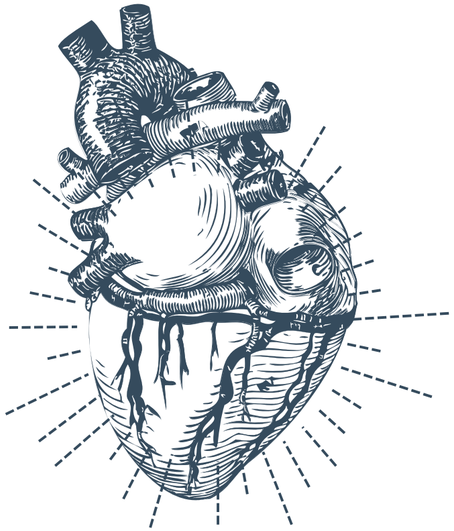
1 2 9 0

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE D COIMBRA

Classifying Heart Sounds



ADVANCED DATA ANALYSIS TECHNIQUES | **MACHINE LEARNING PROJECT**

Nicole Duarte 2018282646 | Sílvia Santos 2018282660

Project Brief

1 **Project Description**

2 **Objectives**

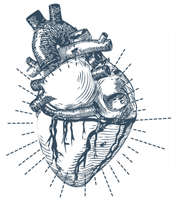
3 **Procedure**

4 **Results**

5 **Conclusion**

6 **Future Work**

Project Description | Objectives



According to the World Health Organisation, cardiovascular diseases (CVDs) are the number one cause of death globally.

CHALLENGE 1
Heart Sound Segmentation

Produce a method that can locate S1(lub) and S2(dub) sounds within audio data, segmenting the Normal audio files in both datasets.

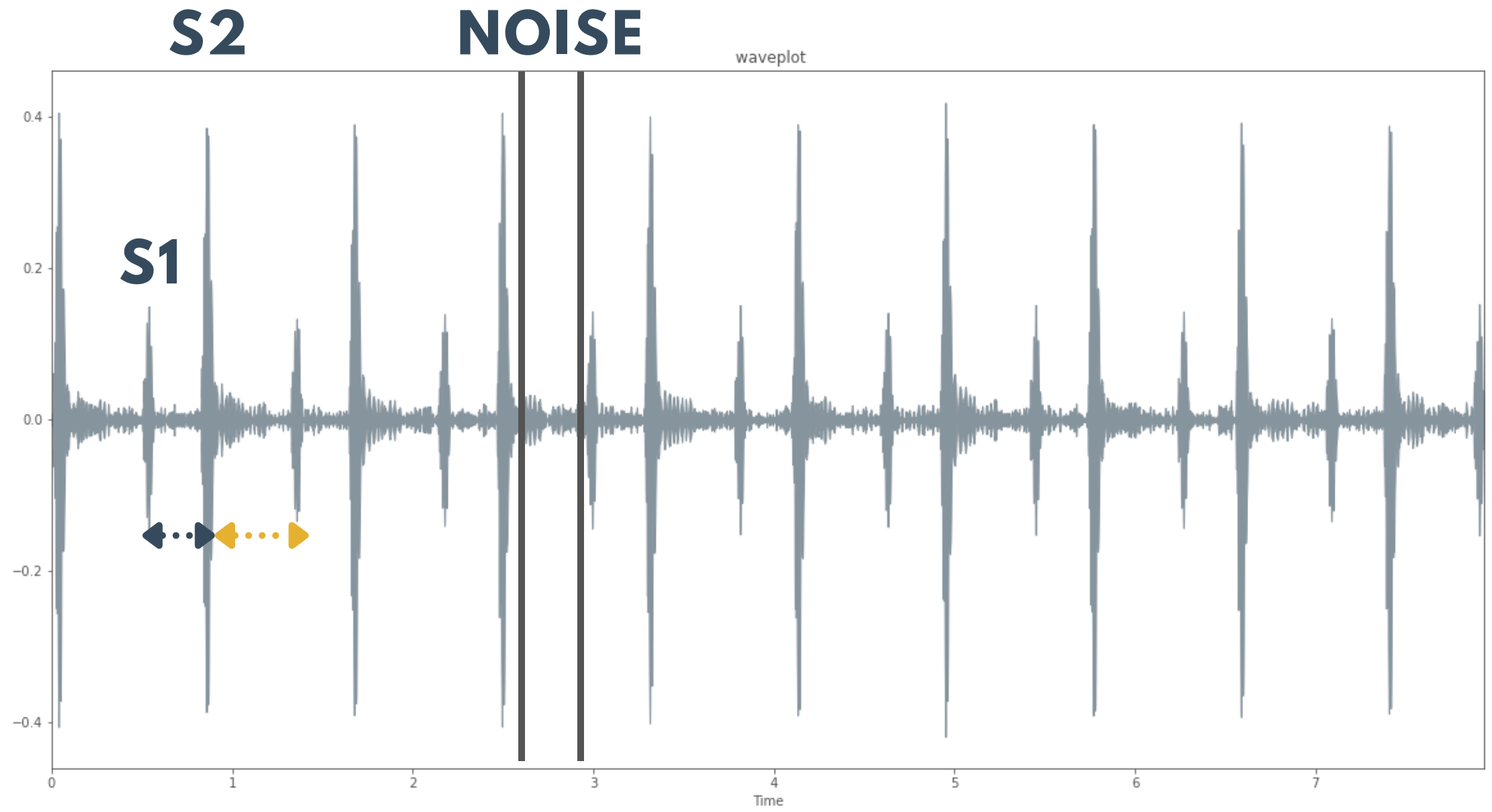
CHALLENGE 2
Heart Sound Classification

Produce a method that can classify real heart audio into one of four categories: Normal, Murmur, Extra Heart Sound and Artifact.

Procedure

CHALLENGE 1

S1 | S2 | Noise



Procedure

CHALLENGE 1

Resources

Locations of S1 and S2 sounds in Atraining_normal.csv.

Python

Librosa

Python package for music and audio analysis.

Keras

Open source neural network library written in Python.

Audacity

Audio Editing Software.

Procedure

CHALLENGE 1

Steps

1

Audios' Split
S1, S2, Noise

2

Features' Extraction

3

Model
Training and Testing

4

Predict

Procedure

CHALLENGE 1

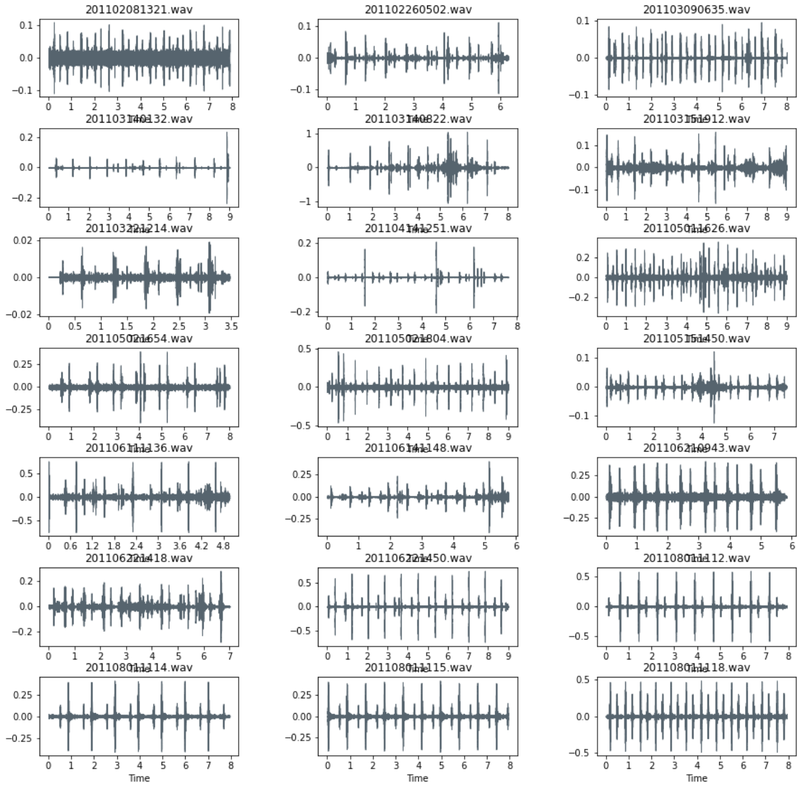
1

Audios' Split

S1, S2, Noise

The positions (in samples) given in the csv file did not all follow the same identification rule.

Do a new cvs with the positions where S1 and S2 start, using Audacity.



Procedure

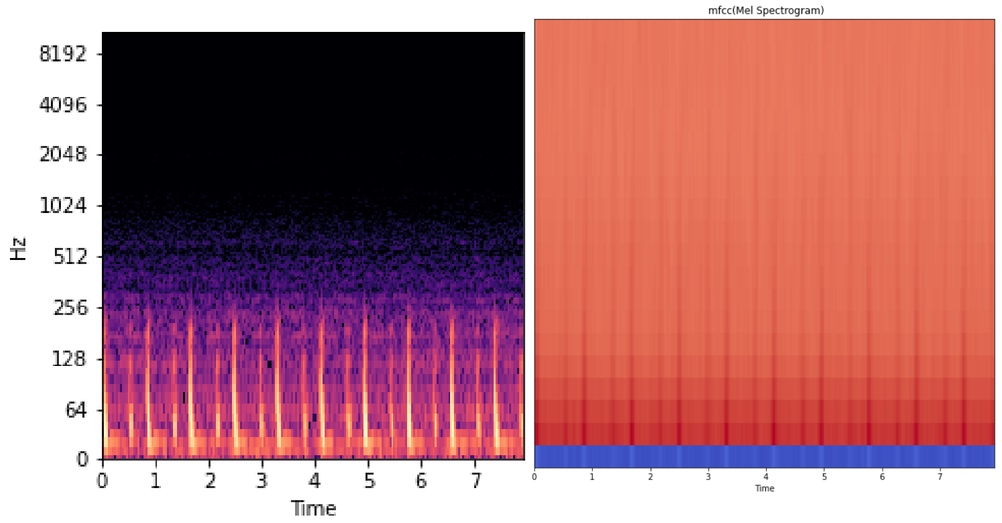
CHALLENGE 1

2 **Features' Extraction**

Define the best method for extracting the features from the audios.

The chosen method: Extraction through **Mel Spectrogram**

1. Compute a Mel-Scaled Spectrogram
2. Use a pre-computed log-power Mel Spectrogram
3. Extract Features: librosa.feature.mfccs() | mfccs = 20
4. Average



Procedure

CHALLENGE 1

3 — **Model**

Overfitting

1. Insert **Dropout** layers
2. Decrease the complexity
   -> Decrease the **Density**

```
model.add(Dropout(0.5))
model.add(Dense(124))
```

Procedure

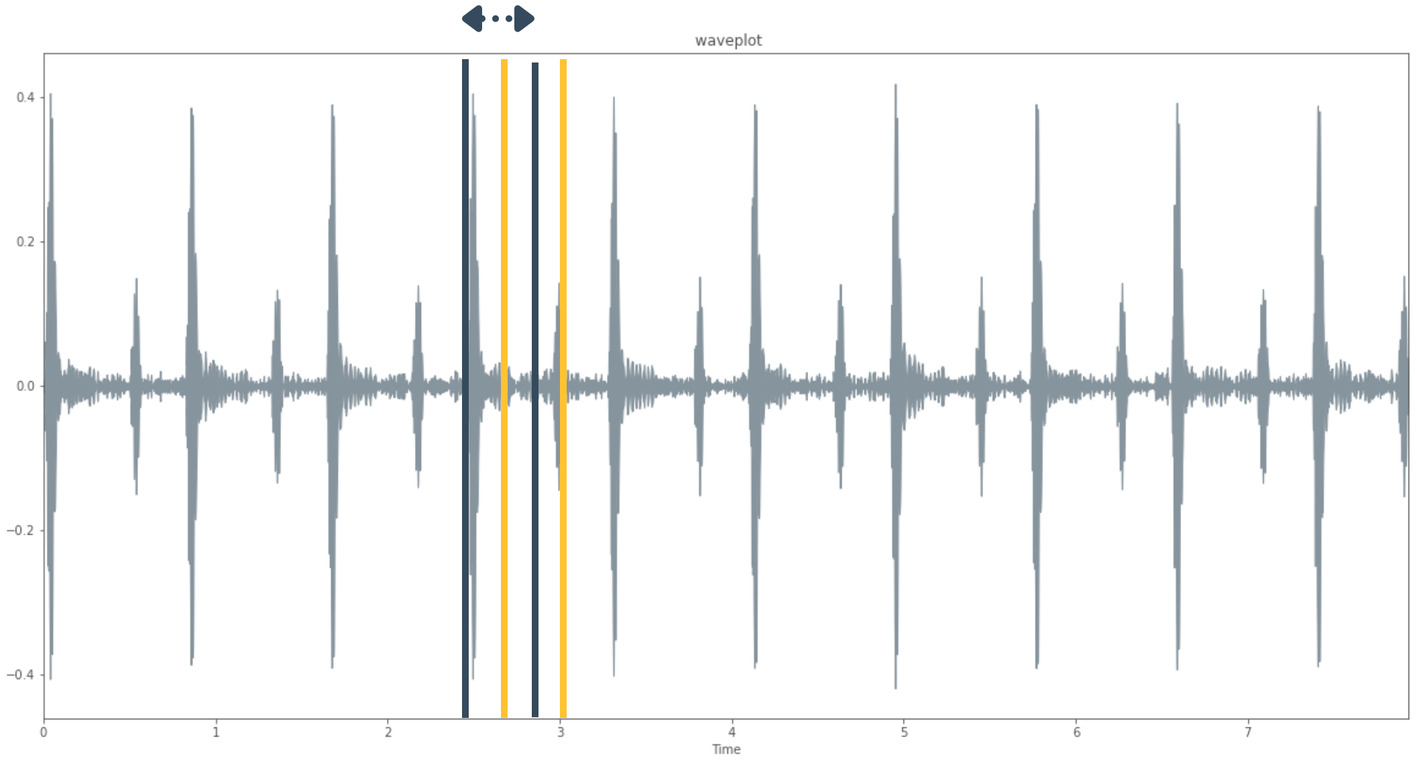
CHALLENGE 1

4

Predict

Define the size of the Slide Window.
MFCCs entry order

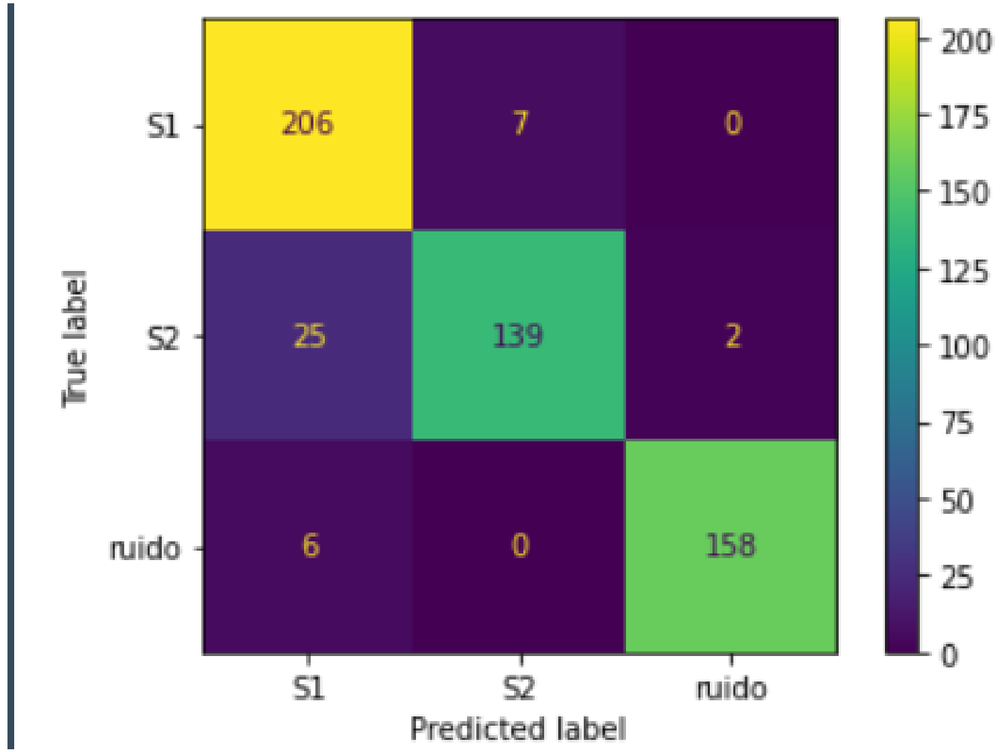
Attempts

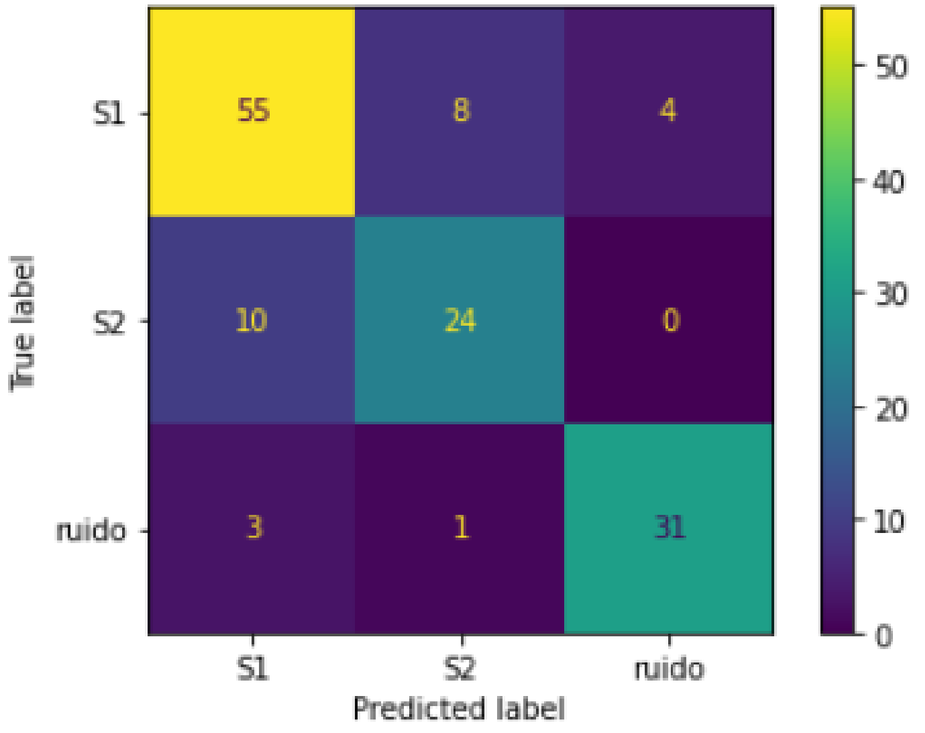


Results

CHALLENGE 1

| Model - Confusion matrix





Results

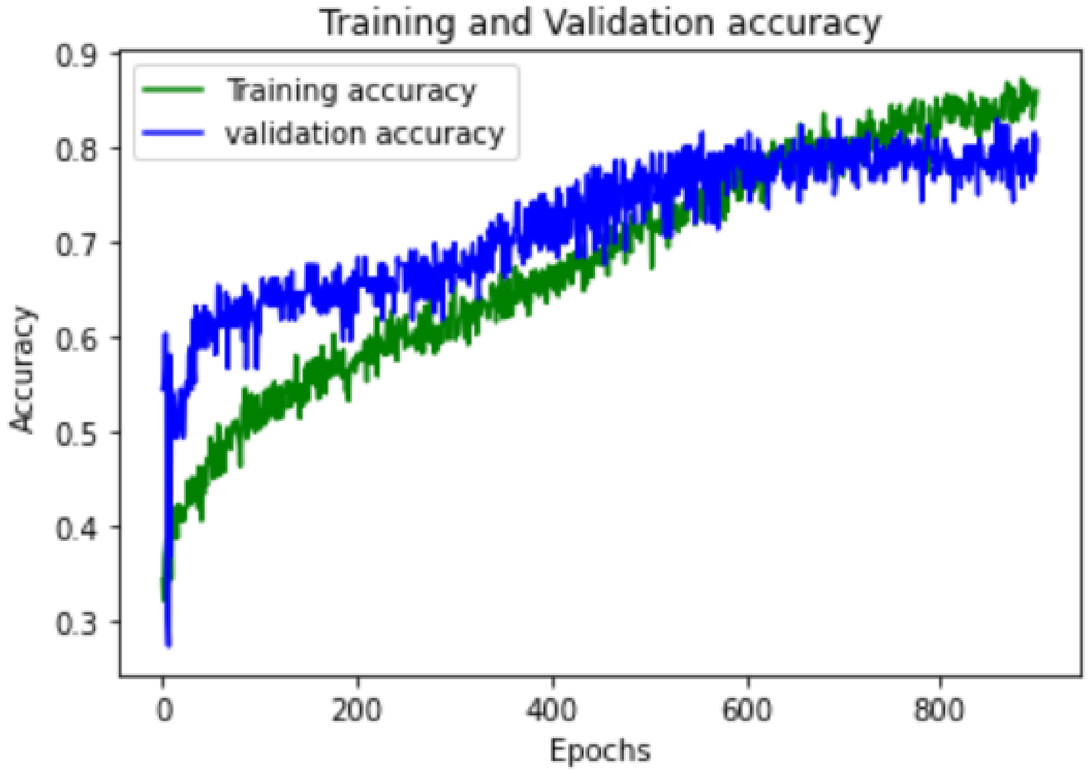
CHALLENGE 1

| Model - Training and Testing

Training - 80% of the sounds
Test - 20% of the sounds
epochs = 900

Training Accuracy: 86.00%

Testing Accuracy: 80.88%



Results

CHALLENGE 1

Predict

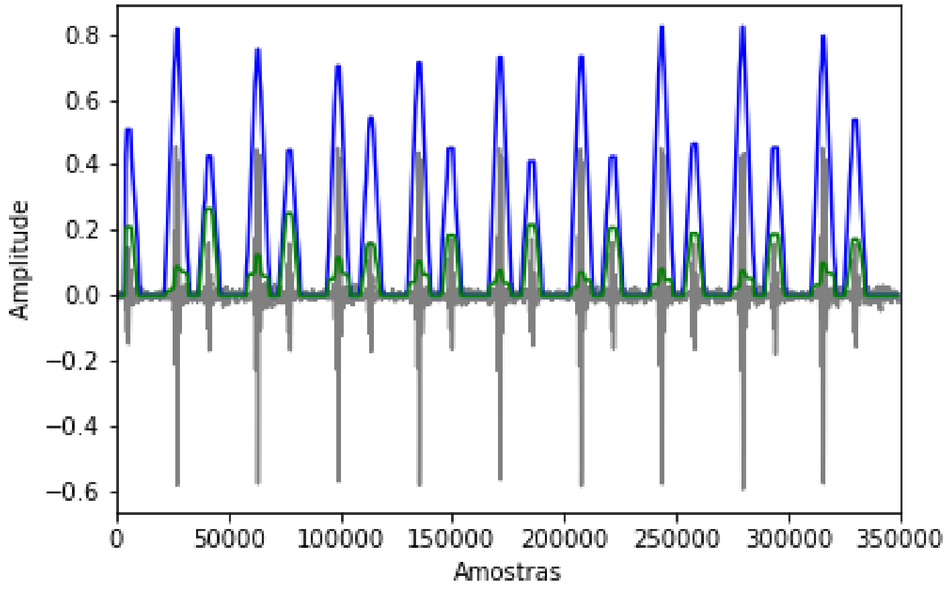
Window_size - 2700 samples

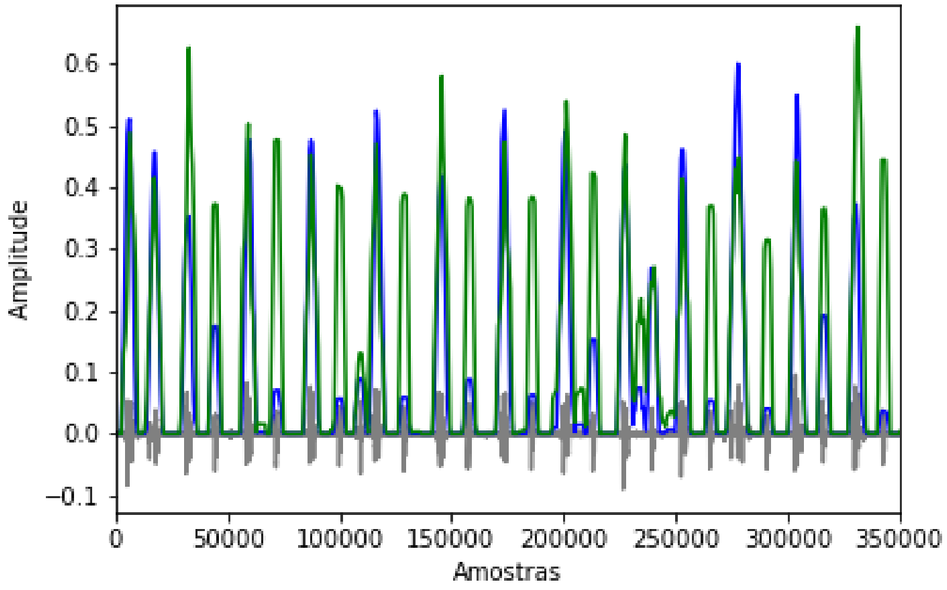
Steps - Window_size/5

S1 Probability - Blue

S2 Probability - Green

Sound - Gray

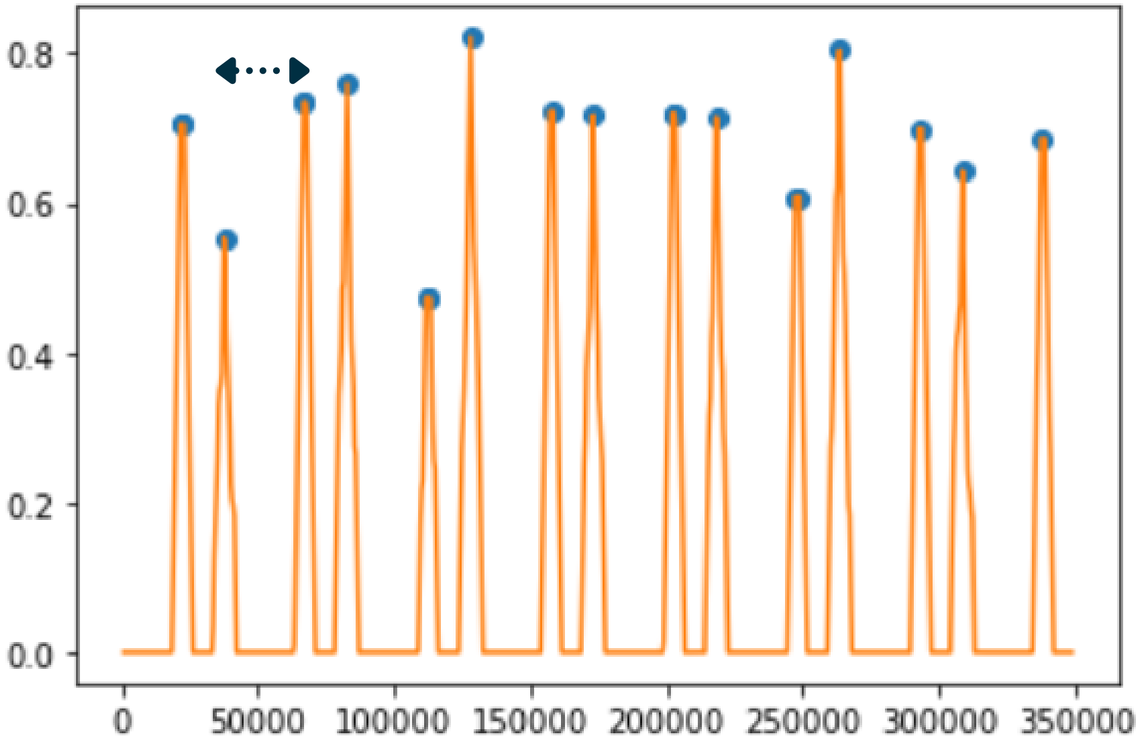




Results

CHALLENGE 1

Heart Rate



Conclusion

CHALLENGE 1

Heart Sound Segmentation

Produce a method that can locate S1(lub) and S2(dub) sounds within audio data, segmenting the Normal audio files in both datasets.

Results

Model

Predict

Heart Rate

Future Work

CHALLENGE 2

Heart Sound Classification

Produce a method that can classify real heart audio into one of four categories: Normal, Murmur, Extra Heart Sound and Artifact.

Procedure

1. Training only with Mermuriums
2. Identify if there are multiple S1 and S2
3. Check for sounds in addition to S1, S2 and noise - Artifacts.
4. Check when some element is missing

References

Web

https://librosa.org/doc/latest/index.html

http://www.peterjbentley.com/heartchallenge/

https://towardsdatascience.com/getting-to-know-the-mel-spectrogram-31bca3e2d9d0

Acknowledgments

Professor Filipe Veloso

Thank you for your attention!