UNIVERSITAT DE BARCELONA

SMALL PROJECT

CLASSIFICATION OF ECG SIGNALS

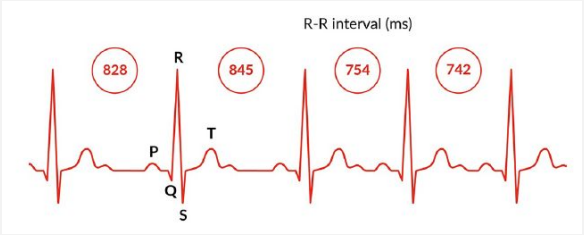
Biomedical Engineering - Biomedical Signals

Sergi Marsol

INTRODUCTION AND MOTIVATION

An important goal ⟶ automatic detection of arrhythmias

Atrial Fibrillation is defined as “Tachyarrhythmia characterized by predominantly uncoordinated atrial activation with consequent deterioration of atrial mechanical function”



One method for AF detection is to analyse the HRV signal, based on the detection of the R peak.
Excessive HRV ⟶ correlated with atrial fibrillation

GOALS AND METHODS OF THE PROJECT

GOAL 1

Separate CLEAN ECG from NOISY ECG

- Read ECG files using Matlab
- Separate CLEAN from NOISY signals

GOAL 2

Design and program our own R wave detector and extract main features

- Pre-Process ECG signals
- Find QRS peaks
- Extract features from original and RR signals
- Create a table of features

GOAL 3

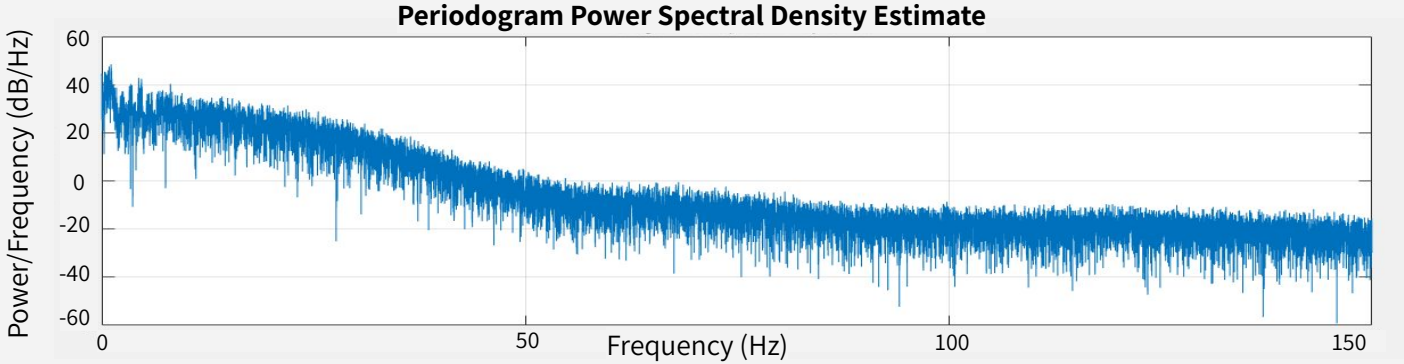
Provide a classifier for the 4 classes of the clean signal

- Build a classifier function
- Validate the predictions for the 4 classes and evaluate the results of the classifier

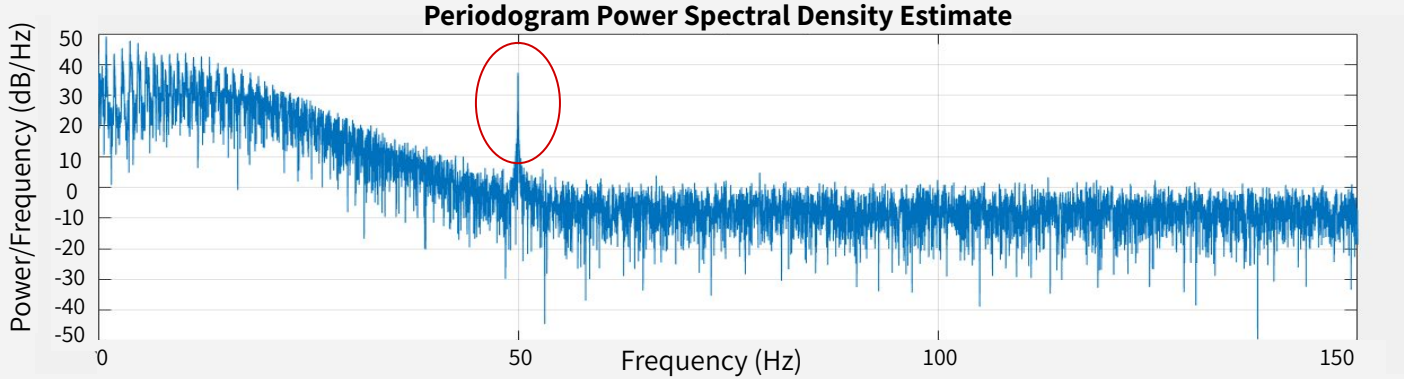
GOAL 1: Separation of clean and noisy signals

PERIODOGRAM

Clean signals



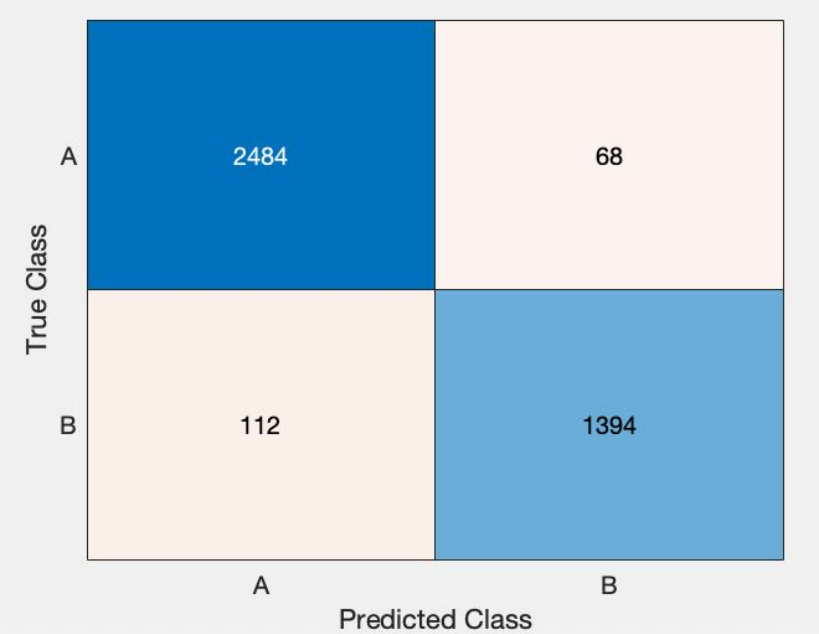
Noisy signals



CLASSIFIERS

Clean and noisy signals separation using 1 feature→ BANDPOWER at **frequency 50 Hz**

1. K-NEAREST NEIGHBORS



Hyperparameter optimization

distance: ***chebyshev***

k = 17

➔ **ACCURACY = 0.9556**

Class	Precision	Recall
A	0.957	0.973
B	0.954	0.926

Total F-Score = 0.952

2. CLASSIFICATION TREE

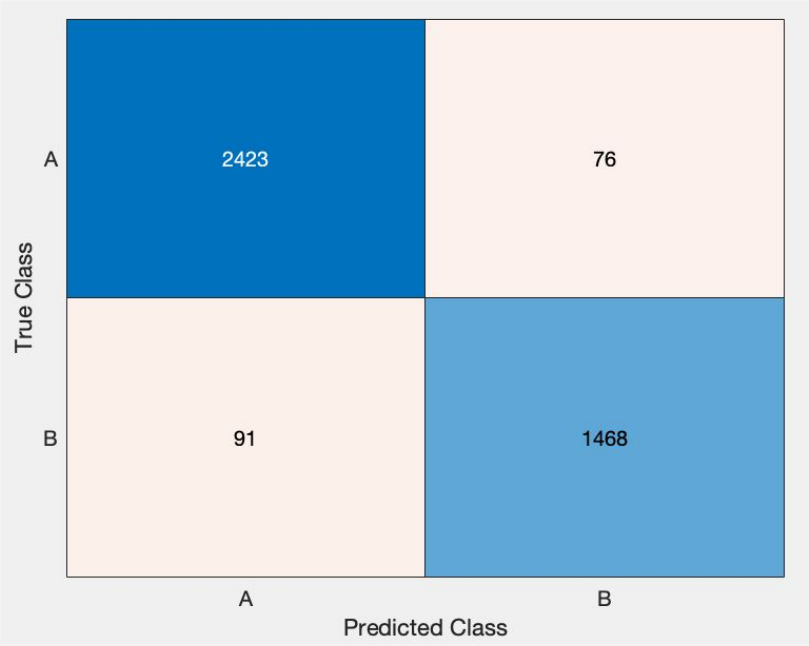
Hyperparameter optimization

MinLeafSize = 78

➔ **ACCURACY = 0.957**

Class	Precision	Recall
A	0.964	0.970
B	0.951	0.942

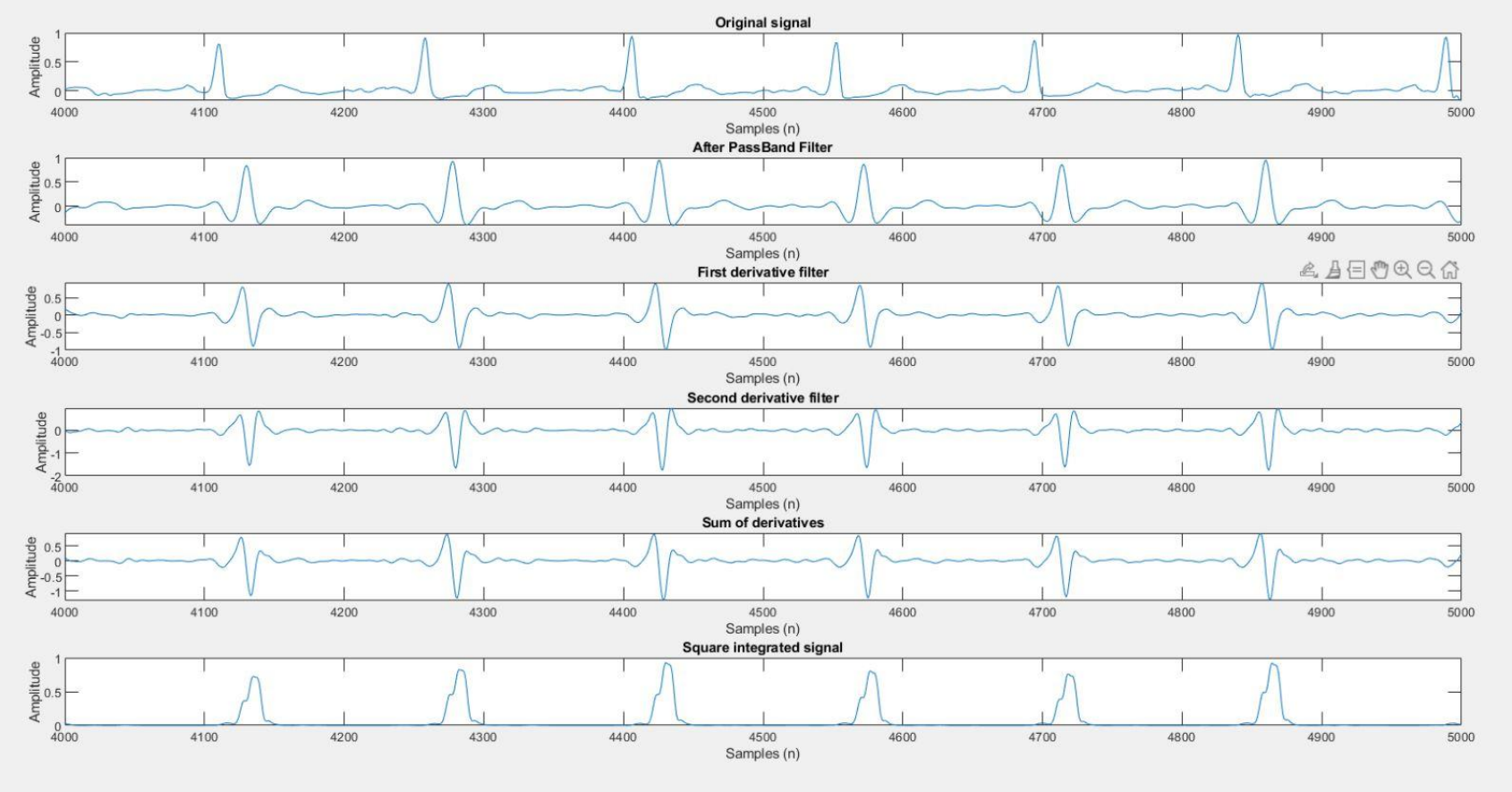
Total F-Score = 0.957



GOAL 2: R wave detector and main features extraction

PAN-TOMPKINS ALGORITHM

- Original signal
- Bandpass filter
- First derivative
- Second derivative
- Sum of derivative
- Square Integrator



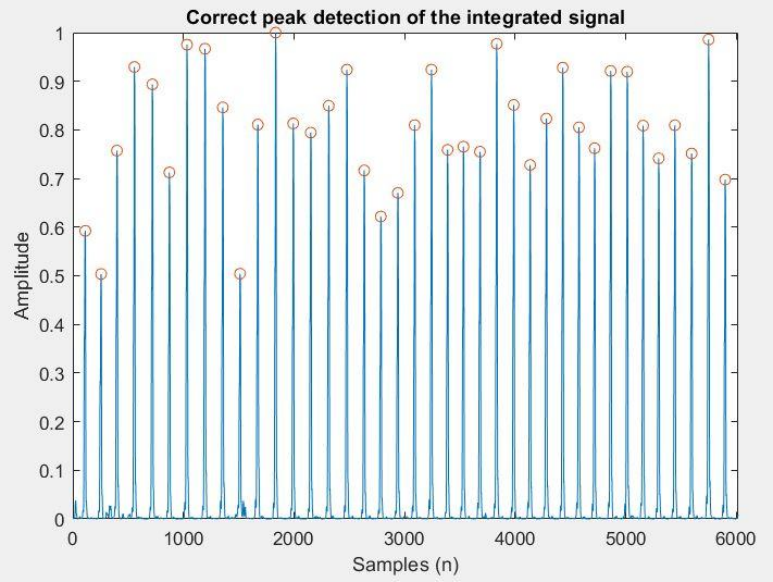
QRS PEAK DETECTOR

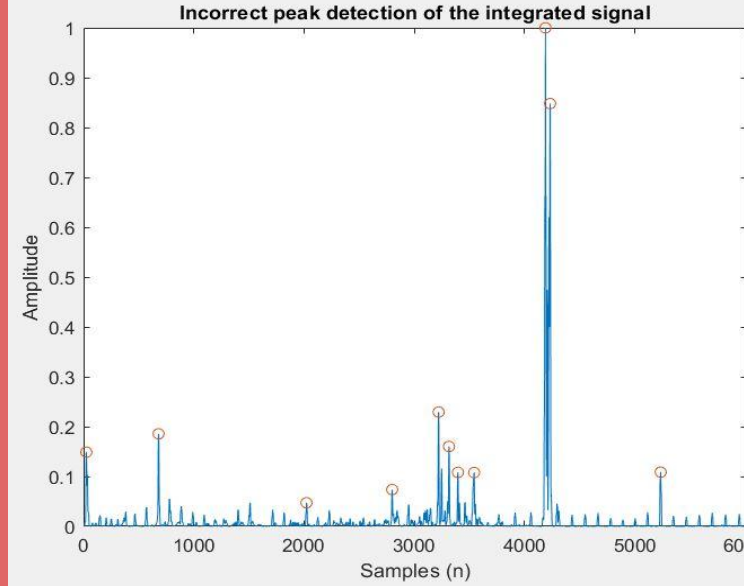
Thr1 = **npk1** + 0,25(**spk1** - **npk1**)

Thr2 = 0,5**Thr1**

spk1 = 0,125**peak1** + 0,875**spk1**

npk1 = 0,125**peak1** + 0,875**npk1**





FEATURE EXTRACTION

ECG signal

- Mean *mA*
- Standard deviation *sdA*
- Outlier fraction *outlierA*

Autocorrelation

- Mean *autoA1*
- Standard deviation *autoA2*
- Maximum value *autoA3*

Bandpower

- Bp [0,5-5]Hz *bp1*
- Bp [5-15]Hz *bp2*
- Bp [15-30]Hz *bp3*
- Bp [30-45]Hz *bp4*

- Number of Peaks *num_pks*
- Peaks detected correctly *pks_ok*

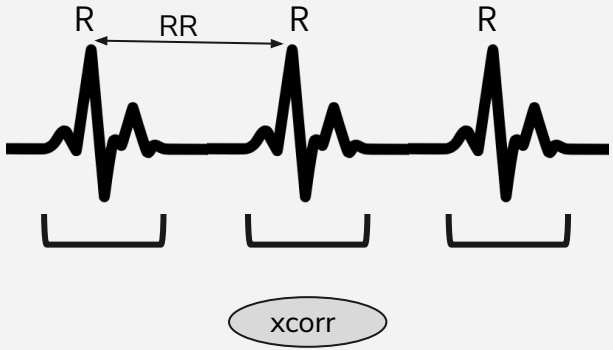
RR signal

- Mean *mRR*
- Standard deviation *sdRR (HRV)*
- Outlier fraction *outlierRR*

- Standard deviation of successive differences *sdsd*

Cross-correlation

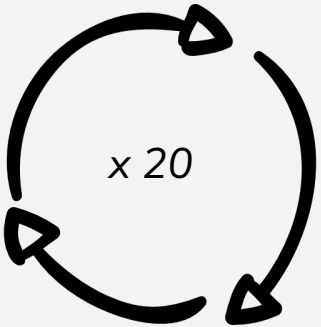
- Correlation of the length of consecutive beats *corrlenNN*



FEATURE EXTRACTION

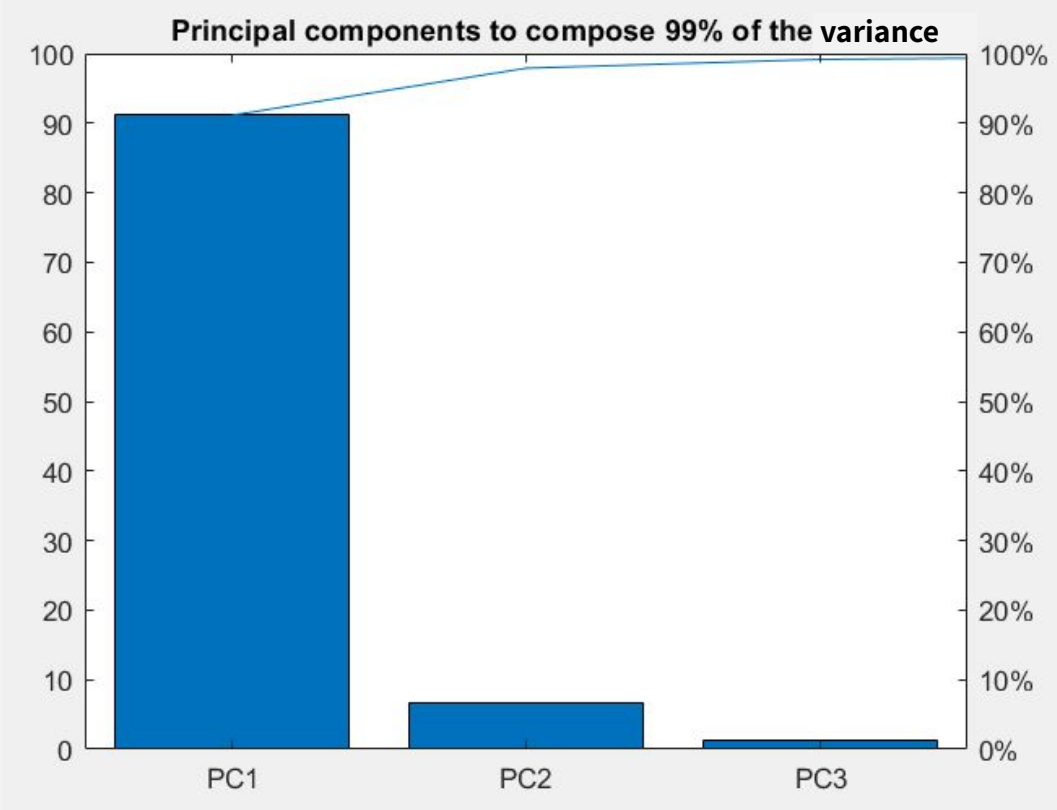
RR signal

- pNN20
- pNN50



pNN10
pNN20
pNN30
...
pNN200

- PC1 *PC1_pNN*
- PC2 *PC2_pNN*
- PC3 *PC3_pNN*

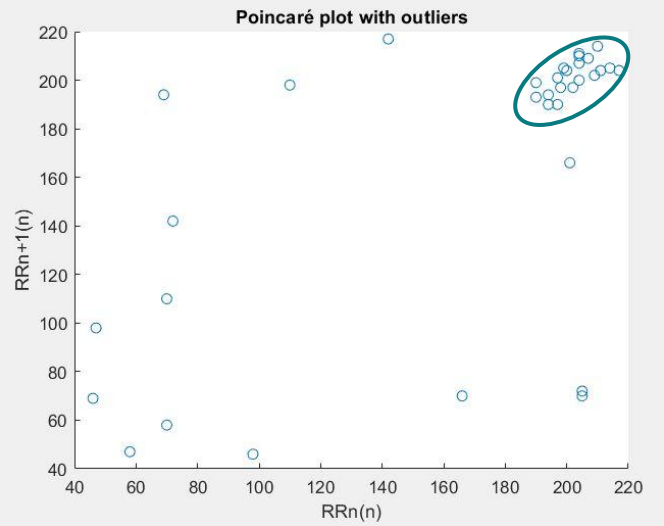


FEATURE EXTRACTION

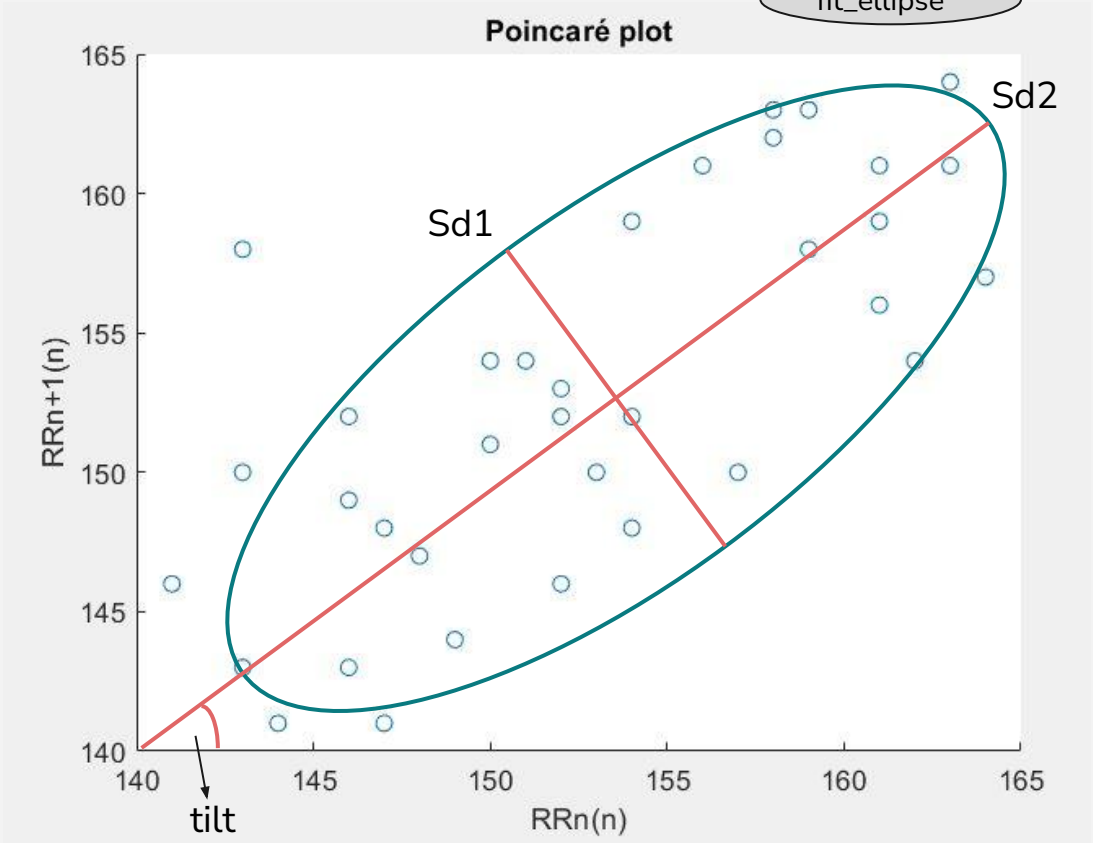
Poincaré

Scatter plot of nn-1 vs nn

- Sd1
- Sd2
- tilt



fit_ellipse



FEATURE EXTRACTION

ANOVA

Analysis of variance

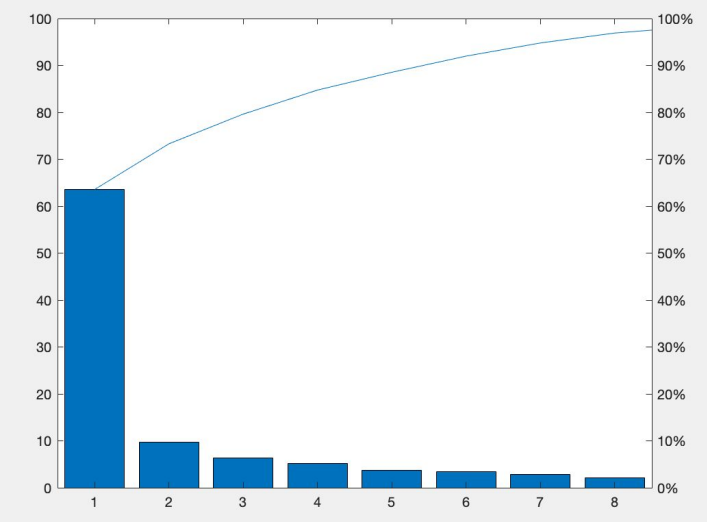
Features	P-value	Features	P-value	Features	P-value
PC1	0	*corrlenNN*	1,8126e-40	*bp3*	0,0057
PC3	1,2338e-146	*sd1*	2,3819e-29	*bp2*	0,0088
PC2	3,0789e-116	*mRR*	2,1524e-17	*tilt*	0,0105
outlierA	3,6195e-97	*bp1*	1,1365e-11	*autoA3*	0,0450
sdsd	2,3383e-73	*autoA1*	1,5348e-07	*bp4*	0,8737
sdRR	1,1718e-65	*autoA2*	1,3441e-05		
pks_ok	1,2913e-62	*mA*	4,2079e-04		
sd2	5,5304e-51	*num_pks*	4,7556e-04		
outlierRR	2,2740e-43	*sdA*	0,0056		

Lower p-value ✓

Higher p-value ✗

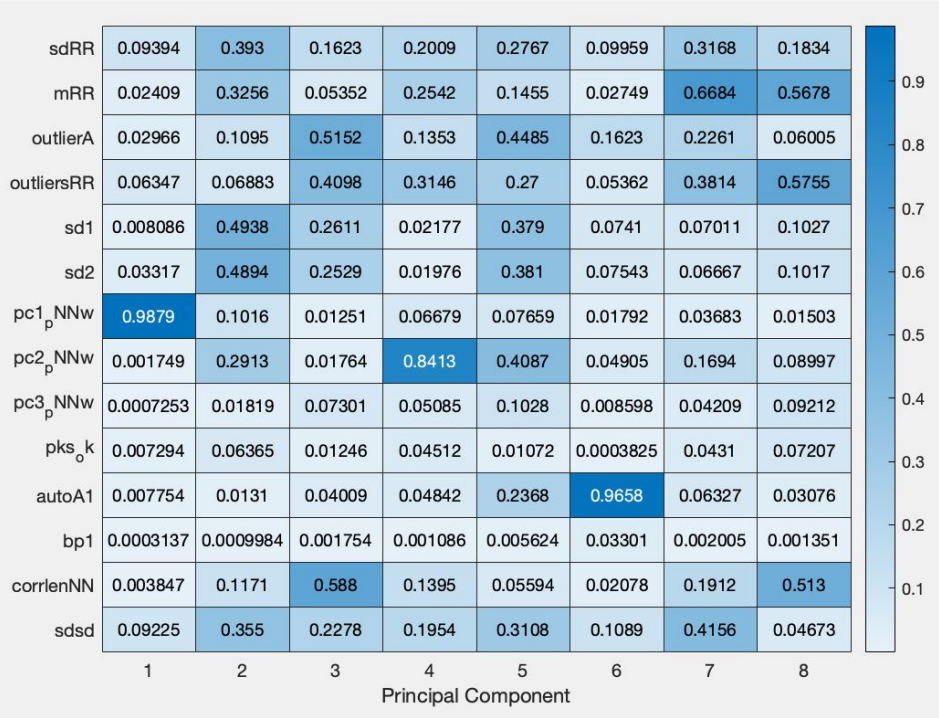
GOAL 3: Build a classifier for all 4 clean ECG recordings

1. PRINCIPAL COMPONENT ANALYSIS (PCA)

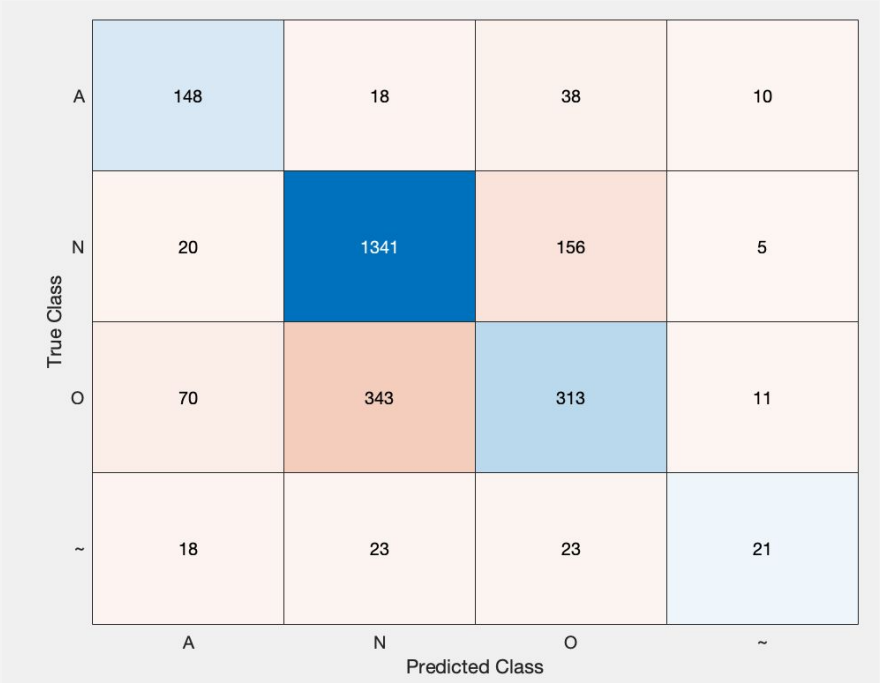


Principal components explaining 95% of variance

We will build the classification model using **8 principal components**.



2. K-NEAREST-NEIGHBOR (KNN)



Hyperparameter optimization

distance: *cityblock*

k = 14

➔ **ACCURACY = 0.7217**

Class	Precision	Recall
AF	0.578	0.692
Normal	0.777	0.881
Other	0.591	0.425
Noisy	0.447	0.247

Total F-Score = 0.567

3. TREE CLASSIFIER

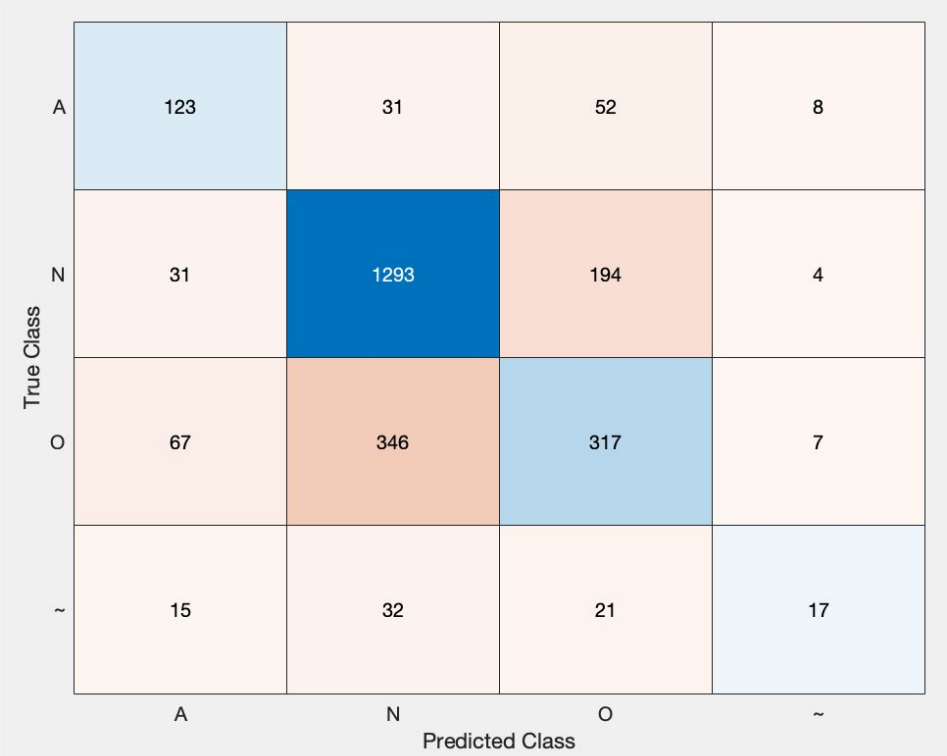
Hyperparameter optimization

MinLeafSize = 49

➔ **ACCURACY = 0.6841**

Class	Precision	Recall
AF	0.521	0.575
Normal	0.76	0.85
Other	0.543	0.43
Noisy	0.472	0.2

F-Score = 0.528



CONCLUSIONS

GOAL 1 → We have been able to separate clean from noisy ECG using the **BANDPOWER** as the main feature of the classifiers.

GOAL 2 → We have performed an analysis in amplitude and frequency of the ECG signals and have extracted **14 features**.

↓ PCA has allowed dimensionality reduction from 14 features to 8 principal components.

GOAL 3 → We have built the classification models to **separate all clean ECG classes** using only 8 features and kNN is giving the best results.