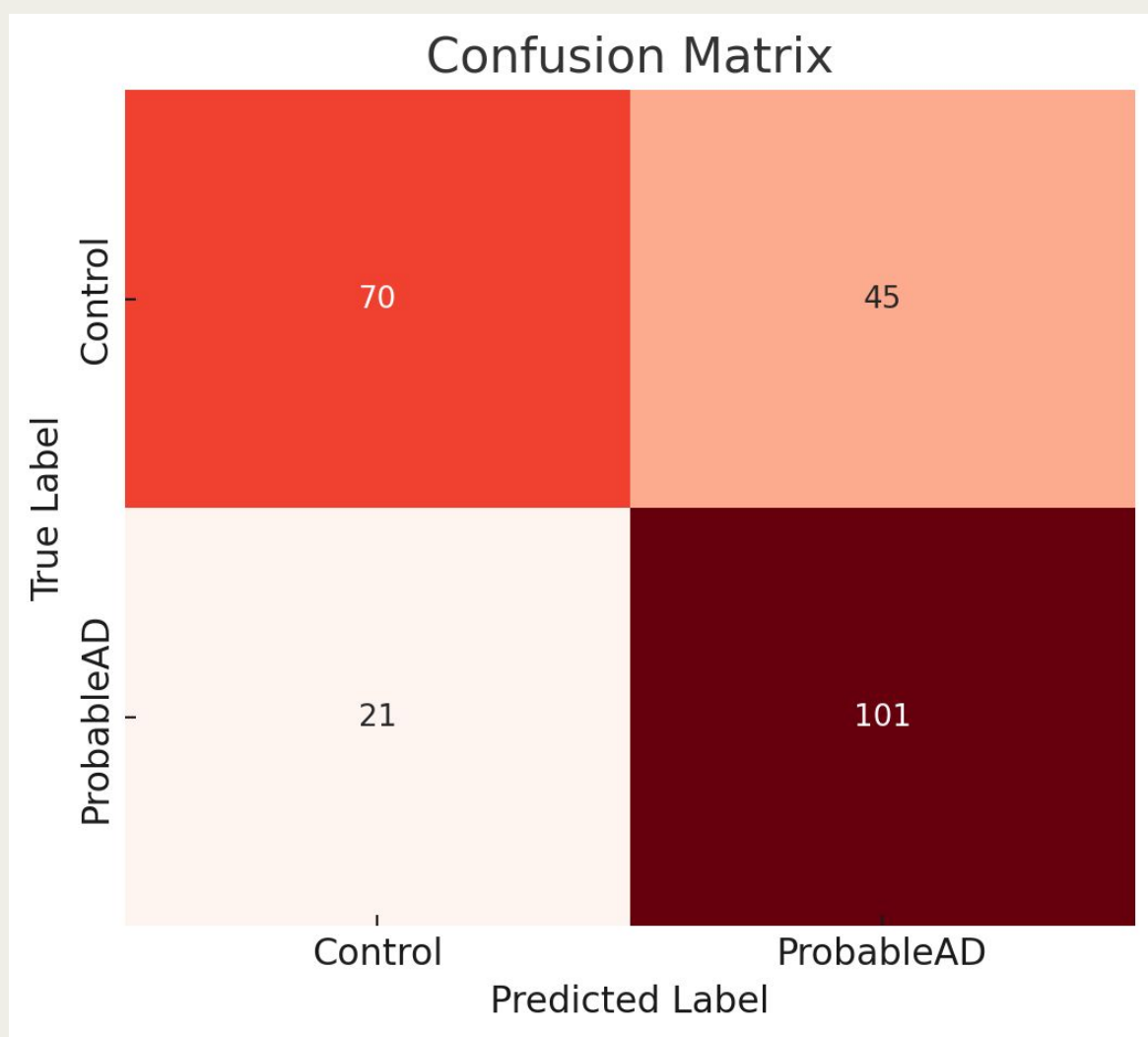
FUSED DATASET - SVM - RBF KERNEL

TRAIN-TEST SPLIT
test_size=0.3, random_state=16



Accuracy: 0.77
F1 Score Control: 0.73
F1 Score ProbableAD: 0.8

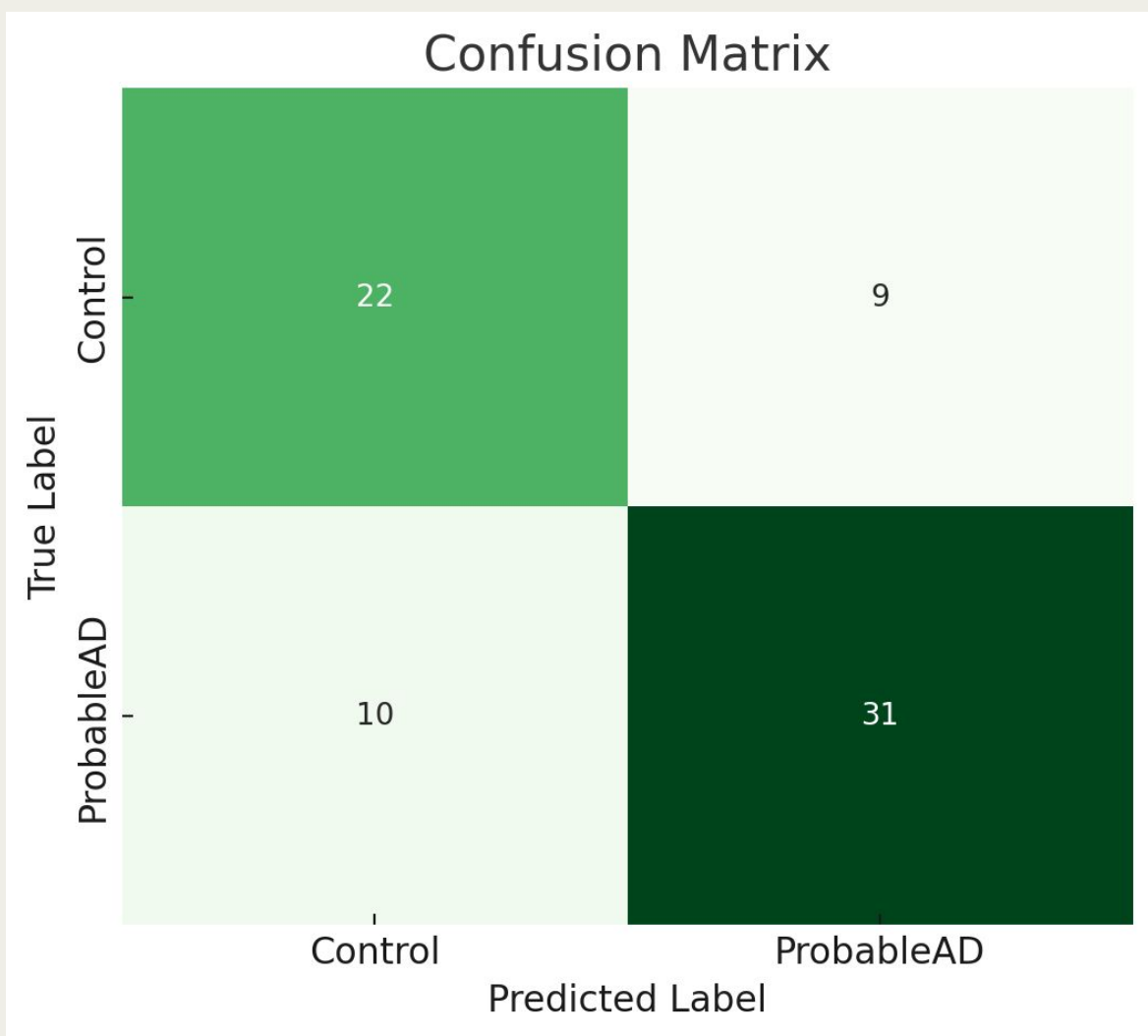
K-FOLD
KFold(n_splits=10, shuffle=True, random_state=42)



Mean Accuracy: 0.72
F1 Score Control: 0.68
F1 Score ProbableAD: 0.75

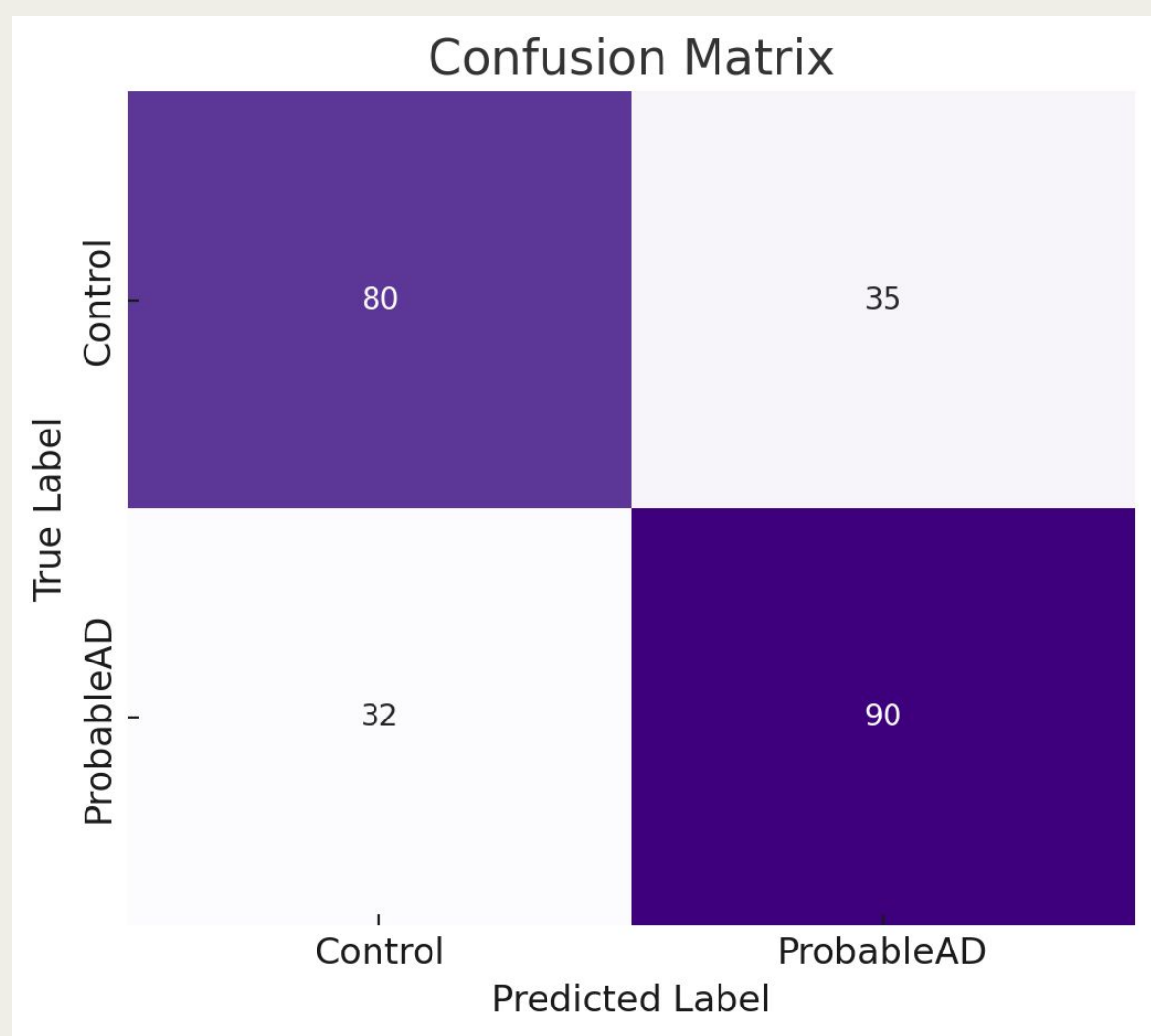
FUSED DATASET - LDA - EIGEN SOLVER

TRAIN-TEST SPLIT
test_size=0.3, random_state=16



Accuracy: 0.736
F1 Score Control: 0.698
F1 Score ProbableAD: 0.765

K-FOLD
KFold(n_splits=10, shuffle=True, random_state=42)



Mean Accuracy: 0.716
F1 Score Control: 0.704
F1 Score ProbableAD: 0.728

ENSEMBLE VOTING APPROACH

Στην περίπτωση του majority vote , αποδείχθηκε αρκετά πιο δύσκολο να εξαχθούν αποδοτικά μοντέλα για κάθε διαφορετικό dataset. Τα mfcc features πάσχιζαν να ξεφύγουν από την τυχαιότητα κυρίως εξαιτίας του μικρού correlation με τις διαγνώσεις μένοντας γύρω στο 60%, ενώ τα text features πλησίαζαν στο 70%.

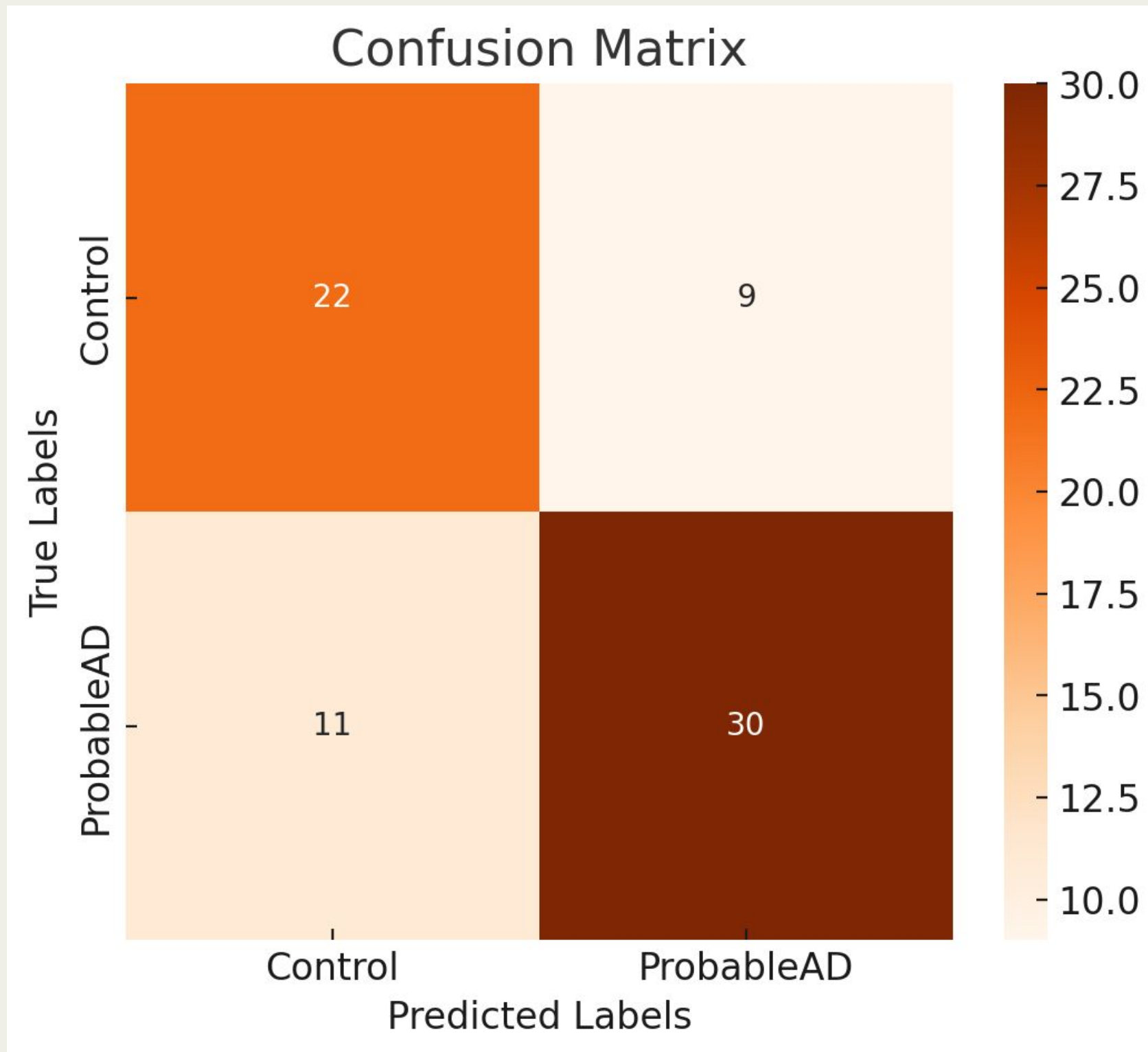
Από την άλλη πλευρά το ground truth dataset έχοντας αρκετά highly correlated features με τις διαγνώσεις, έδειχνε καλύτερα αλλά όχι κάτι που ξέφευγε εύκολα από το 72-73%.

Παράλληλα η μέθοδος των K-folds έδειχνε να μειώνει το accuracy ακριβώς όπως και στα fused datasets.

Για την μέθοδος χρησιμοποιήθηκαν **SVM** classifiers για τα mfcc, **Logistic Regression** για το groundtruth dataset και **LDA** για τα text features.

ENSEMBLE VOTING APPROACH - TRAIN-TEST SPLIT

test_size=0.3, random_state=16



Accuracy: 0.722
F1 Score Control: 0.69
F1 Score ProbableAD: 0.75

MFCC top 5 classifier:

model3 = svm.SVC(kernel='rbf', C=0.8)

Training Groundtruth classifier:

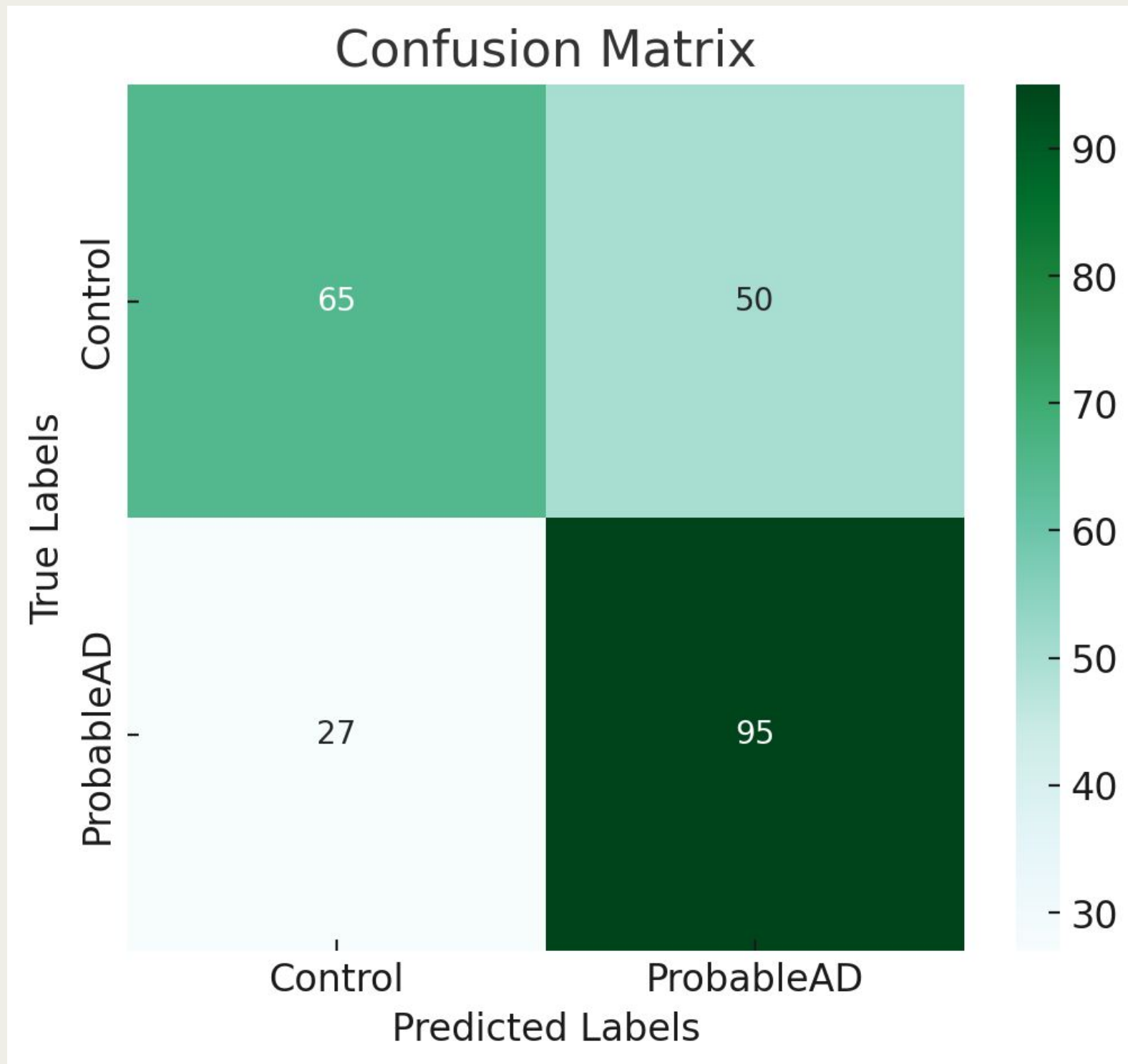
*LogisticRegression(C=0.1, penalty='l2', max_iter=100,
class_weight={'Control': 1, 'ProbableAD': 1.24})*

Text Features classifier:

*LinearDiscriminantAnalysis(solver='eigen',
shrinkage=1)*

ENSEMBLE VOTING APPROACH - K-FOLD

(*n_splits=10, shuffle=True, random_state=42*)



Accuracy: 0.675
F1 Score Control: 0.627
F1 Score ProbableAD: 0.705

MFCC top 5 classifier:

model3 = svm.SVC(kernel='rbf', C=0.8)

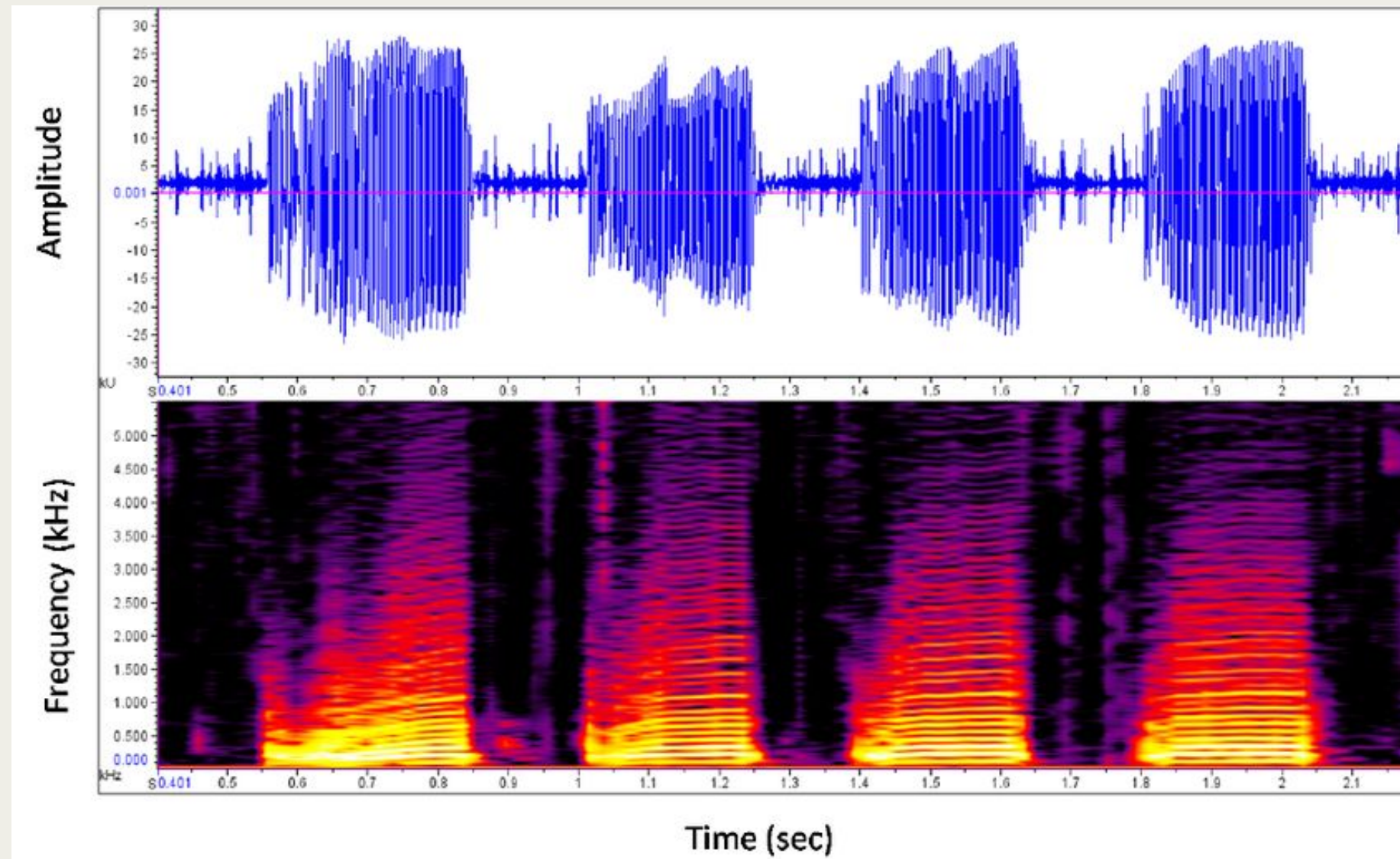
Training Groundtruth classifier:

*LogisticRegression(C=0.1, penalty='l2', max_iter=100,
class_weight={'Control': 1, 'ProbableAD': 1.24})*

Text Features classifier:

*LinearDiscriminantAnalysis(solver='eigen',
shrinkage=1)*

IMPROVEMENTS - SPECTOGRAMS



Εξαγωγή spectrogram από τα audios και στη συνέχεια επεξεργασία εικόνας ώστε να εξαχθούν νέα features και να συμπεριληφθούν σε majority vote.

Thank you!
