Heartbeat audio classification

Advanced Biomedical
Machine Learning project
University of Pavia

Andrea Alberti
Davide Ligari

TABLE OF CONTENTS

INTRODUCTION

1st

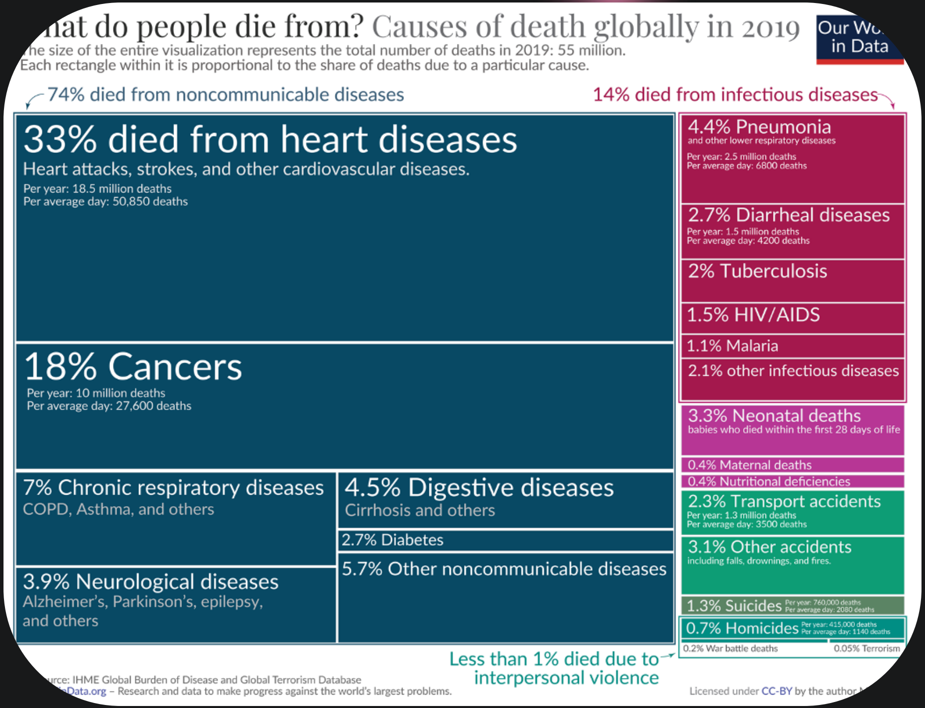
Death cause in the world

50.850

Average deaths per day

18,5 M

Death per year



STATE OF ART

AUTHORS	MODELS	FEATURES	RESULTS	ANNO	DATASET
W. Zhang et al [1]	SVM	Spectrogram	0.76 Precision	2017	N, M, EH, AR
SW. Deng et al [2]	SVM	DWT	0.76 Precision	2016	N, M, EH, AR
A. Raza et al [3]	LSTM	1D time series	0.80 Accuracy	2019	N, M, ES
T. Alafif et al [4]	2D-CNN + transfer learning	MFCC	0.89 Accuracy	2020	N, A
Noman et al [5]	Ensemble CNN	1D time series + MFCC	0.89 Accuracy	2019	N, A
Chen et al [6]	2D CNN	WT + Hilbert-Huang	0.93 Accuracy	2018	N, M, ES
Our Model	Ensemble Model (MLPs + RF)	MFCC + Chroma + ZCR	0.88 Accuracy	2024	AR, M, N, EH, ES

LEGEND: AR: Artifacts, M: Murmurs, N: Normals, EH: Extra Heart Sound, ES: Extra Systole, A: Abnormal

OUR GOALS

PREVENTION

Providing an accessible method for everyone to identify potential cardiac cycle anomalies, enabling proactive intervention.

SUPPORT

Offering physicians enhanced support to accurately identify cardiac issues.

DATASET

SOURCE

kaggle

Dangerous Heartbeat Dataset

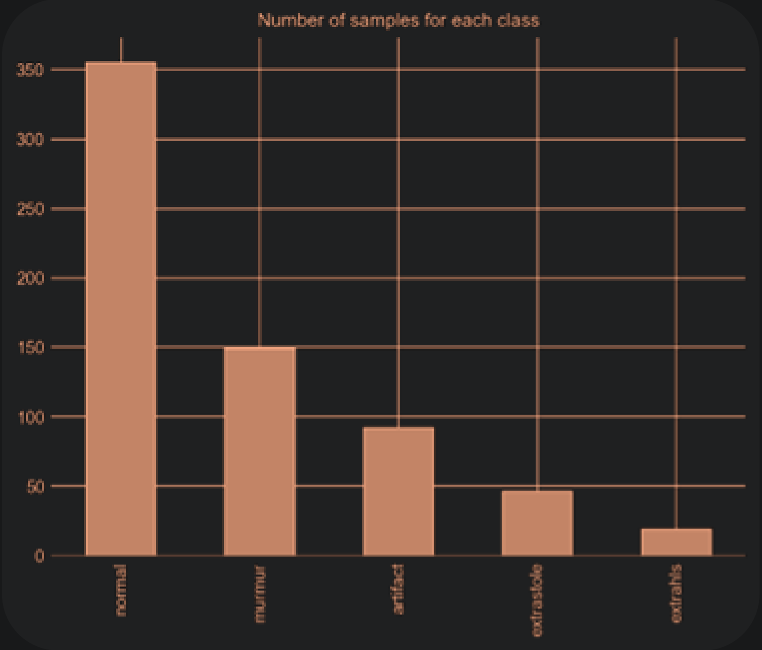
STRUCTURE

Heartbeat audio

Different sources (Type A, B and C)

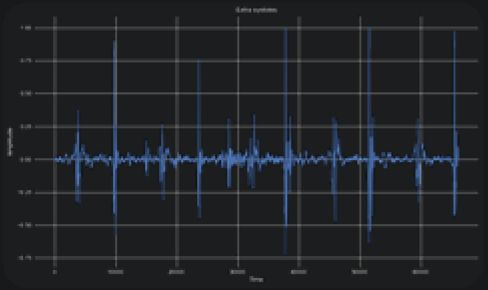
CLASS BALANCE

Highly imbalanced classes

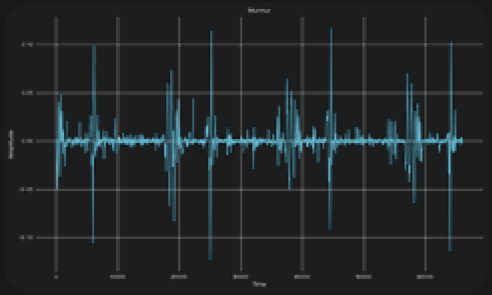


THE CLASSES

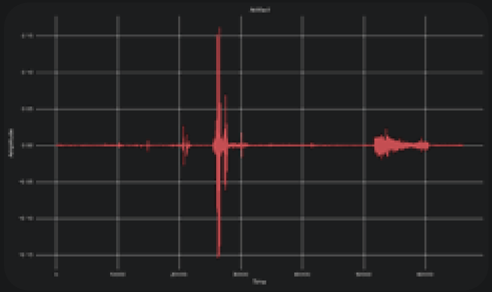
EXTRA SYSTOLES



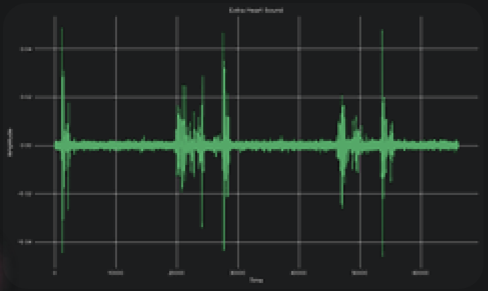
MURMUR



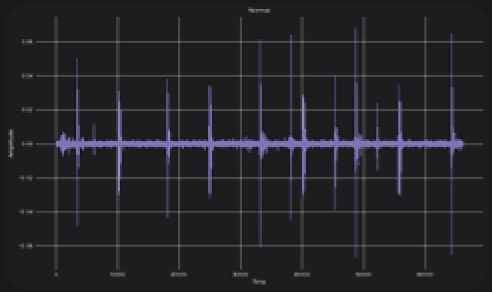
ARTIFACT



EXTRA HEARTBEAT



NORMAL



PREPROCESSING

NOISE REDUCTION

Audio clipped to reduce noise and irrelevant information

CORRUPTED FILES

Corrupted files were deleted

RESAMPLING

All files resampled to a common frequency of 4000 Hz

SEGMENTATION

Files clipped into a fixed length of 1 second

OUTLIERS

Long lasting audios have been removed

FEATURES

MFCC

MFFCs can reflect the different perceived quality of heart sounds, such as the presence of murmurs or other anomalies

Root Mean Square Energy

RMS is used for detecting volume changes and can help identify different types of heartbeats based on their energy levels

Zero Crossing Rate

ZCR can help differentiate between normal and abnormal sounds by highlighting changes in signal periodicity

Spectral Rolloff

Can be used to differentiate between sounds with a concentrated or dispersed energy distribution

Chroma STFT

Captures the 12 pitch classes, useful for identifying harmonic patterns in heartbeat audio, aiding in anomaly detection

Constant-Q Transform

Uses a logarithmic frequency scale, capturing more detail at lower frequencies, which is useful for analyzing the low-frequency components of heart sounds

Spectral Bandwidth

This feature helps in understanding the spread of the frequency components in the heart sounds

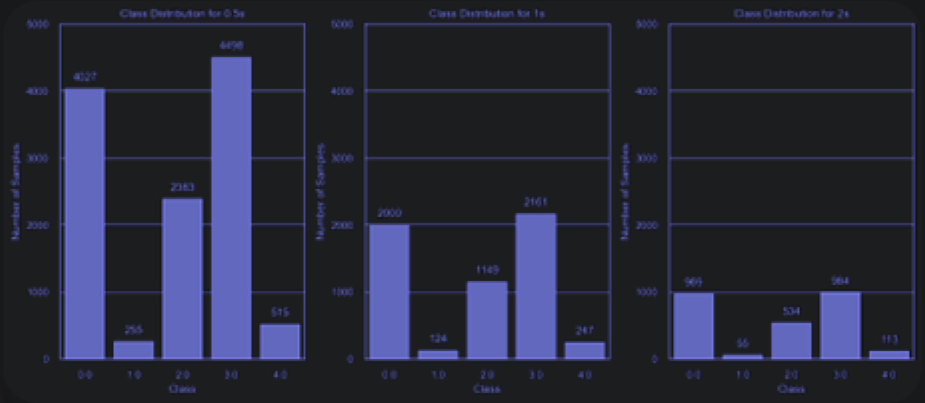
Spectral Centroid

Represents the 'brightness' of a sound by indicating the center of mass of the spectrum

FEATURES EXTRACTION

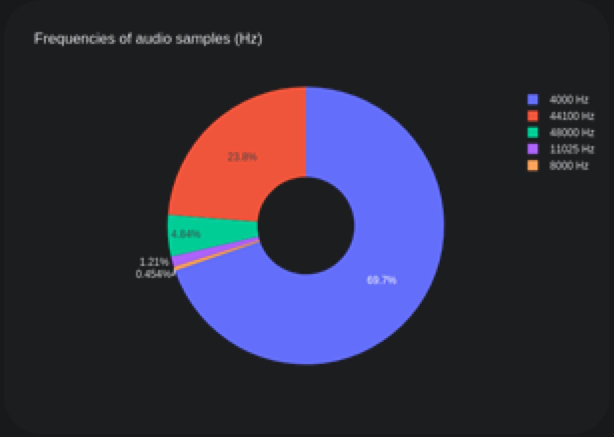
EXTRACTION INTERVAL

Each audio is divided into chunks of 1 second as it offers good performance while augmenting at the same time the number of samples



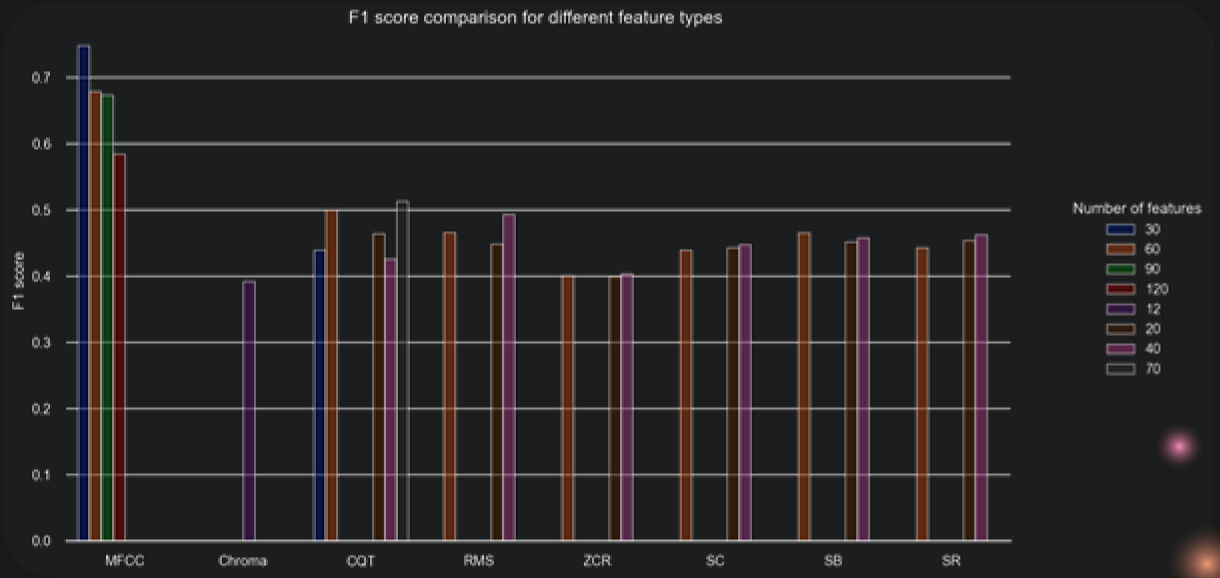
SAMPLING RATE

All audios were resampled at the most common frequency (4000 Hz)



FEATURES EXTRACTION

TYPE	N° FEATURES
MFCC	30
Chroma STFT	12
CQT	70
RMSE	40
ZCR	40
SB	60
SC	40
SR	40



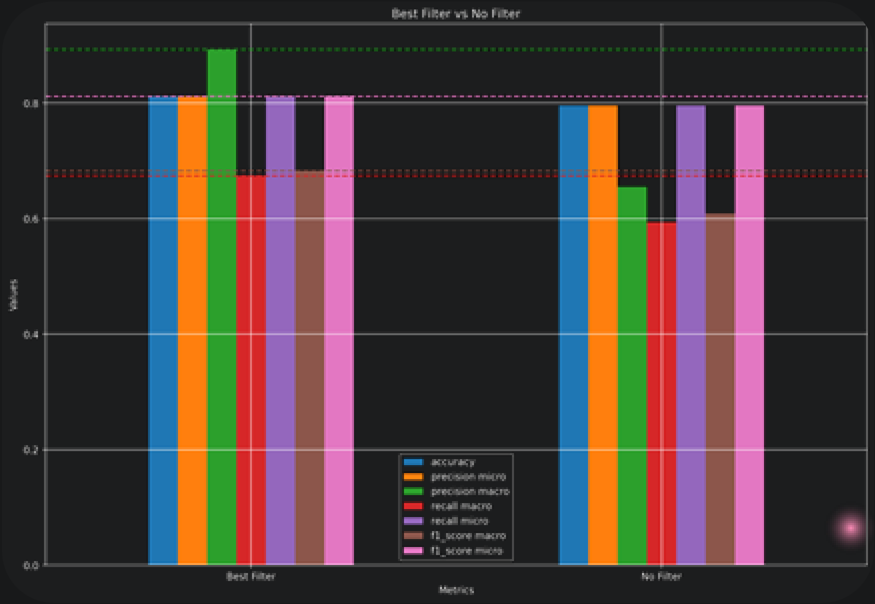
CORRELATION ANALYSIS

Correlation with Target

Removed features with a correlation lower than a threshold

Feature Correlation

Removed features with a correlation higher than a threshold with more than N other features



88 % features reduction

MODELS

PREVENTION MODEL

Used for prevention and focused on distinguishing a normal heartbeat, from an abnormal one or an artifact

SUPPORT MODEL

Specialized in identifying the specific disease and with explanaibility methods implemented

ARCHITECTURES

Name	Architecture (hidden layers)
Random Forest	-
XGBoost	-
CatBoost	-
LightGBM	-
MLP Basic	(128, 64, 32)
MLP Ultra	(512, 256, 128, 64, 32)
MLP Large	(256, 128, 64, 32)
MLP Small	(64, 32)
MLP Tiny	(32, 16)
MLP Reverse	(32, 64, 128, 256, 512, 256, 128, 64, 32)
MLP Bottleneck	(512, 64, 32)
MLP Rollercoaster	(512, 128, 256, 128, 256, 64, 32)
MLP Hourglass	(512, 256, 128, 64, 32, 64, 128, 256, 512)

Name	Architecture (hidden layers)
MLP Pyramid	(1024, 512, 256, 128, 128, 128, 64, 32)
MLP Wide	(1024, 1024)
MLP WideUltra	(1024, 1024, 128, 32)
MLP Sparse	(32, 16, 8)
MLP Dropout	(128, 64, 32)
MLP Ensemble1	MLP Basic, Large, Ultra
MLP Ensemble2	RandomForest, MLP Ultra
MLP Ensemble3	MLP Rollercoaster, Large
MLP Ensemble4	MLP Rollercoaster, Large, Ultra
MLP Ensemble5	RandomForest, MLP Ultra, Rollercoaster
MLP Ensemble6	MLP Rollercoaster, Large, Ultra, Wide
ALL Ensemble	All models majority vote
CB ALL Ensemble	All models CatBoost

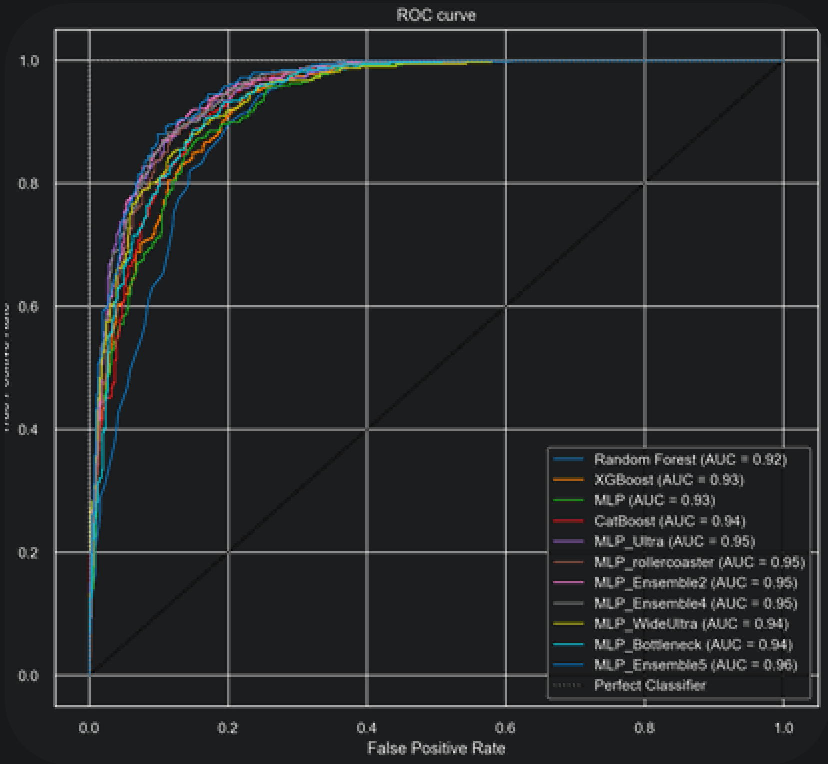
PREVENTION MODEL

Normal FPR Minimization Approach

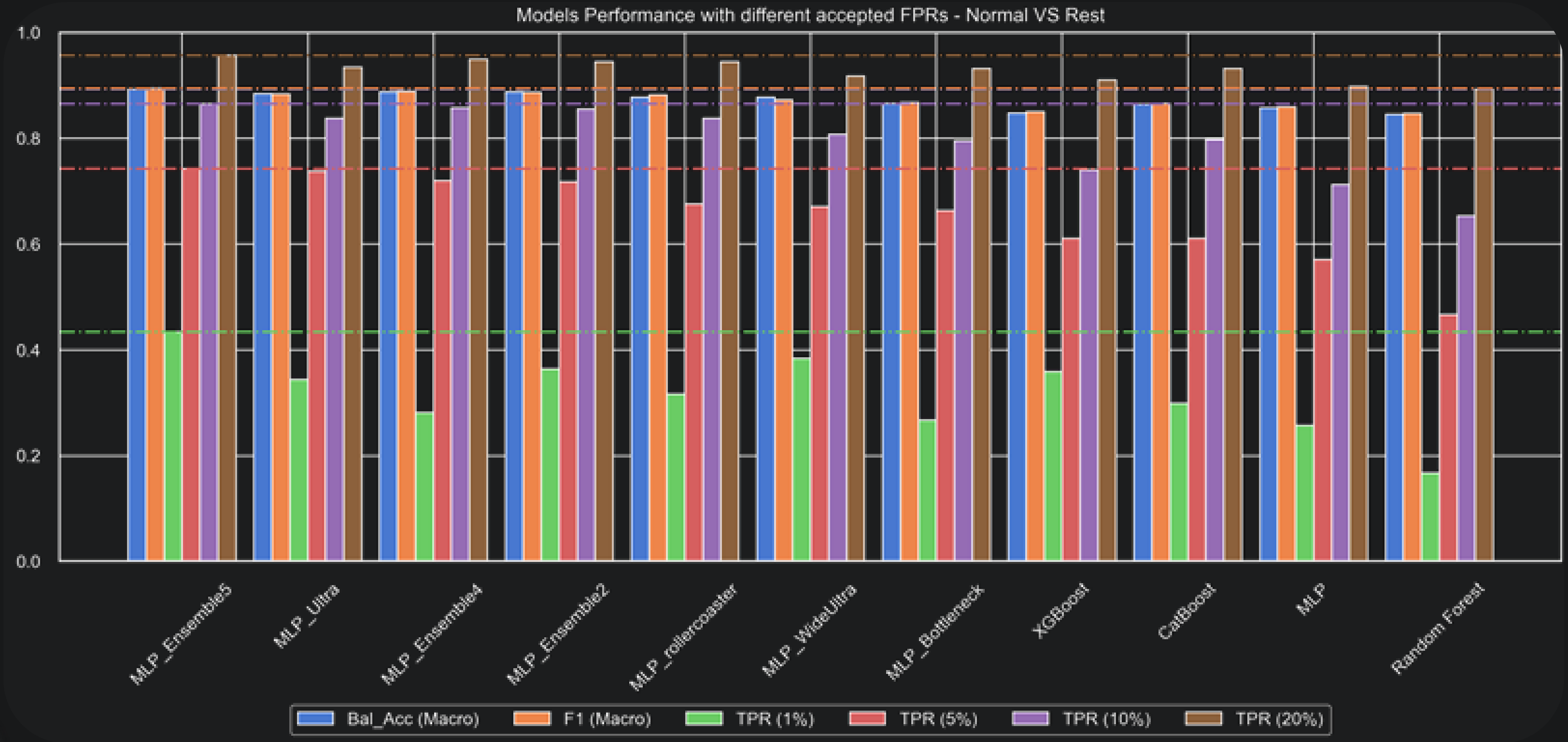
Artifacts or abnormal heart sounds classified as normal are a big threat for the patient safety

At the same time TPR do not have to be disregarded, otherwise the model is unuseful

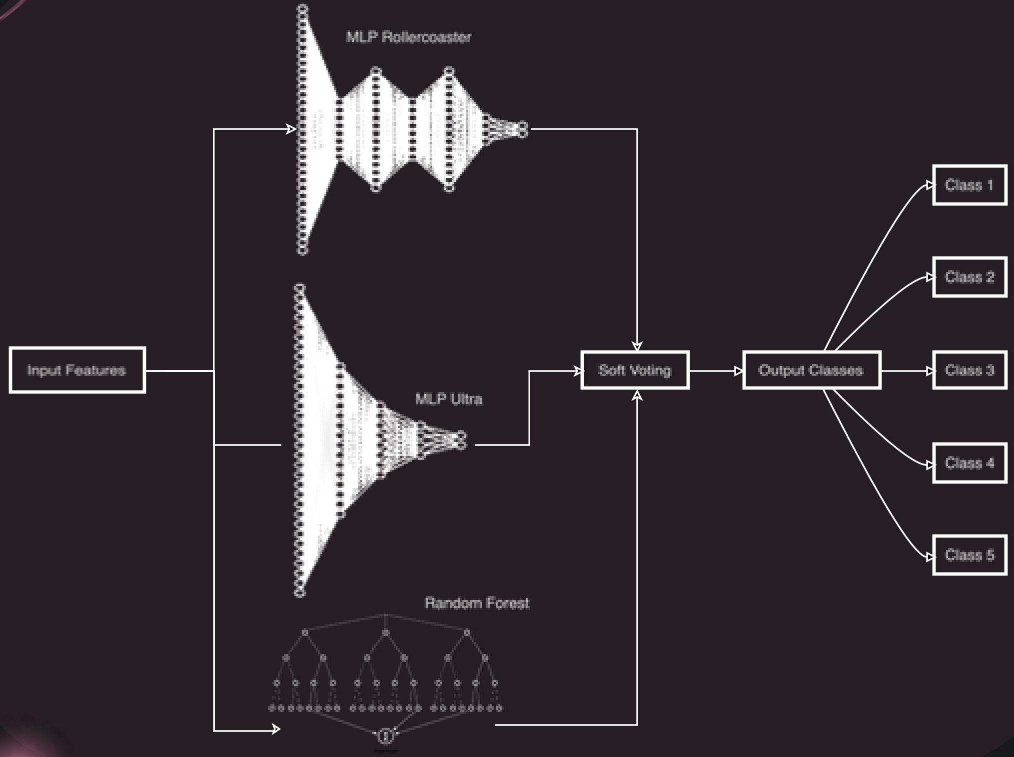
FPR vs TPR trade-off

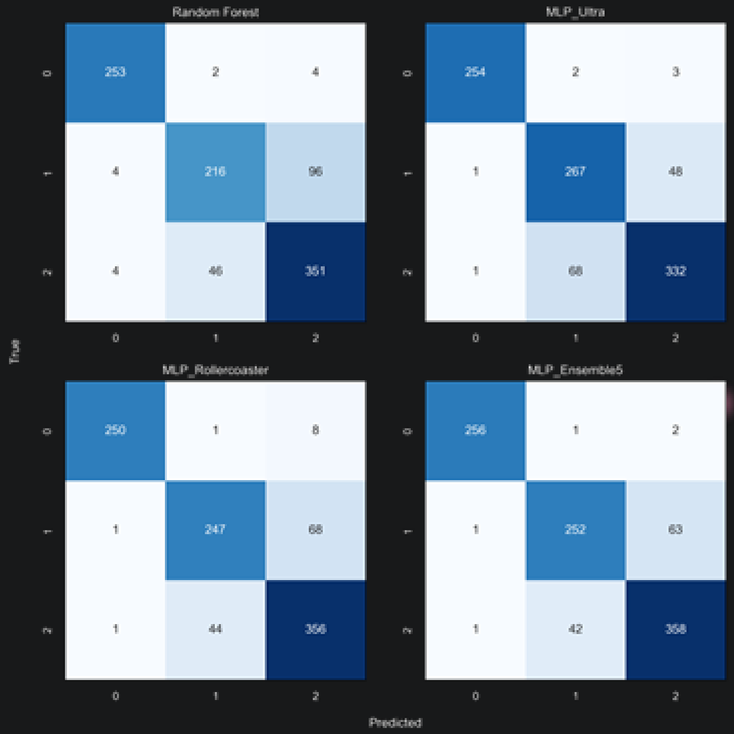


PREVENTION MODEL



PREVENTION MODEL





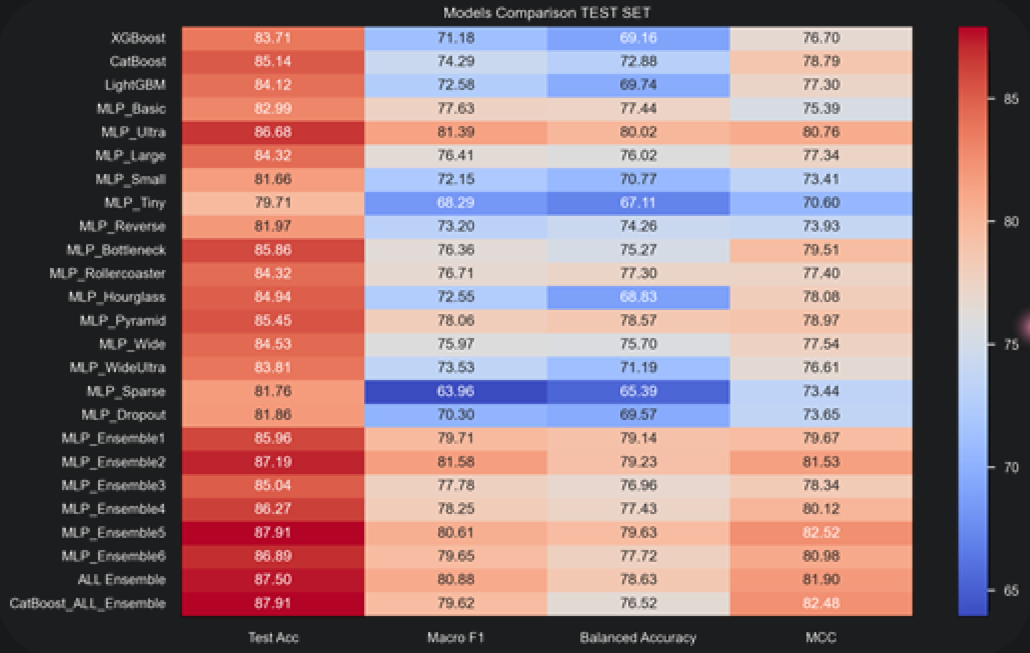
SUPPORT MODEL

Overall metrics Maximization Approach

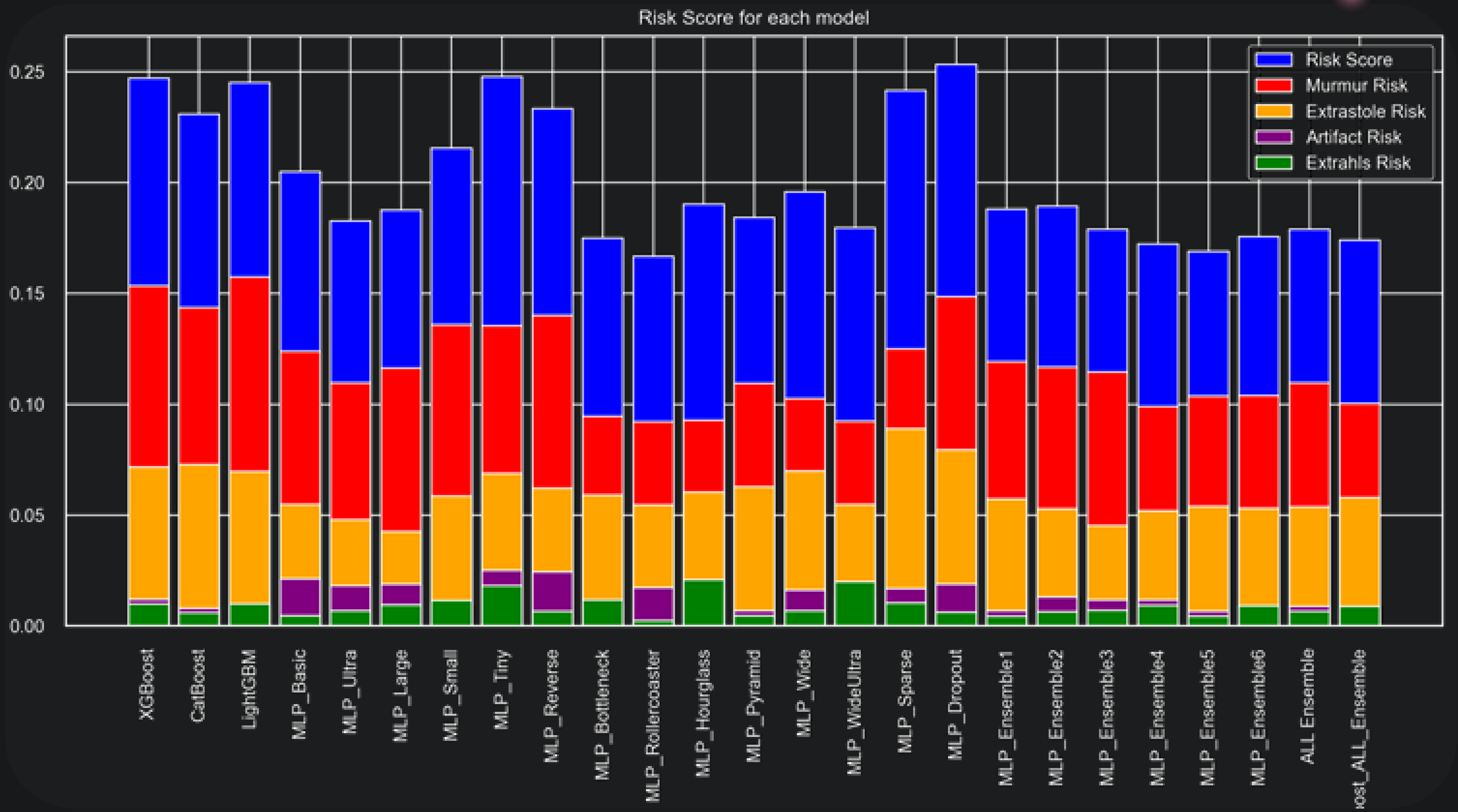
The model must be capable not only to differentiate whether a diseases is present or not, but also which disease is present

Metrics considering the class imbalance were employed

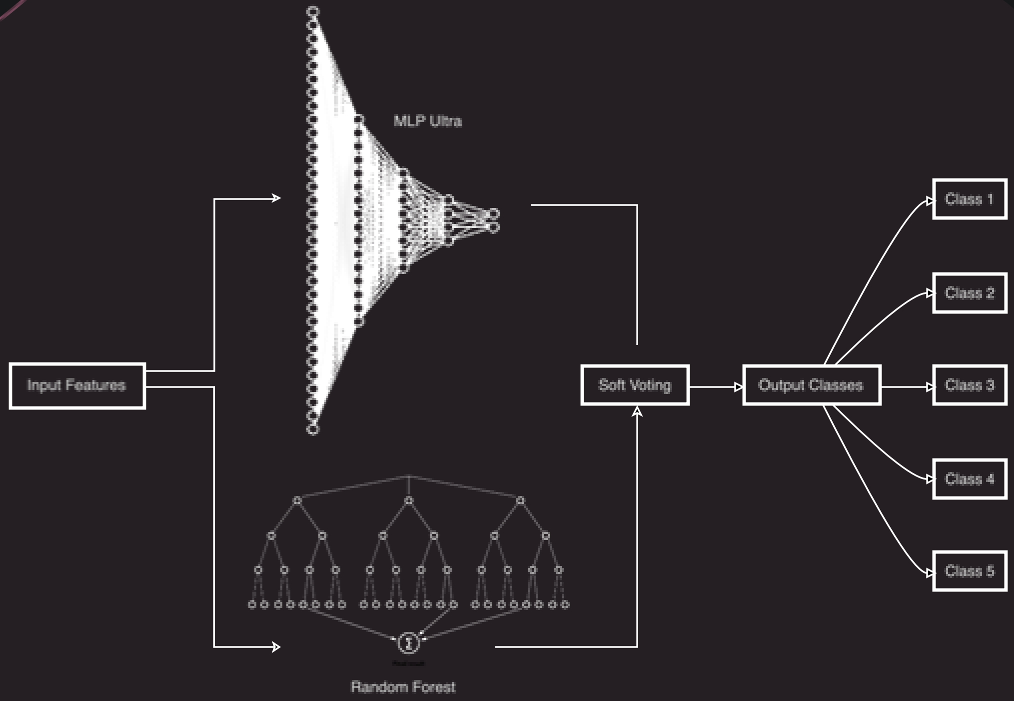
To help the physician understanding the reasons behind each decision, explainability tools were used

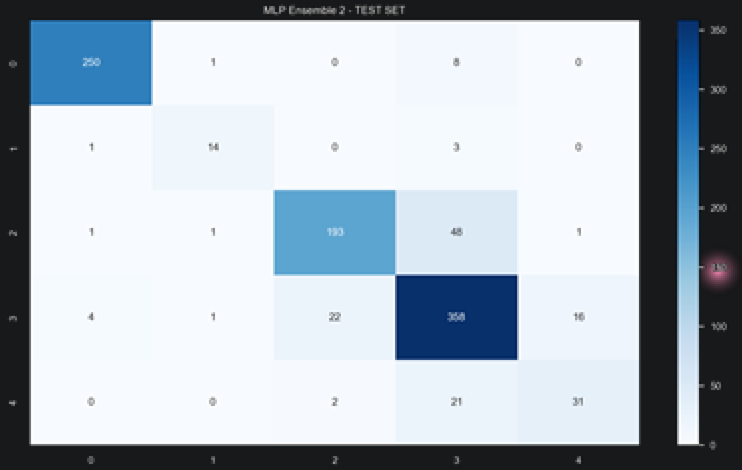


SUPPORT MODEL



BEST MODEL – MLP Ensemble 2



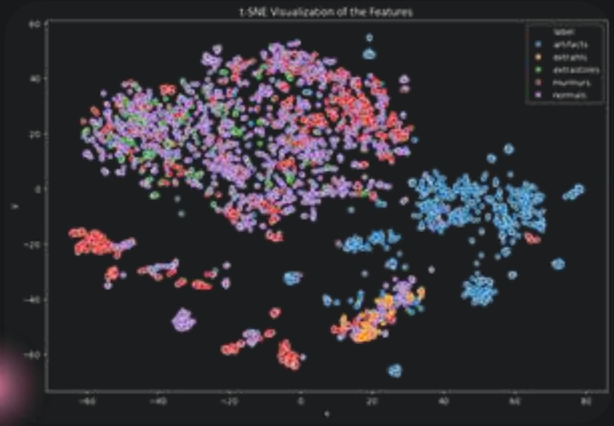


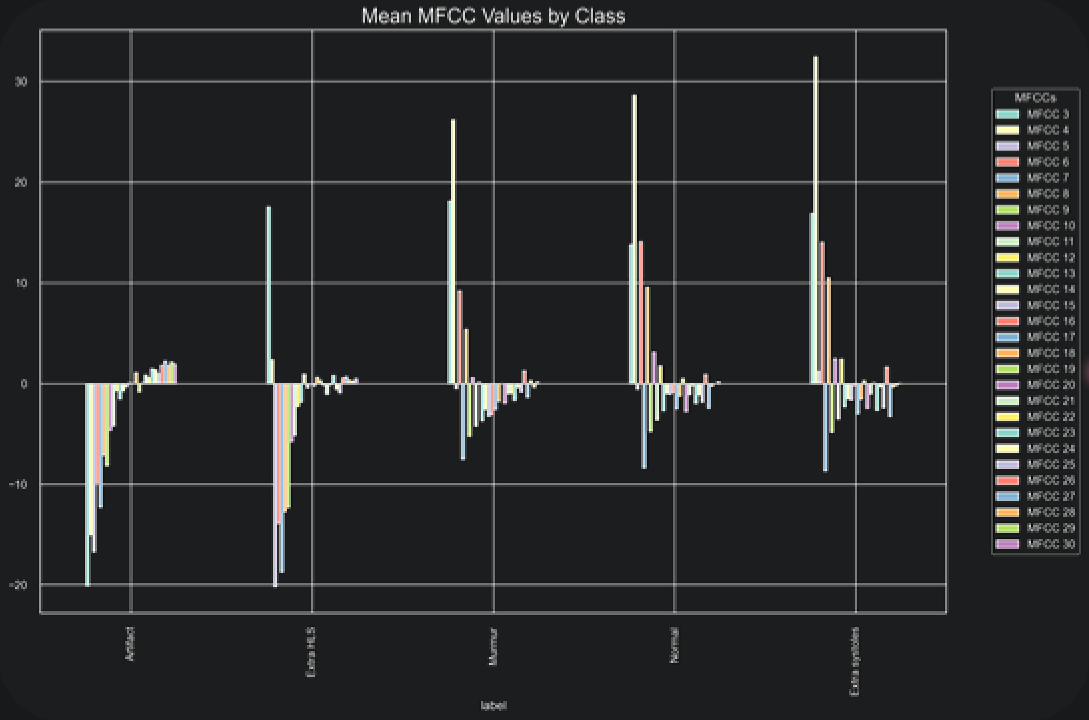
BEST MODEL – MLP Ensemble 2

Mean values across classes

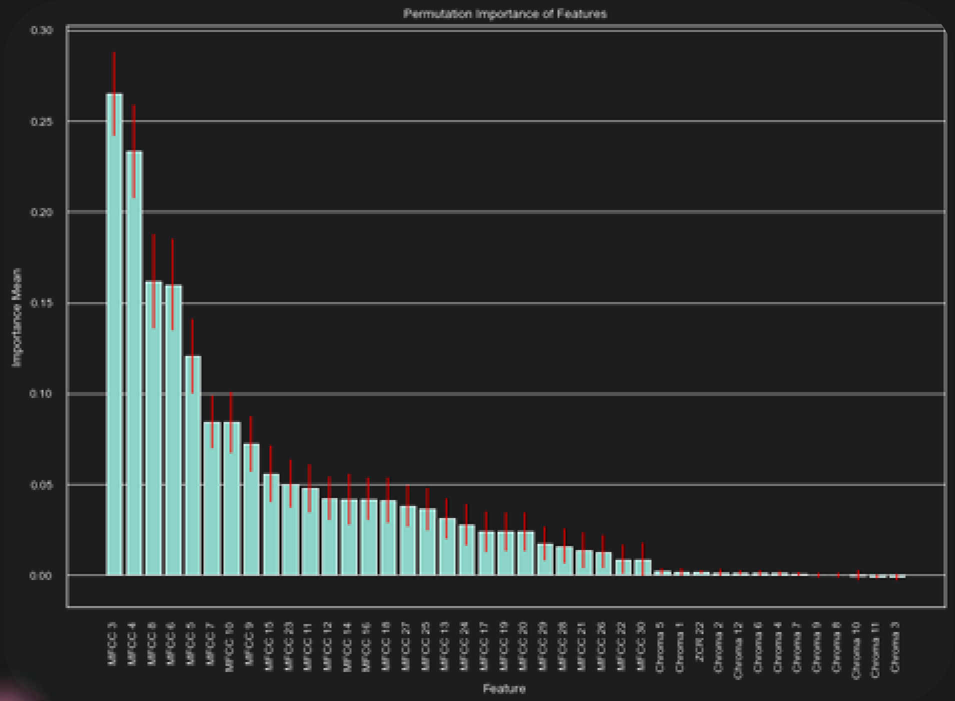
Murmur, Normal and Extra Systole have similar MFCCs values

New features need to be introduced, to help discriminate among those classes





EXPLAINABILITY – MLP Ensemble 2



Permutation Importance

Compute different times the importance of the features. Bars are the mean values and red lines are the standard deviations

Main Takeaways

Most important features : MFCCs
Less important : Chroma STFTs

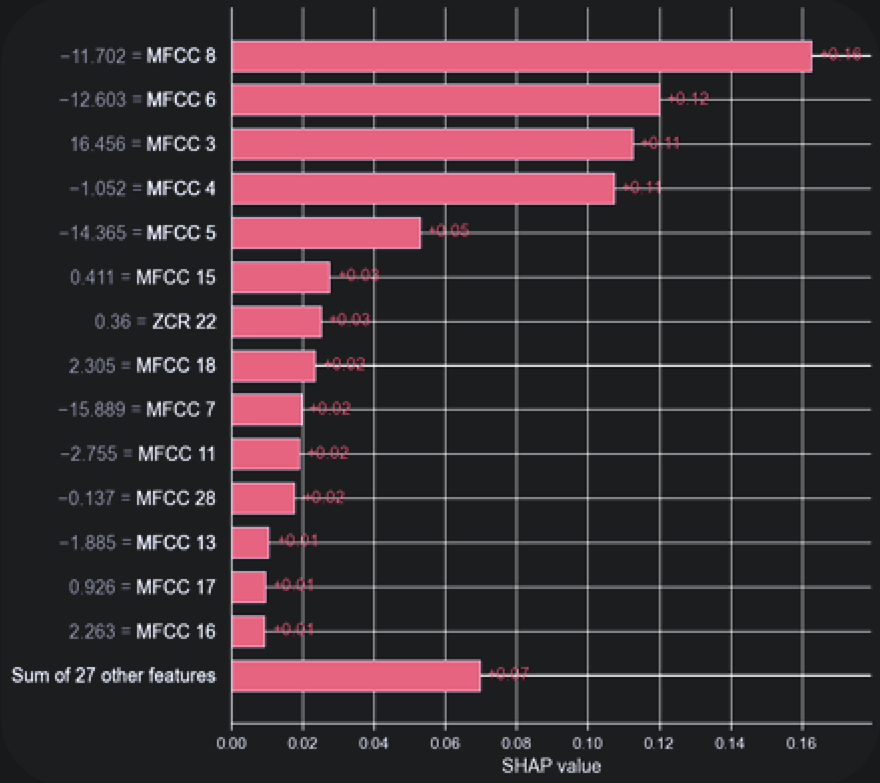
MFCC3 is the most important

EXPLAINABILITY – MLP Ensemble 2

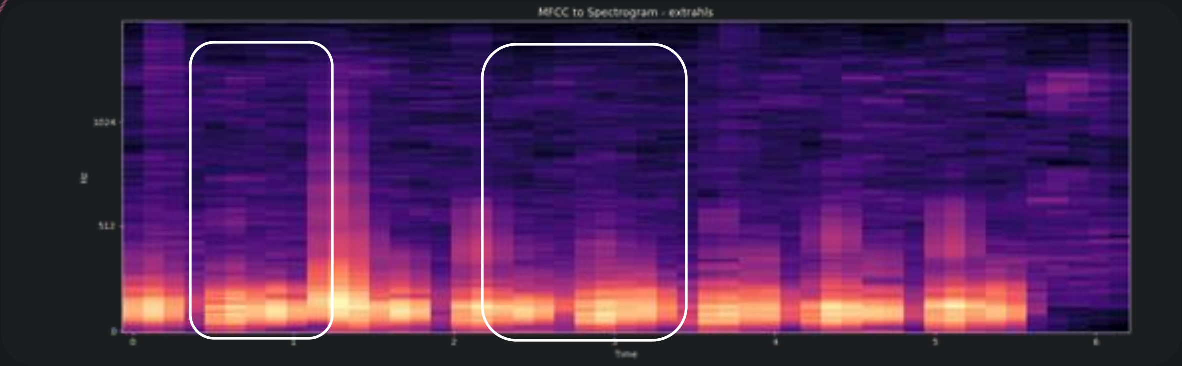
Feature importance single prediction

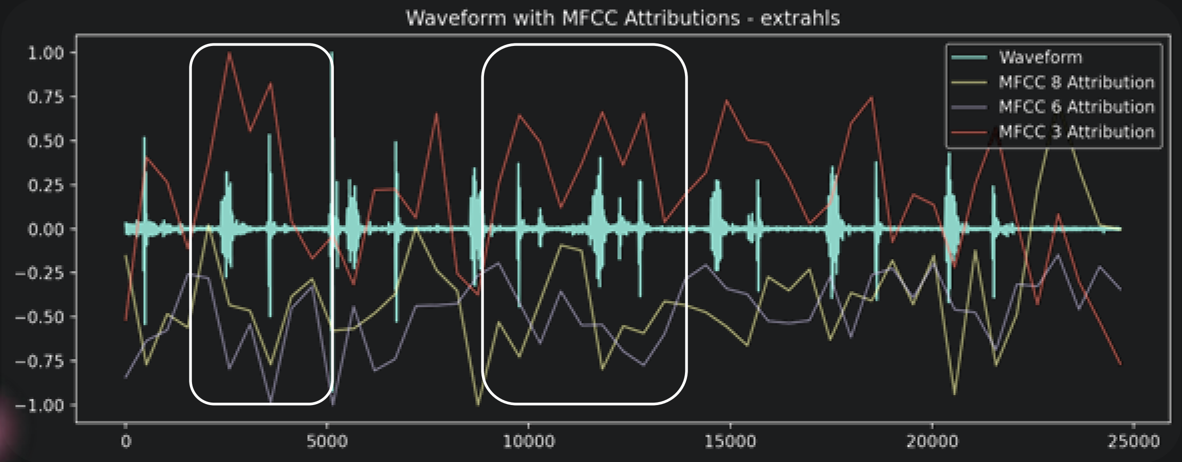
Most impactful features in the predicition of a extrahls sample, according to SHAP

The higher the value of the bar, the higher the probability of having the extrahls class for the given sample



EXPLAINABILITY – MLP Ensemble 2





Frequencies distribution across time

Most meaningful zones for the model

OTHER EXPERIMENTS

CNN Based on Spectral Features

Spectral features converted into images

VGG16 with Batch Normalization as a features extractor

Promising results that may be optimal after fine tuning

CNN Based on waveforms

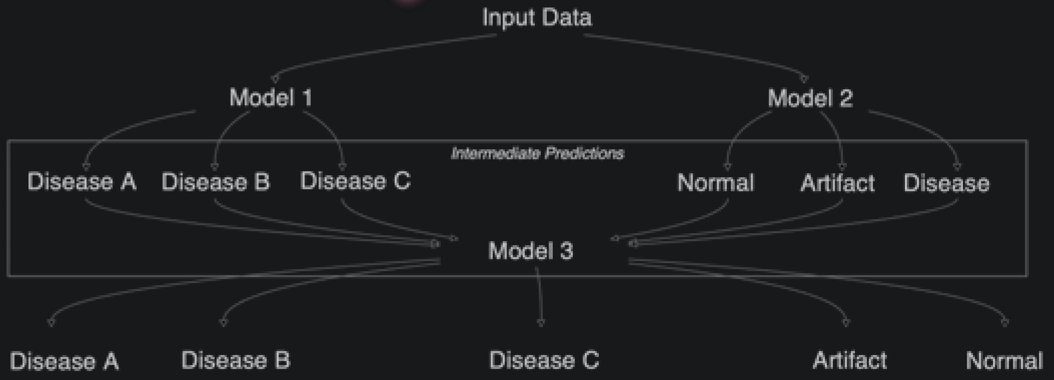
VGG16 with Batch Normalization, using raw waveform images

Features taken from 3rd, 4th and 5th convolutional layer after pooling

Tried different classifiers (RF, MLP, SVM) with no relevant results

OTHER EXPERIMENTS

Tiered Ensemble Model



Data Augmentation

We tried to augment the training data of the under represented classes by adding noise and changing speed and pitch. No relevant improvement showed.

FUTURE WORKS

1. Highly imbalanced classes and difficult to balance (increase the dataset), we had to make many trade-off choices.
2. Inability to create a validation set to find the hyperparameters, test metrics are biased.
3. Evaluate the techniques used with clinicians and their consistency with medical aspects, especially for the explainability phase
4. Experiment with other features. MFCC resulted not discriminant for 3 classes

THANKS!

AndreaAlberti07 Andrea Alberti DavideLigari

REFERENCE

1. Zhang, W.; Han, J.; Deng, S. Heart Sound Classification Based on Scaled Spectrogram and Tensor Decomposition. Expert Syst. Appl. 2017, 84, 220–231

2. Shi-Wen Deng, Ji-Qing Han, Towards heart sound classification without segmentation via autocorrelation feature and diffusion maps, Future Generation Computer Systems, 2016

3. Raza, A.; Mehmood, A.; Ullah, S.; Ahmad, M.; Choi, G.S.; On, B.-W. Heartbeat Sound Signal Classification Using Deep Learning. Sensors 2019

4. Alafif T., Boulares M., Barnawi A., Alafif T., Althobaiti H., Alferaidi A. Normal and Abnormal Heart Rates Recognition Using Transfer Learning; 2020

5. Noman F., Ting C.-M., Salleh S.-H., Ombao H. Short-segment heart sound classification Using an ensemble of deep convolutional neural networks; 2019

6. Chen L., Ren J., Hao Y., Hu X. The Diagnosis for the Extrasystole Heart Sound Signals Based on the Deep Learning. J. Med. Imaging Health Inform. 2018