

Detecting Alzheimer's Through Speech

An Audio-Based Classification Approach

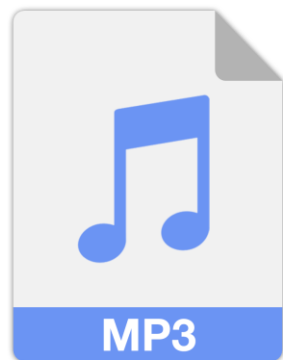
E. Makrimanolaki

T. Giannakopoulos

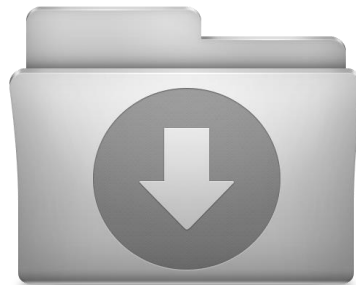




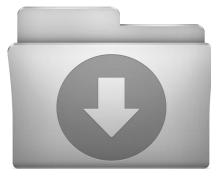
Speech-Based Alzheimer Classification



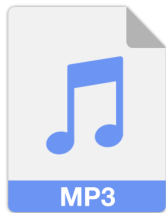
&



Inside the Engine



- pandas
- numpy
- matplotlib.pyplot
- seaborn
- assemblyai
- os
- Nltk
- mutagen



○ 45"-2'

{52% Alzheimer, 48% Control}
Groundtruth dataset



Features

MFCC

- 137 features
- mel-frequency από audio
- Διάρκεια ανάλυσης: Ανά 15"
- pyAudioAnalysis

Groundtruth

- Dataset με 4 features
- Metadata που περιλαμβάνουν δημογραφικές και κλινικές πληροφορίες
 - MMSE Score
 - Gender
 - Age
 - Education

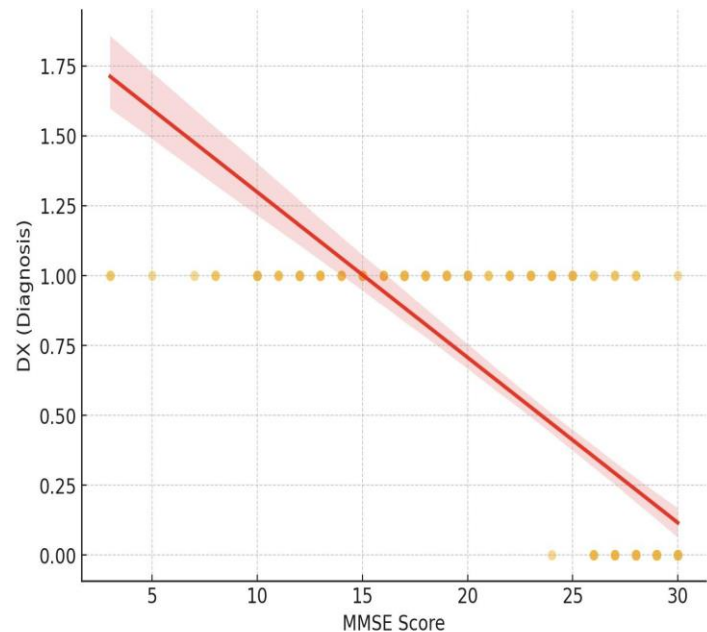
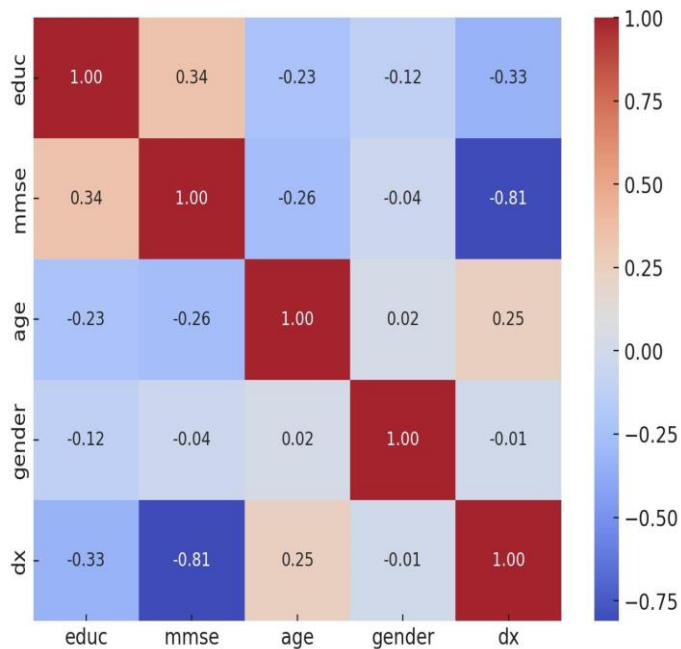
Audio to Text to Features

- 2 features
- Εξαγωγή και ανάλυση κειμένου από audio
- ΠΩΣ;
 - Μετατροπή ομιλίας σε κείμενο με τη βιβλιοθήκη assemblyai
 - Tokenization του κειμένου σε λέξεις
 - Εξαγωγή γλωσσικών χαρακτηριστικών, όπως lexical richness και word variance, με τη βιβλιοθήκη nltk

Data Cleaning & Feature Insights

- Από την εξαγωγή των features, προέκυψαν νέα datasets:
 - **Correlation matrices & linearity analysis** χρησιμοποιήθηκαν για την καλύτερη κατανόηση των δεδομένων και την εξαγωγή πιθανών σχέσεων μεταξύ μεταβλητών
 - Το **training ground truth dataset** έδειξε **συσχέτιση** μεταξύ του **MMSE** και του **educ** σε σχέση με τη διάγνωση της νόσου **Alzheimer**
- **Data Cleaning**
 - Το **ground truth dataset** καθαρίστηκε από **null values**
 - Τα **categorical values** του χαρακτηριστικού gender μετατράπηκαν σε δυαδική μορφή (0 και 1)
- **MFCC features**
 - **Ελάχιστα** (από τα 137 χαρακτηριστικά) εμφάνισαν γραμμική σχέση με το ground truth
 - Αυτό επιβεβαιώθηκε και μέσω **importance analysis** με χρήση του RandomForestClassifier από τη βιβλιοθήκη sklearn
- **Text Features**
 - Τα δύο χαρακτηριστικά κειμένου ανέδειξαν φαινόμενο γραμματικότητας, το οποίο ήταν ιδιαίτερα έντονο στο χαρακτηριστικό **Hapax Legomena**

From Patterns to Predictions

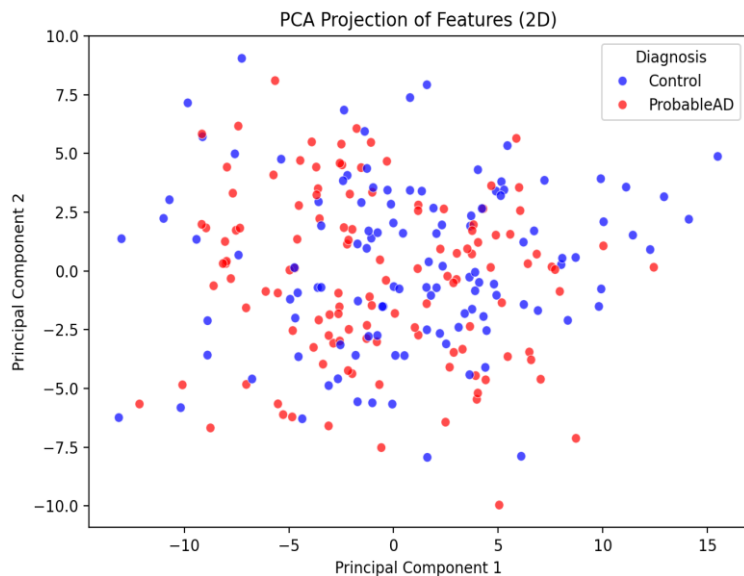


Finalized Data: Fusions&Extraction

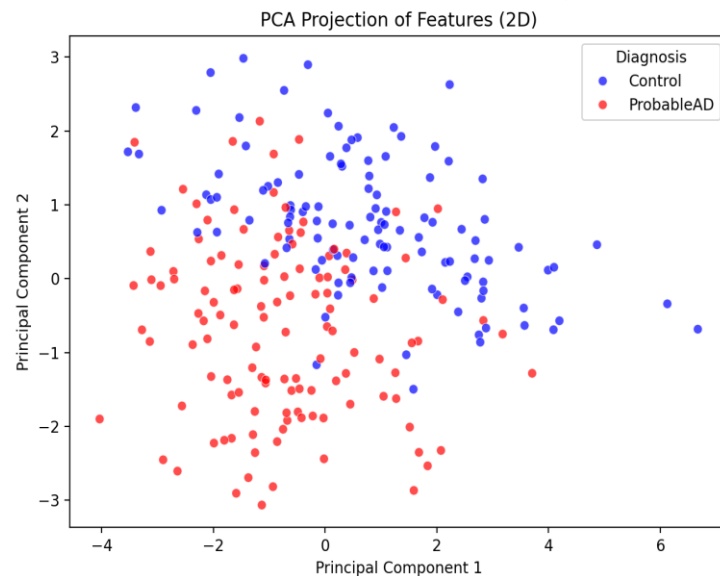
- Fuse όλων των χαρακτηριστικών
 - MFCC + Text Features + Ground Truth → (137 + 2 + 4 features)
 - Onset detection: Identifies speech start and stop points
 - Pause duration: Measures silence length in speech
 - Μεμονωμένες Ομάδες Χαρακτηριστικών
 - MFCC: (137 features)
 - Text Features: (2 features)
 - Ground Truth: (4 features)
 - Επιλογή Βέλτιστων MFCC Χαρακτηριστικών
 - Top 5 MFCC Features → (5 features)
 - Top 10 MFCC Features → (10 features)
 - Συγχώνευση Επιλεγμένων Χαρακτηριστικών
 - Fuse (MFCC top 5 + Text Features + Ground Truth) → (5 + 2 + 4 features)
- Για εύκολη επαναχρησιμοποίηση στους classifiers:
Τα σύνολα δεδομένα εξήχθησαν σε CSV

PCA Variance & Data Distribution

MFCC Features



Top 5 MFCC Features

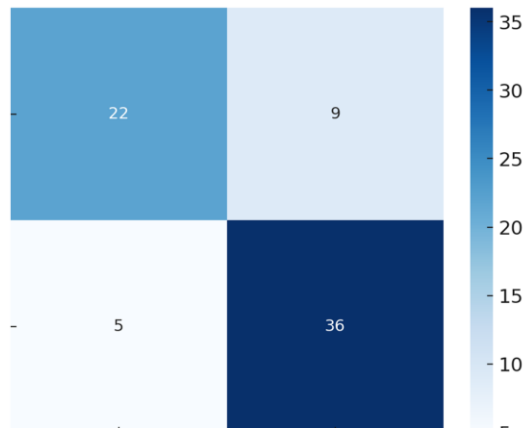


Candidate Classifiers & Selection Strategy

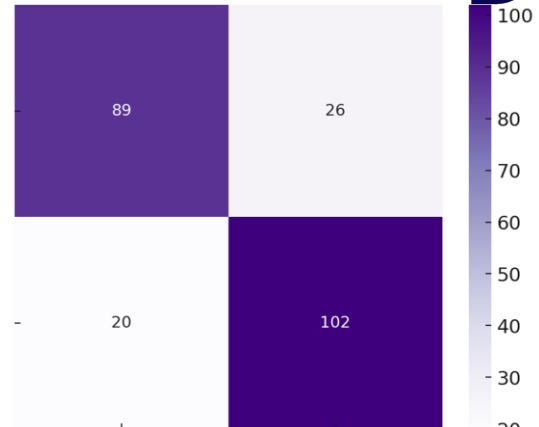
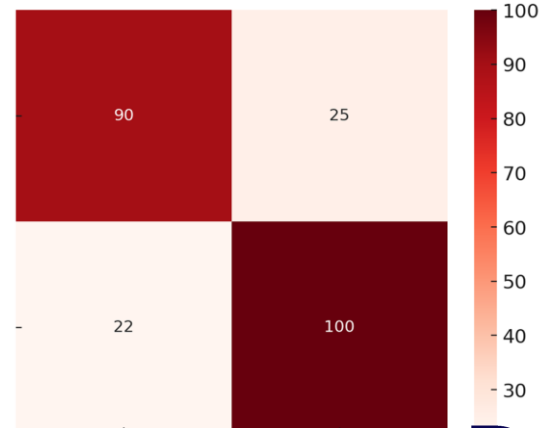
Classifier	Χαρακτηριστικά & Χρήση
Support Vector Machines (SVM)	Ιδανικό για εύρεση του βέλτιστου υπερεπιπέδου σε δεδομένα υψηλών διαστάσεων
Logistic Regression & LDA	Χρησιμοποιούνται όταν υπάρχει γραμμική σχέση στα χαρακτηριστικά. Χρησιμεύουν ως baseline για αρχικές δοκιμές
Ensemble Voting Classifier	Συνδυάζει διαφορετικούς classifiers για MFCC, text features και ground truth. Χρησιμοποιεί majority voting (2/3), π.χ., SVM για MFCC, LDA για text features, Logistic Regression για ground truth
Άλλοι Classifiers	Δοκιμάστηκαν για λόγους σύγκρισης και ευρύτερων τεστ

○ Κάθε dataset υποβάλλεται σε normalization ή standardization πριν από το modelling, ανάλογα με τις ιδιαιτερότητές του και τις απαιτήσεις του κάθε classifier

Fused

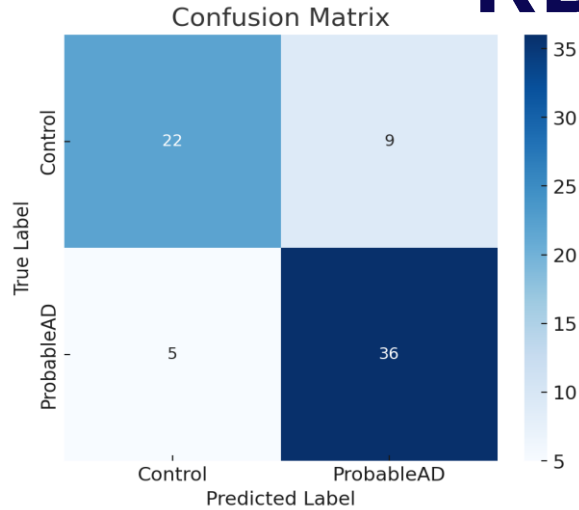


Dataset

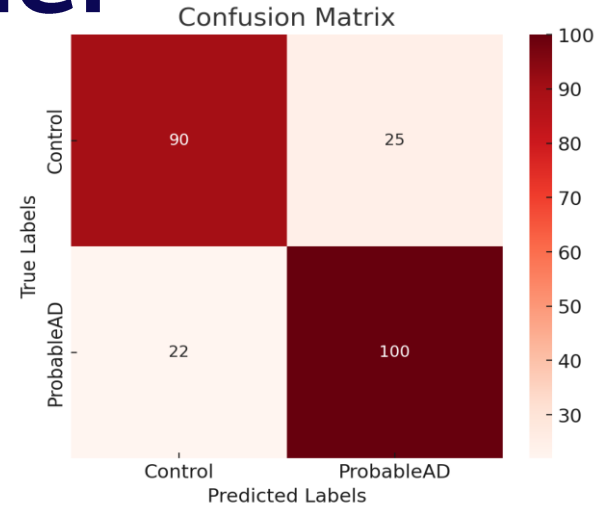


SVM & RBF Kernel

TRAIN-TEST SPLIT
test_size=0.3, random_state=16



Metric	Value
Accuracy	0.8
F1 Score Control	0.758
F1 Score Probable AD	0.837

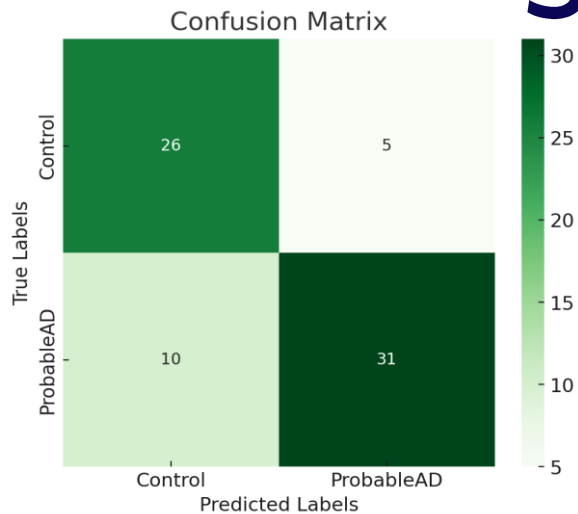


Metric	Value
Mean Accuracy	0.801
F1 Score Control	0.79
F1 Score Probable AD	0.809

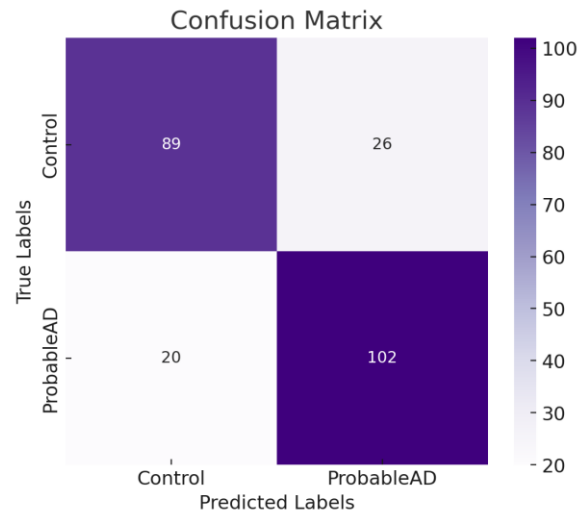
K-FOLD
KFold(n_splits=10, shuffle=True, random_state=42)

LDA & Eigen Solver

TRAIN-TEST SPLIT
test_size=0.3, random_state=16



Metric	Value
Accuracy	0.79
F1 Score Control	0.776
F1 Score Probable AD	0.805



Metric	Value
Mean Accuracy	0.805
F1 Score Control	0.79
F1 Score Probable AD	0.816

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

Ensemble Voting Strategy

Η χρήση **majority voting** αποδείχθηκε αρκετά δύσκολη, καθώς δεν ήταν εύκολο να εξασφαλιστεί η αποδοτική απόδοση των μοντέλων σε κάθε διαφορετικό dataset.



Προσέγγιση με Weighting

Δίνεται μεγαλύτερη βαρύτητα στο ground truth dataset, ακόμη κι αν επιτρέπει ένα μικρό overfit.
Η προσέγγιση αυτή βοηθά στη μείωση των λάθος προβλέψεων, διορθώνοντας το impact των χαμηλότερων accuracies των υπόλοιπων μοντέλων.

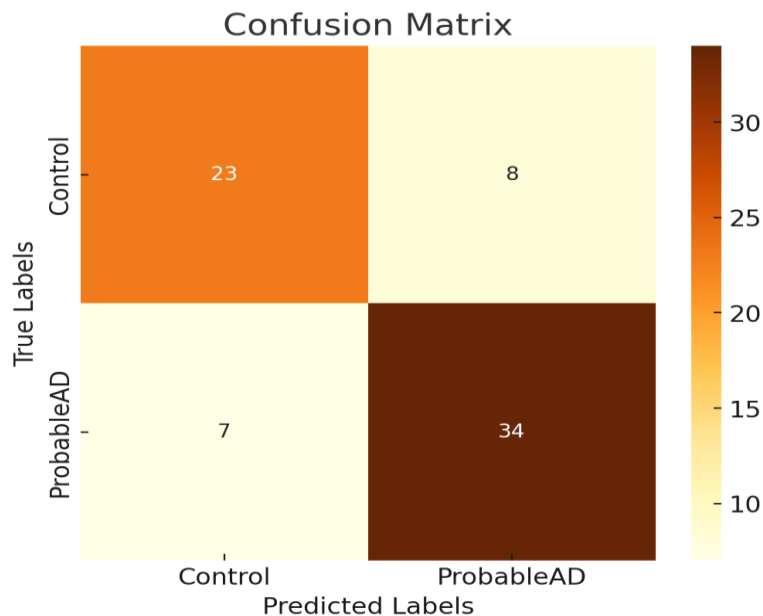


- Τα MFCC features παρουσίασαν χαμηλό correlation με τις διαγνώσεις, με accuracy περίπου 60%, καθιστώντας δύσκολη τη διάκριση μεταξύ των κλάσεων.
- Τα text features εμφάνισαν ελαφρώς καλύτερη απόδοση (~70%), αλλά εξακολουθούσαν να έχουν περιορισμένη διακριτική ικανότητα.
- Το ground truth dataset, από την άλλη, περιείχε ισχυρά συσχετισμένα features με τις διαγνώσεις, οδηγώντας σε πιθανό overfitting.

- SVM για MFCC και Ground Truth datasets
- LDA για Text Features, καθώς διαχειρίζεται καλύτερα τη γραμμικότητα των δεδομένων

Ensemble Voting Strategy & Train-Test Split Results

test_size=0.3, random_state=16



MFCC Top 5 Classifier

- SVM μ RBF kernel, $C=0.8$
- `model3 = svm.SVC(kernel='rbf', C=0.8)`

Training Groundtruth Classifier

- SVM μ RBF kernel, $C=0.05$
- `svm.SVC(kernel='rbf', C=0.05)`

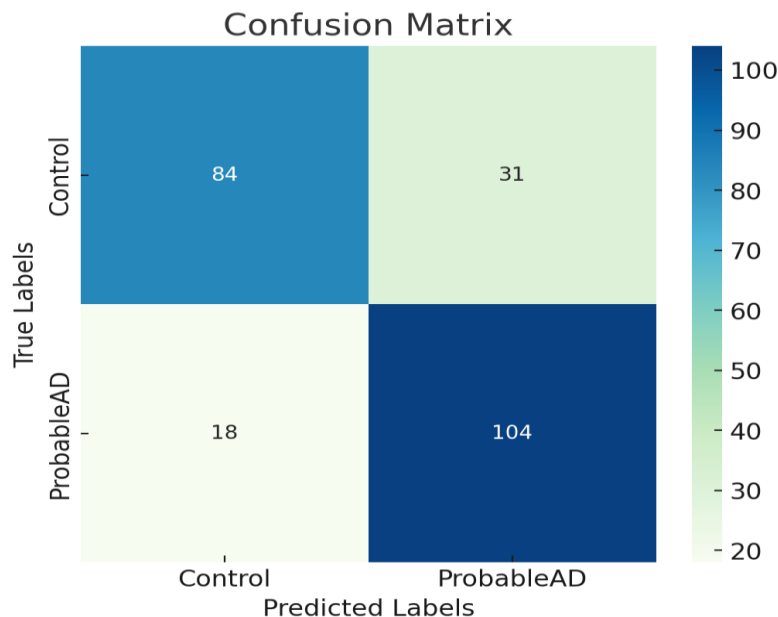
Text Features Classifier

- LDA μ eigen solver
- `LinearDiscriminantAnalysis (solver='eigen', shrinkage=1)`

Metric	Value
Accuracy	0.791
F1 Score Control	0.754
F1 Score Probable AD	0.819

Ensemble Voting Strategy & K-Fold

n_splits=10, shuffle=True, random_state=42



MFCC Top 5 Classifier

- SVM μ ε RBF kernel, C=0.8
- `model3 = svm.SVC(kernel='rbf', C=0.8)`

Training Groundtruth Classifier

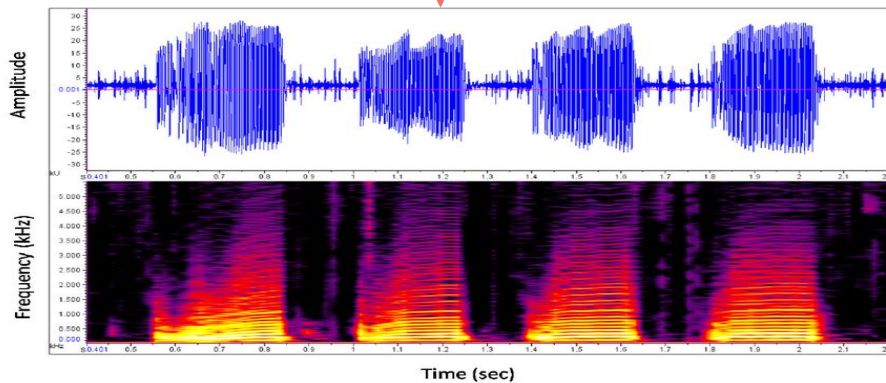
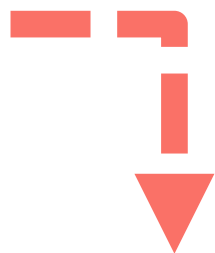
- SVM μ ε RBF kernel, C=0.05
- `svm.SVC(kernel='rbf', C=0.05)`

Text Features Classifier

- Linear Discriminant Analysis (LDA) μ ε eigen
- `solverLinearDiscriminantAnalysis (solver='eigen', shrinkage=1)`

Metric	Value
Accuracy	0.793
F1 Score Control	0.77
F1 Score Probable AD	0.8

From Today to Tomorrow...



Thank you!

