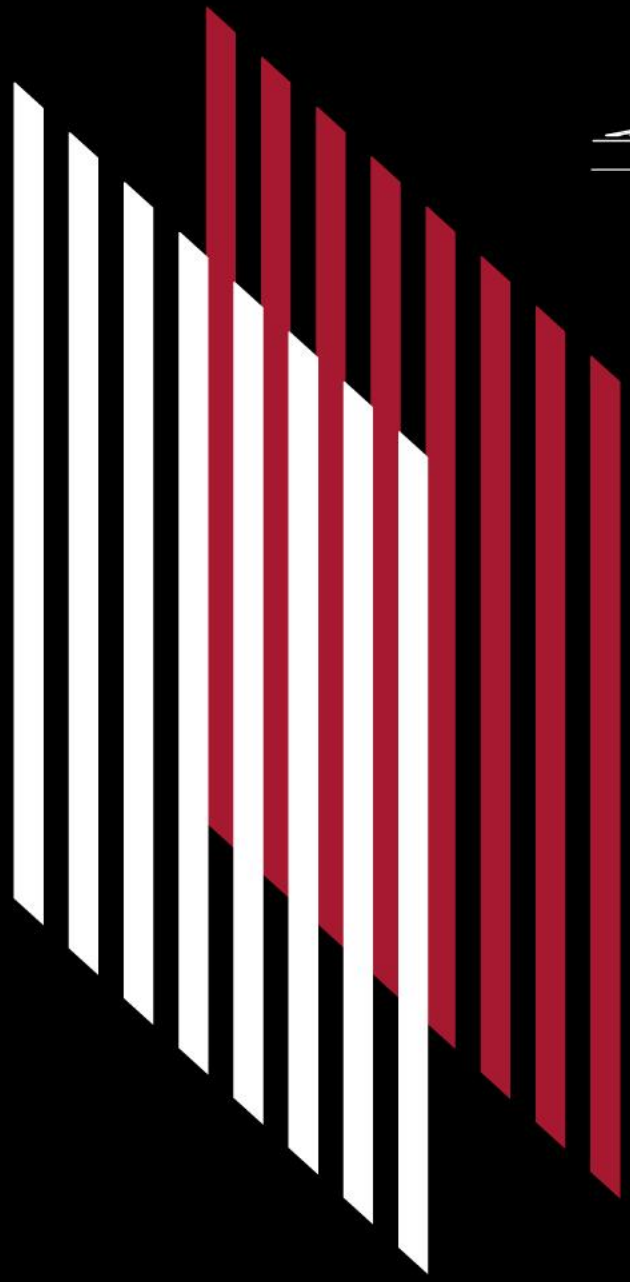
Murmur Detection Challenge

Andrés Felipe Romero Gómez

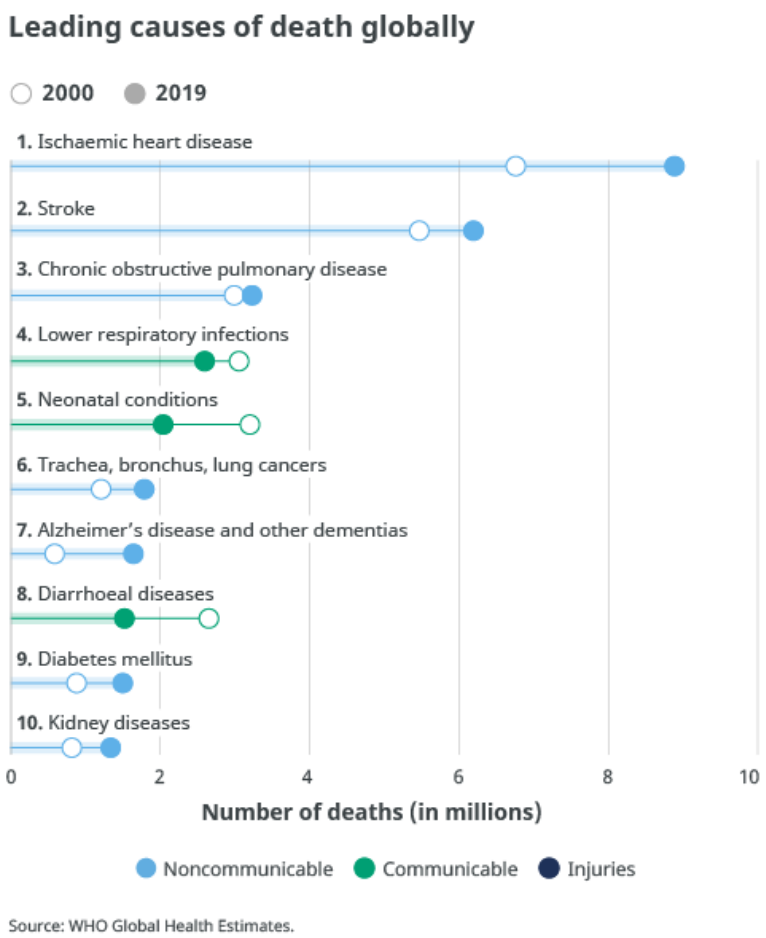


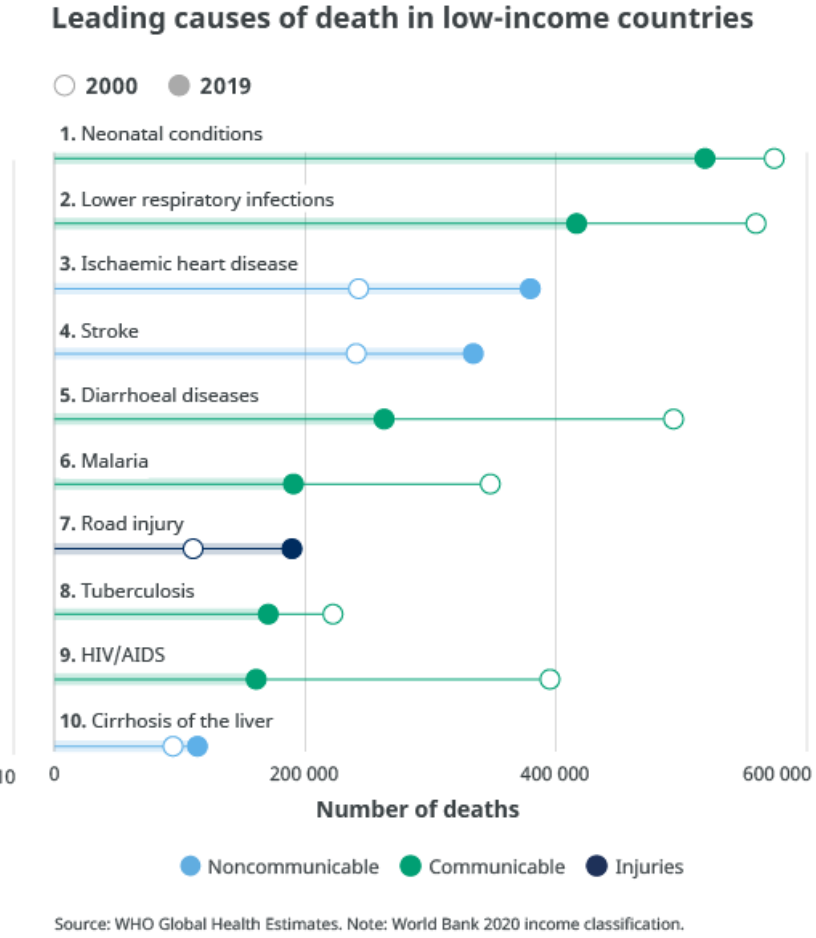
ESCUELA COLOMBIANA DE INGENIERÍA JULIO GARAVITO

UNIVERSIDAD

Problem

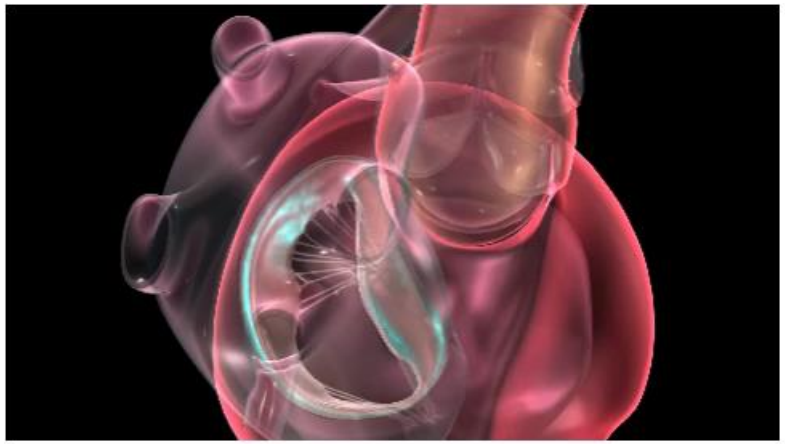
- Congenital heart diseases affect about 1% of newborns, representing an important morbidity and mortality factor.
- Acquired heart disease represents a major public health problem in developing regions.
- Difficulties in diagnosing and treating due to the lack of infrastructure and cardiology specialists.



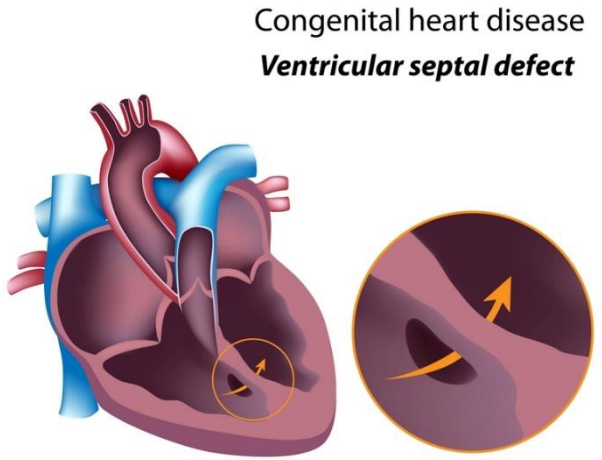


Hearth Murmur

- Cardiac auscultation and the analysis of the phonocardiogram (PCG) can unveil fundamental clinical information regarding heart malfunctioning caused by congenital and acquired heart disease in pediatric populations.
- Murmurs are abnormal waves generated by turbulent blood flow in cardiac and vascular structures.
- Can be related with specific diseases.
- Innocent or abnormal heart murmurs
- Diseases that can reduce the quality of life and can be potentially fatal.



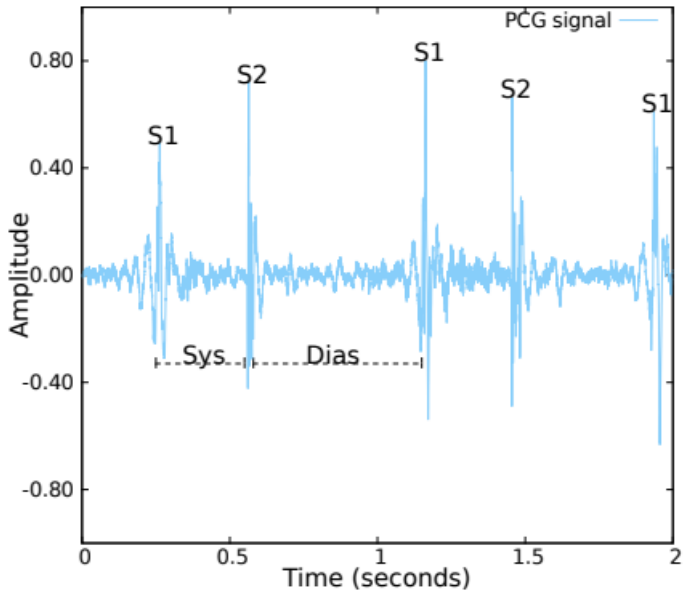
Taken from: https://my.clevelandclinic.org/health/diseases/17083-heart-murmur



Taken from: https://universityhealthnews.com/daily/heart-health/what-is-a-heart-murmur/

Heart Murmur Detection from Phonocardiogram Recordings: The George B. Moody PhysioNet Challenge 2022

Matthew A. Reyna, Andoni Elola, Jorge Oliveira, Francesco Renna, Annie Gu, Nadi Sadr, Erick A. Perez Alday, Yashar Kiarashinejad, Sandra Mattos, Miguel T. Coimbra, Reza Sameni, Ali Bahrami Rad, Gari D. Clifford



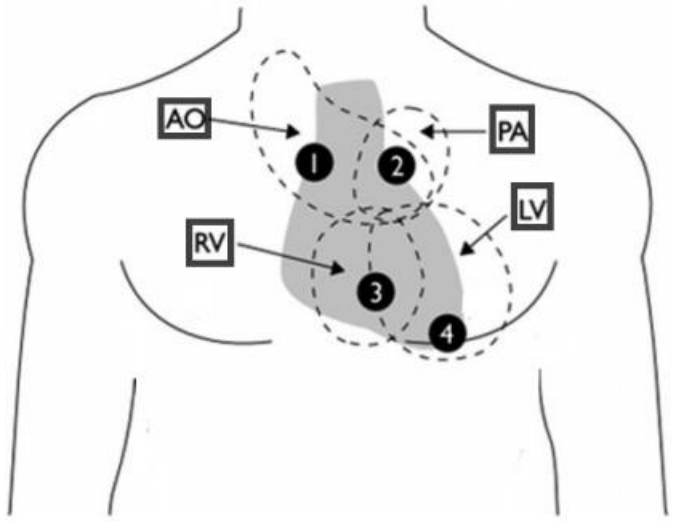


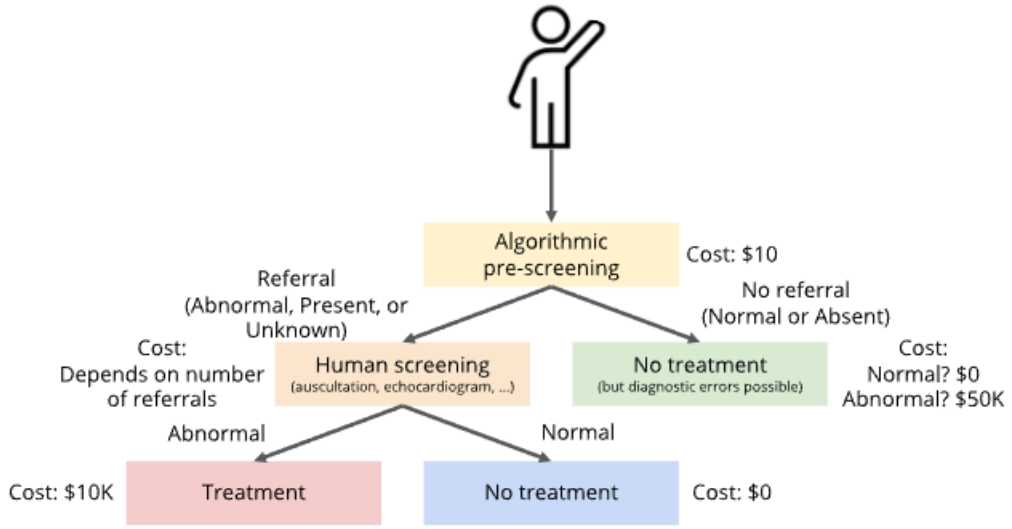
Fig. 1. Cardiac auscultation spots (image adapted from [20]); AO = aortic area; LV = left ventricle; PA = pulmonary area; RV = right ventricle; 1 = right second intercostal space; 2 = left second intercostal space; 3 = midleft sternal border (tricuspid); 4 = fifth intercostal space, midclavicular line.

- Data recorded from 1568 patients.
- Demographic and signal annotated data.
- 60% for training data (available to the public).

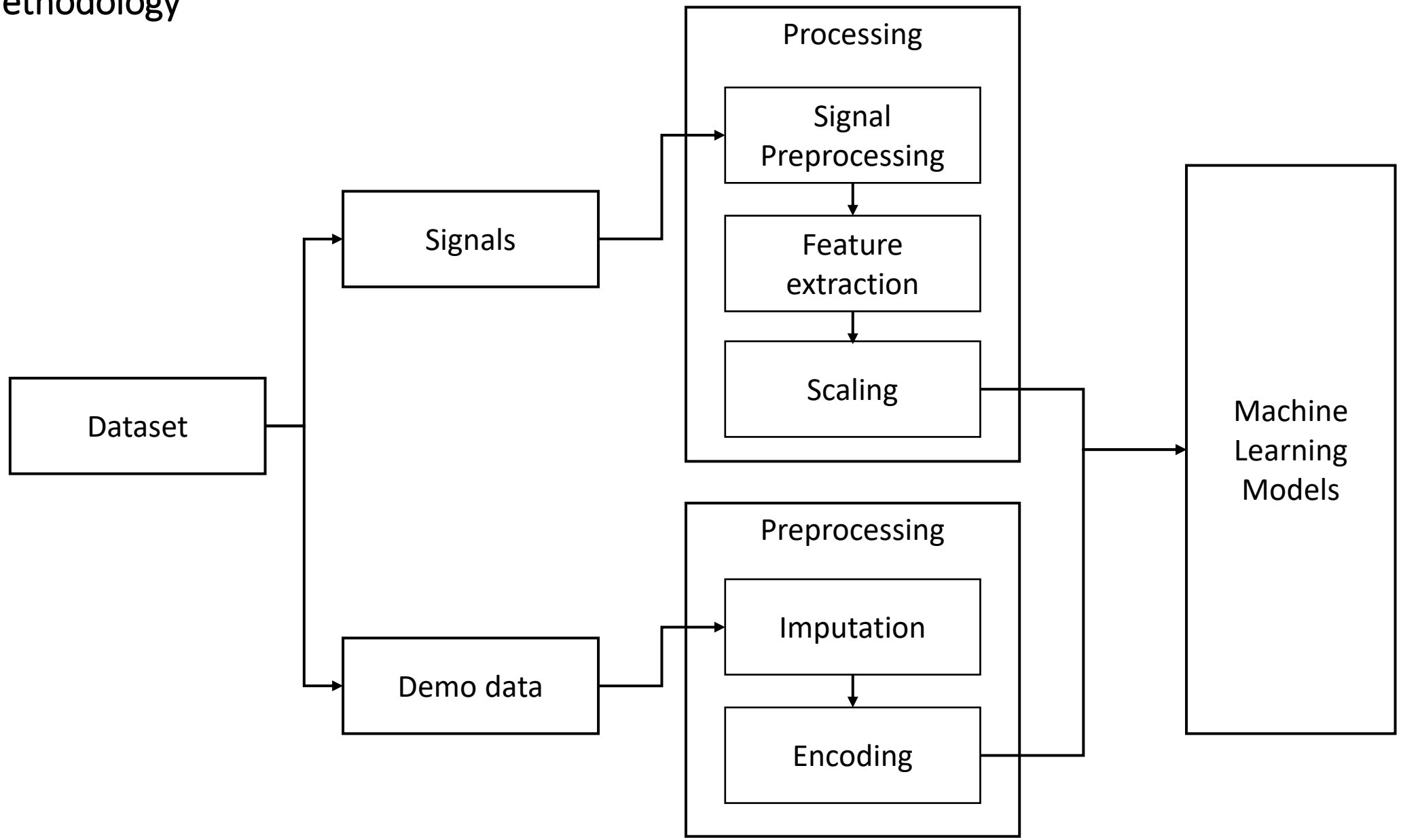
Objective

- Challenge: The goal of the Challenge is to identify the presence, absence, or unclear cases of murmurs and the normal vs. abnormal clinical outcomes from heart sound recordings collected from multiple auscultation locations on the body using a digital stethoscope.

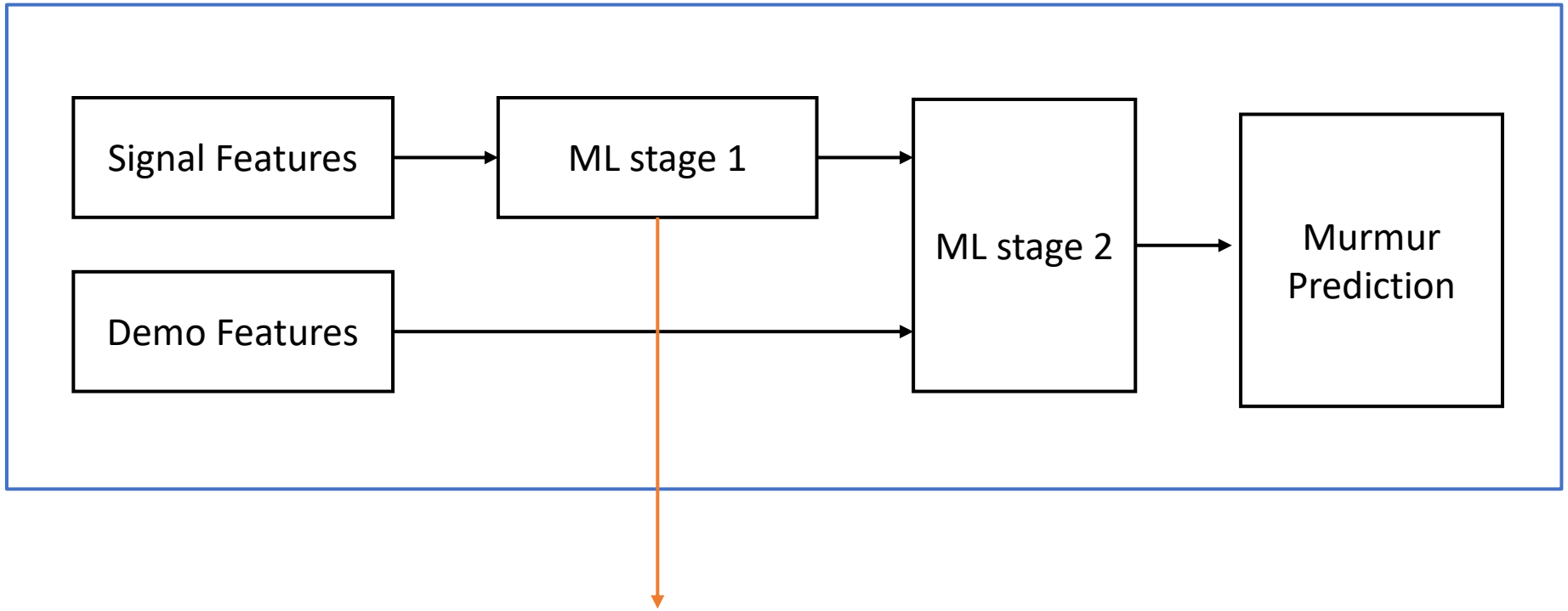
- Classify between patients with and without the presence of murmurs using demographic data and PCGs.



Methodology



Methodology for Machine Learning (ML) Models

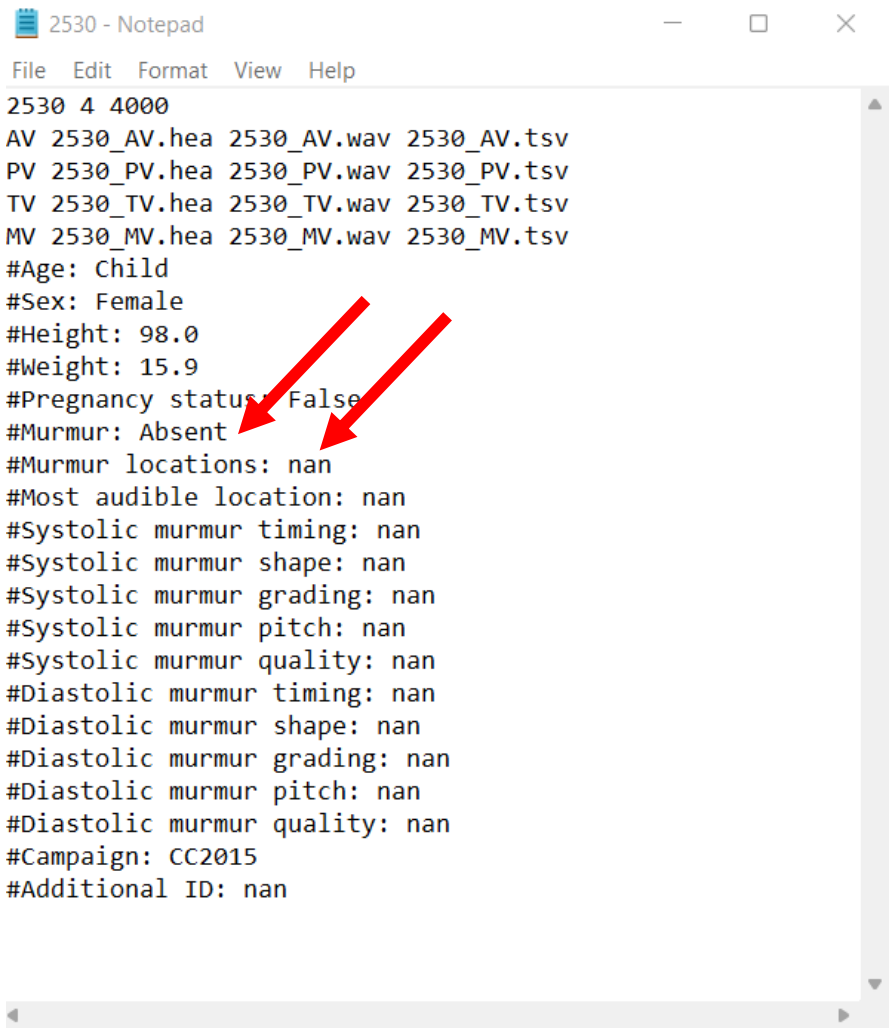


- Hyperparameter Exploration
- Resampling techniques
- Feature Selection

Dataset General Description

- 3163 recordings from 942 patients

- A wave recording file (*.wav*)
- A header file (*.hea*)
- A segmentation data file (*.tsv*)
- A subject description text file (*.txt*)



Demographic data (*.txt* files)

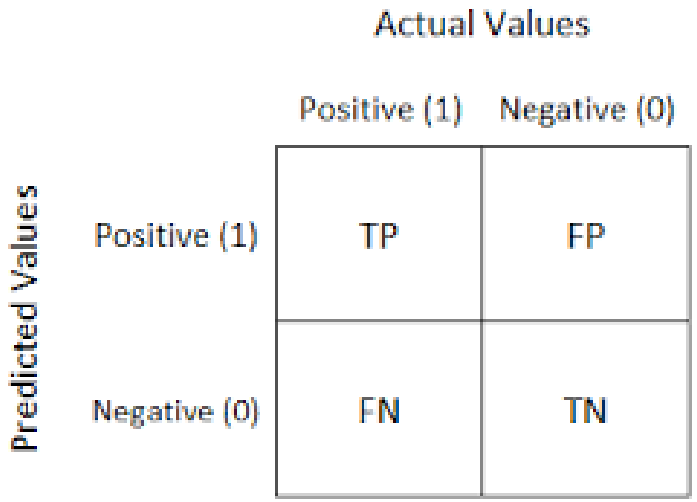
ESCUELA COLOMBIANA DE INGENIERÍA JULIO GARAVITO
UNIVERSIDAD

Pandas Profiling
Profile Report

Performance Metrics

Value	Count	Frequency (%)
Absent	695	73.8%
Present	179	19.0%
Unknown	68	7.2%

#Murmur locations: nan



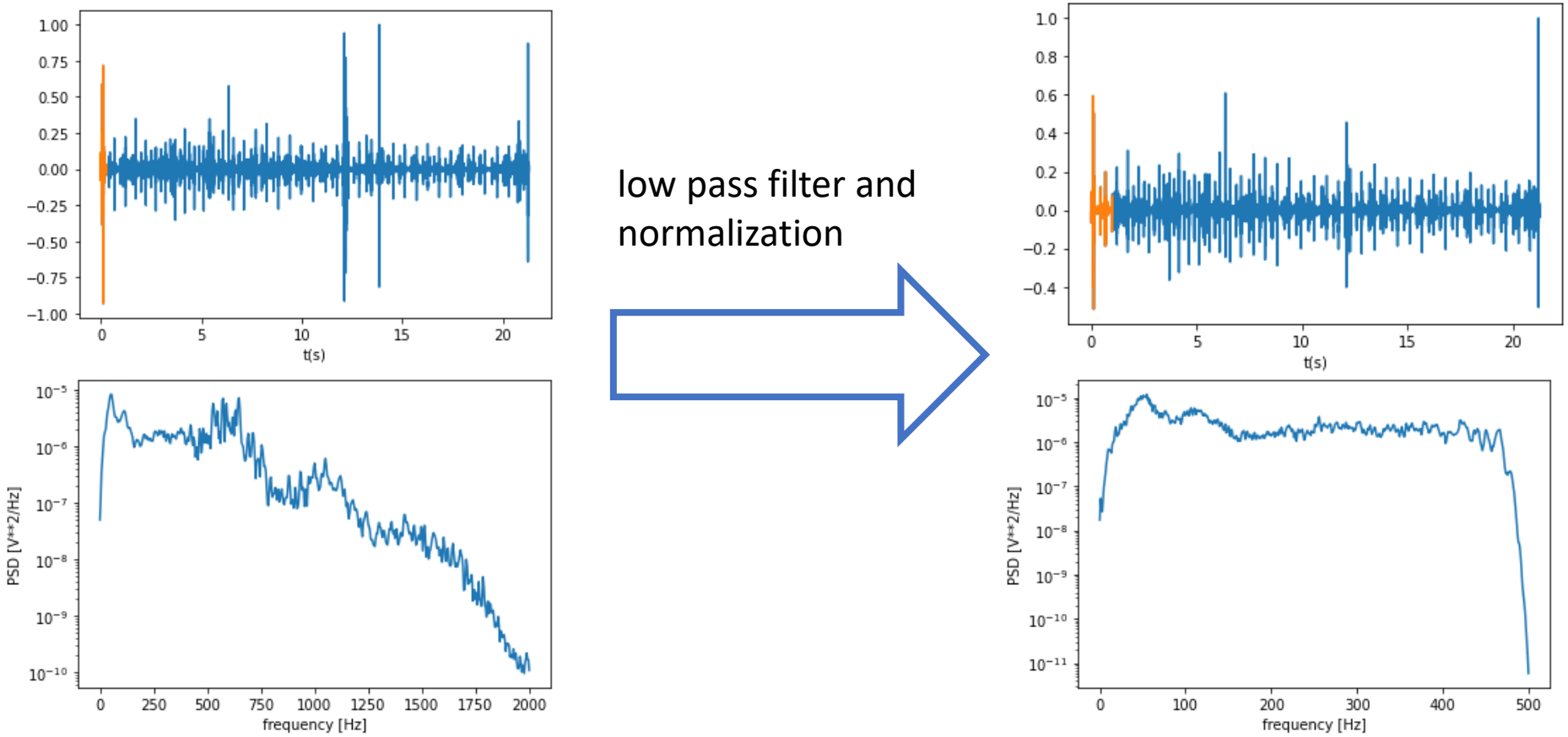
$$\texttt{balanced-accuracy} = \frac{1}{2}\left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP}\right)$$

$$Precision = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN}$$

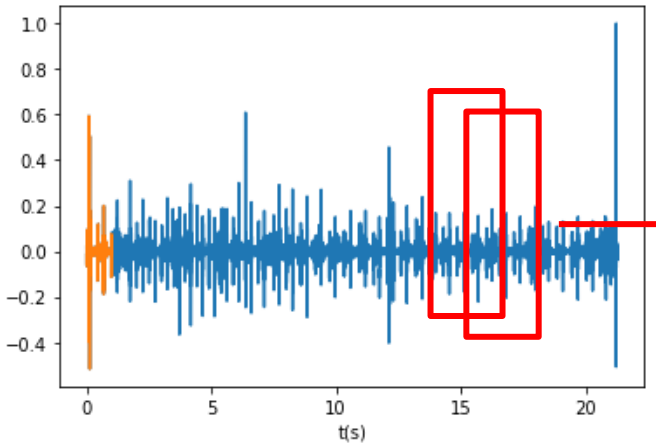
$$F1\text{-}score = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

PCG Signals



Arnott, P. J., Pfeiffer, G. W., & Tavel, M. E. (1984). Spectral analysis of heart sounds: Relationships between some physical characteristics and frequency spectra of first and second heart sounds in normals and hypertensives. Journal of Biomedical Engineering, 6(2), 121–128. doi:10.1016/0141-5425(84)90054-2

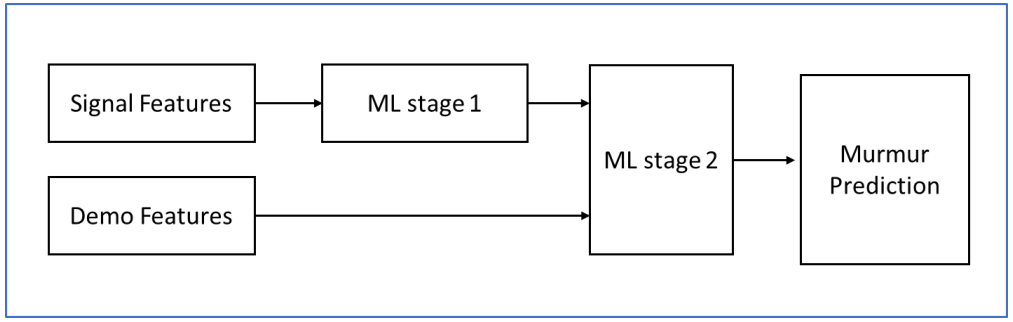
PCG Signals



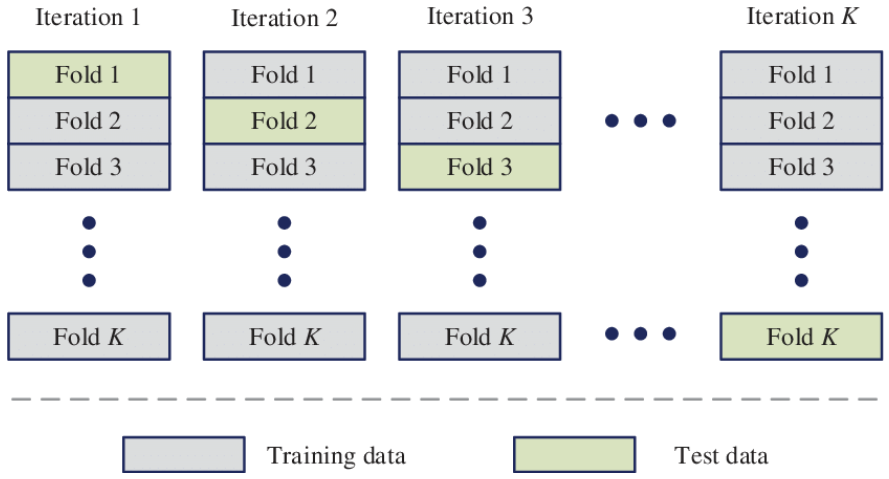
- 13 MFCCs
  - Minimum
  - Maximum
  - Mean
  - Median
  - Standard Deviation
- Zero Crossing rate
- RMS
- Spectral Rollof 90%
- Spectral Rollof 10%
- Spectral Centroid
  - Minimum
  - Maximum
  - Mean
  - Median
  - Standard Deviation

90 features per signal

ML stages



Hyperparameter exploration:



https://www.researchgate.net/figure/K-fold-cross-validation-method_fig2_331209203

Resampling:

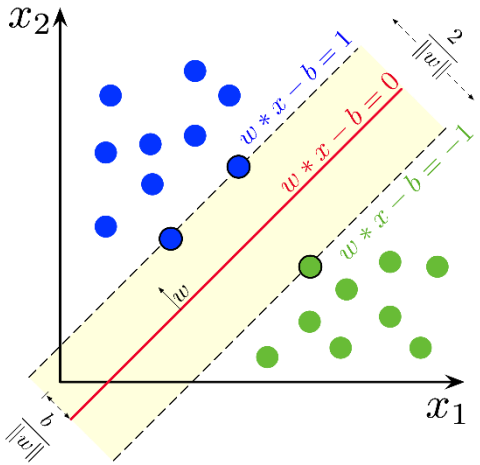
- Oversampling
- Over- and Under-sampling

Feature Selection (Signals):

- Variance
- F-value (ANOVA)
- Feature importance (Random Forest)

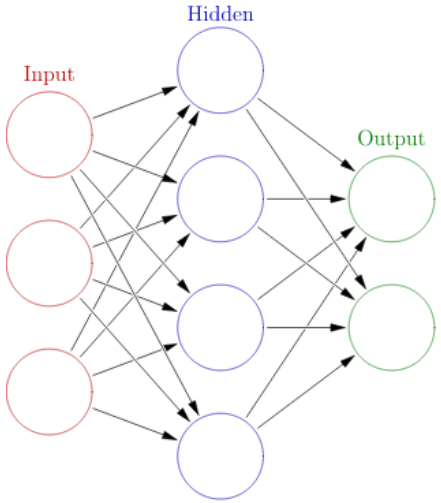
ML Algorithms

Support Vector Machine (SVM)



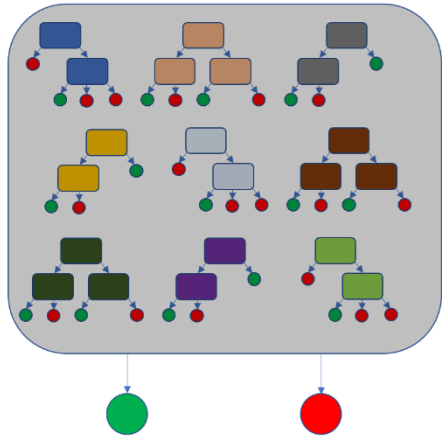
https://commons.wikimedia.org/wiki/File:SVM_margin.png

Multilayer Perceptron (MLP)



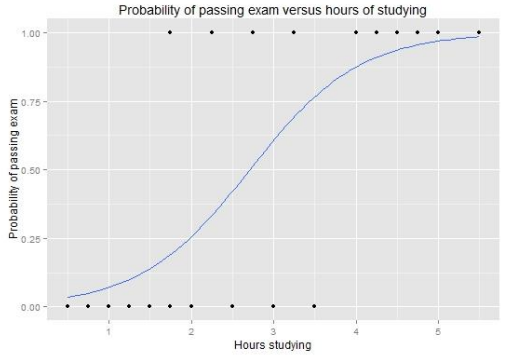
https://commons.wikimedia.org/wiki/File:Colored_neural_network.svg

Random Forest (RF)



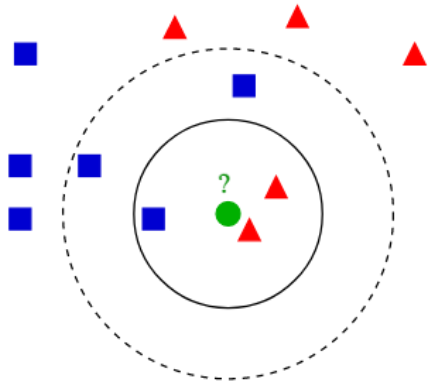
https://commons.wikimedia.org/wiki/File:Decision_Tree_vs._Random_Forest.png

Logistic Regression (LR)



https://commons.wikimedia.org/wiki/File:Exam_pass_logistic_curve.jpeg

k-nearest neighbor classification (KNN)



https://commons.wikimedia.org/wiki/File:KnnClassification.svg

Results

Hyperparameter Exploration

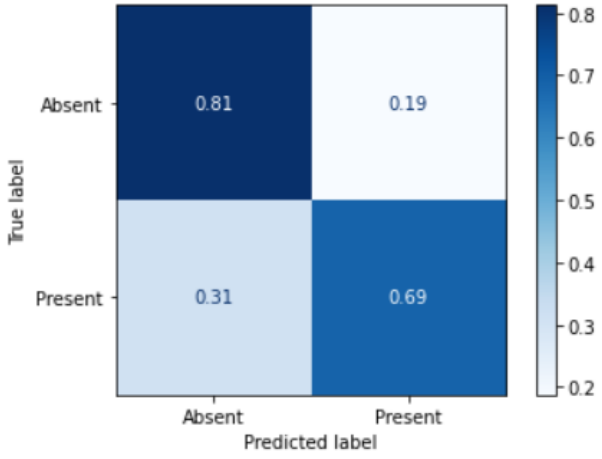
Results saved in excel, the top 10 hyperparameters were saved and will be used

ML algorithm	Hyperparameters
SVM	C Kernel Gamma
MLP	Hidden Layers Sizes Activation Functions Learning Rate Batch Size
RF	# Estimators Min samples split
LR	C Solver
KNN	# Neighbors Weights Algorithm Leaf Size

SVM

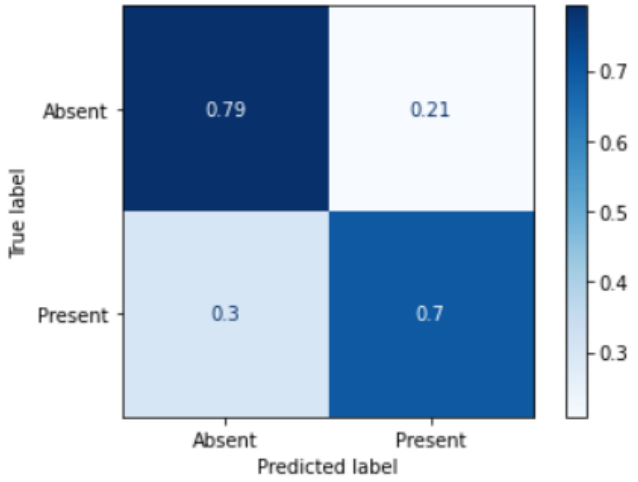
No resampling

	precision	recall	f1-score	support
Absent	0.93	0.81	0.87	533
Present	0.41	0.69	0.51	100
accuracy			0.79	633
macro avg	0.67	0.75	0.69	633
weighted avg	0.85	0.79	0.81	633



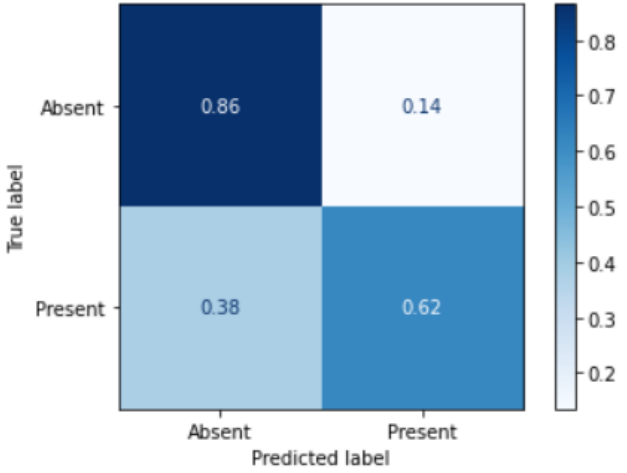
Oversampling

	precision	recall	f1-score	support
Absent	0.93	0.79	0.86	533
Present	0.39	0.70	0.50	100
accuracy			0.78	633
macro avg	0.66	0.75	0.68	633
weighted avg	0.85	0.78	0.80	633



Undersampling

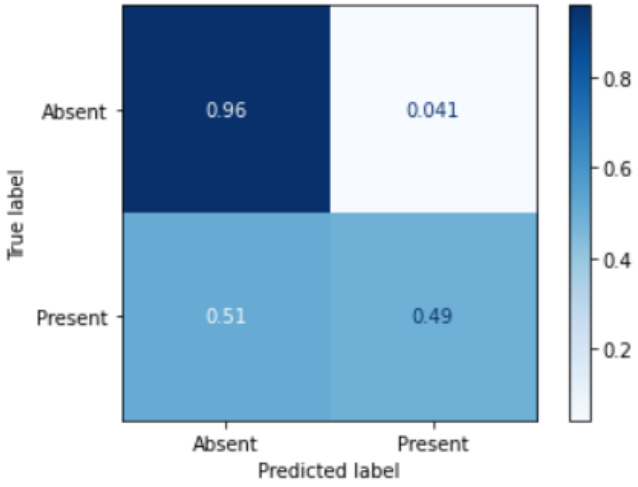
	precision	recall	f1-score	support
Absent	0.92	0.86	0.89	533
Present	0.46	0.62	0.53	100
accuracy			0.83	633
macro avg	0.69	0.74	0.71	633
weighted avg	0.85	0.83	0.84	633



MLP

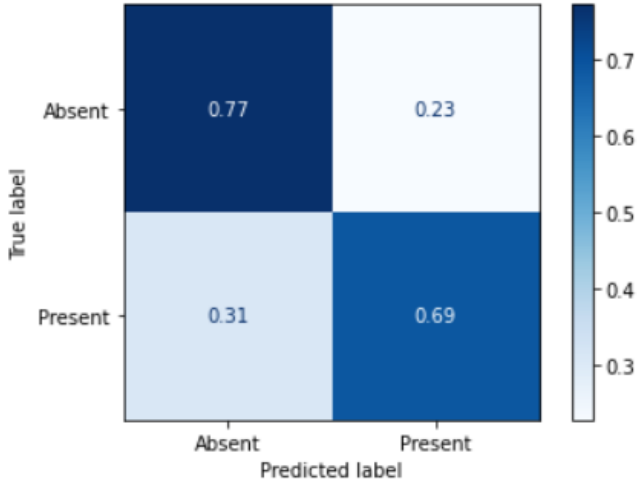
No resampling

	precision	recall	f1-score	support
Absent	0.91	0.96	0.93	533
Present	0.69	0.49	0.57	100
accuracy			0.88	633
macro avg	0.80	0.72	0.75	633
weighted avg	0.87	0.88	0.88	633



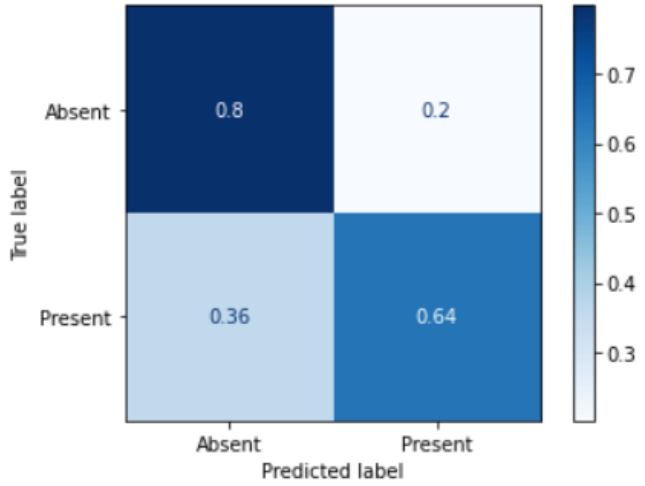
Oversampling

	precision	recall	f1-score	support
Absent	0.93	0.77	0.84	533
Present	0.36	0.69	0.48	100
accuracy			0.76	633
macro avg	0.65	0.73	0.66	633
weighted avg	0.84	0.76	0.79	633



Undersampling

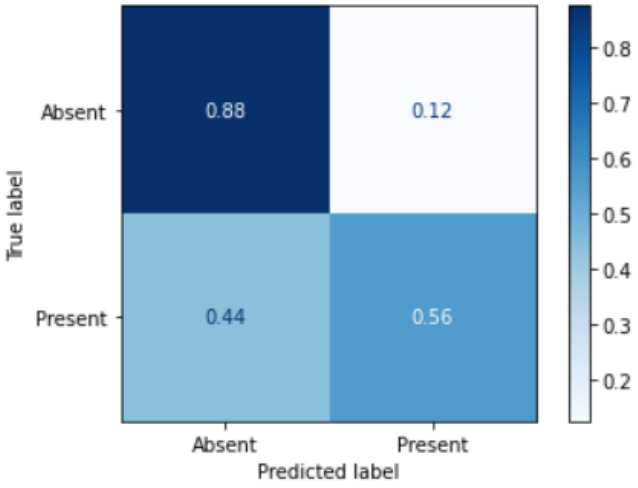
	precision	recall	f1-score	support
Absent	0.92	0.80	0.86	533
Present	0.37	0.64	0.47	100
accuracy			0.77	633
macro avg	0.65	0.72	0.66	633
weighted avg	0.84	0.77	0.79	633



RF

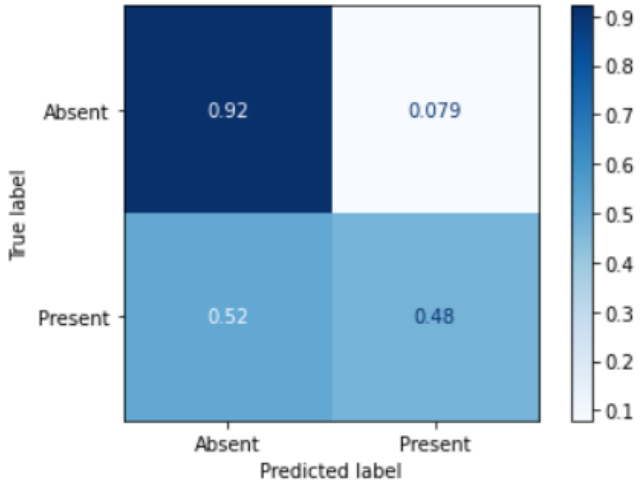
No resampling

	precision	recall	f1-score	support
Absent	0.91	0.88	0.89	533
Present	0.46	0.56	0.50	100
accuracy			0.83	633
macro avg	0.69	0.72	0.70	633
weighted avg	0.84	0.83	0.83	633



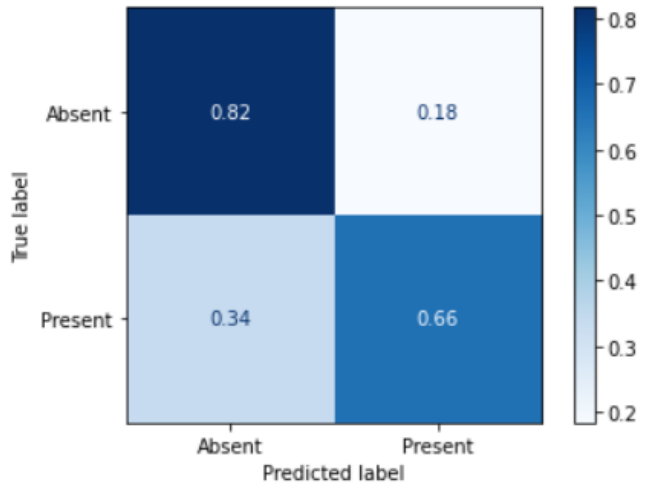
Oversampling

	precision	recall	f1-score	support
Absent	0.90	0.92	0.91	533
Present	0.53	0.48	0.51	100
accuracy			0.85	633
macro avg	0.72	0.70	0.71	633
weighted avg	0.85	0.85	0.85	633



Undersampling

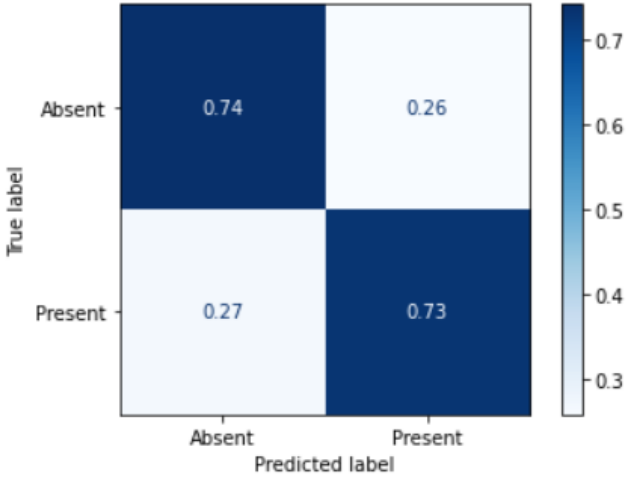
	precision	recall	f1-score	support
Absent	0.93	0.82	0.87	533
Present	0.40	0.66	0.50	100
accuracy			0.79	633
macro avg	0.66	0.74	0.68	633
weighted avg	0.84	0.79	0.81	633



LR

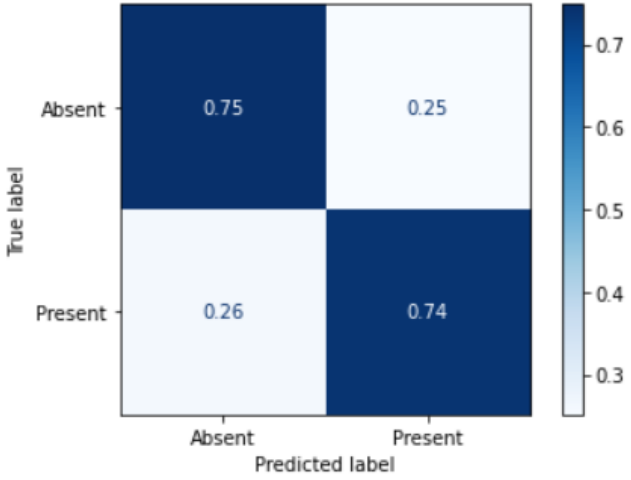
No resampling

	precision	recall	f1-score	support
Absent	0.94	0.74	0.83	533
Present	0.35	0.73	0.47	100
accuracy			0.74	633
macro avg	0.64	0.74	0.65	633
weighted avg	0.84	0.74	0.77	633



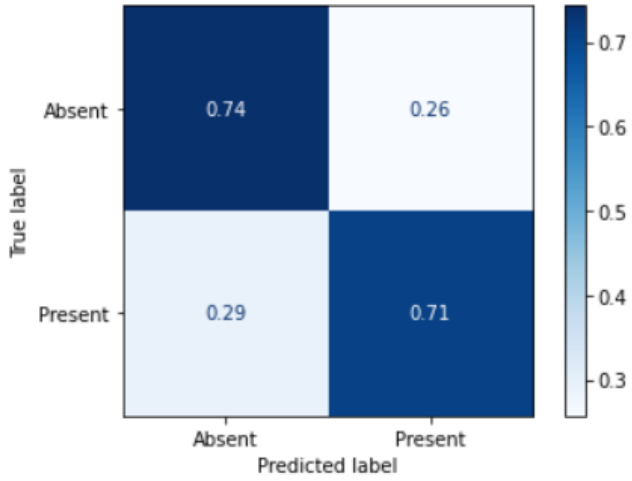
Oversampling

	precision	recall	f1-score	support
Absent	0.94	0.75	0.83	533
Present	0.36	0.74	0.48	100
accuracy			0.75	633
macro avg	0.65	0.74	0.66	633
weighted avg	0.85	0.75	0.78	633



Undersampling

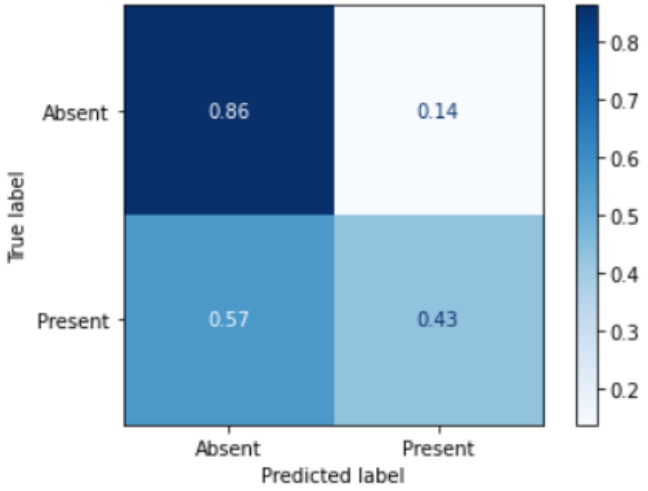
	precision	recall	f1-score	support
Absent	0.93	0.74	0.83	533
Present	0.34	0.71	0.46	100
accuracy			0.74	633
macro avg	0.64	0.73	0.64	633
weighted avg	0.84	0.74	0.77	633



KNN

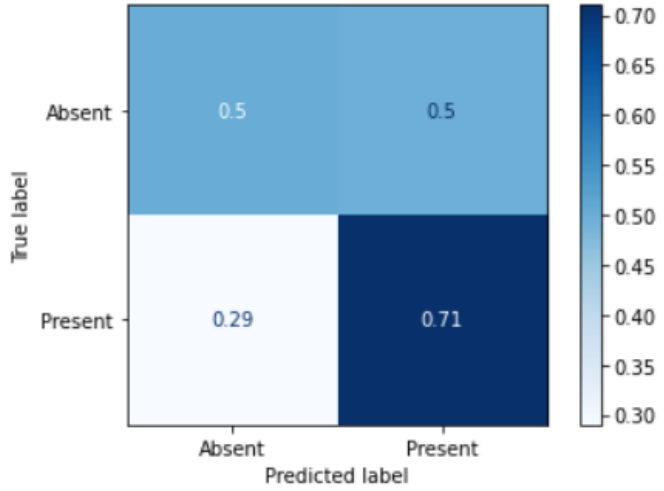
No resampling

	precision	recall	f1-score	support
Absent	0.89	0.86	0.88	533
Present	0.37	0.43	0.40	100
accuracy			0.79	633
macro avg	0.63	0.65	0.64	633
weighted avg	0.81	0.79	0.80	633



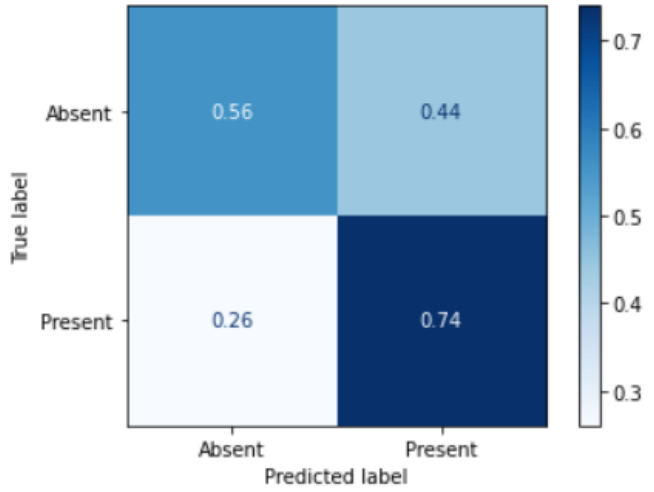
Oversampling

	precision	recall	f1-score	support
Absent	0.90	0.50	0.65	533
Present	0.21	0.71	0.33	100
accuracy			0.54	633
macro avg	0.56	0.61	0.49	633
weighted avg	0.79	0.54	0.60	633



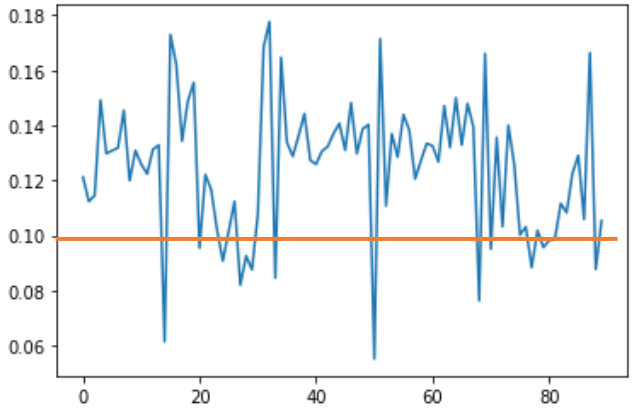
Undersampling

	precision	recall	f1-score	support
Absent	0.92	0.56	0.69	533
Present	0.24	0.74	0.36	100
accuracy			0.59	633
macro avg	0.58	0.65	0.53	633
weighted avg	0.81	0.59	0.64	633



Feature Selection

Standard deviation > 0.1



AND

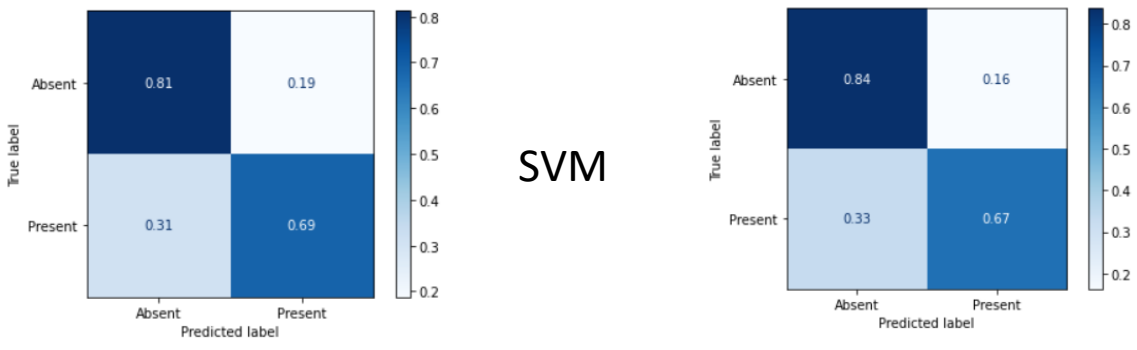
The 70 higher F-values

AND

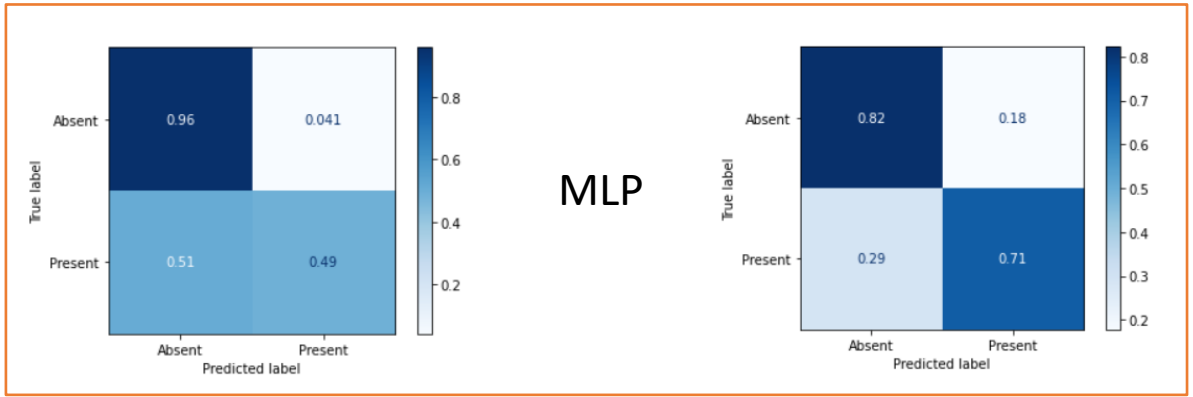
The 70 most important features

49 features

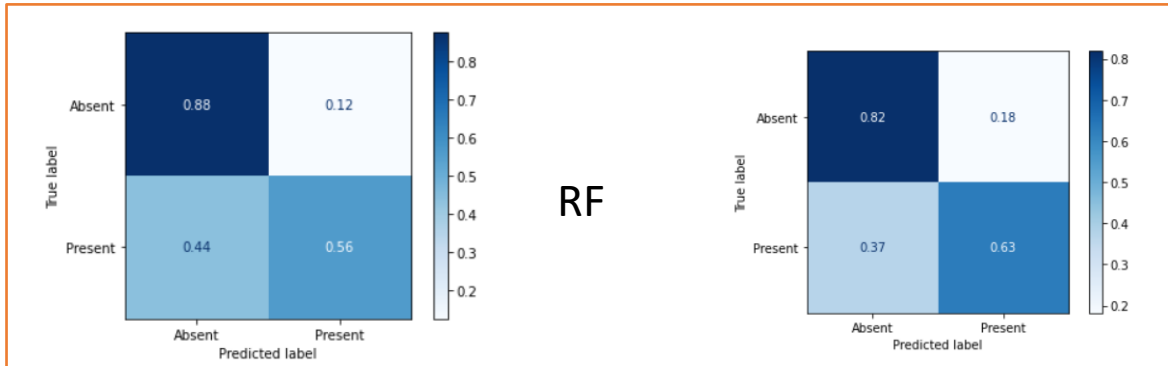
Feature Selection Results



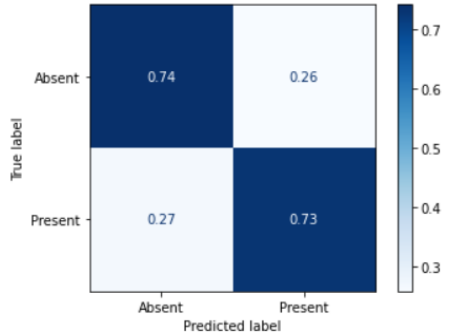
SVM



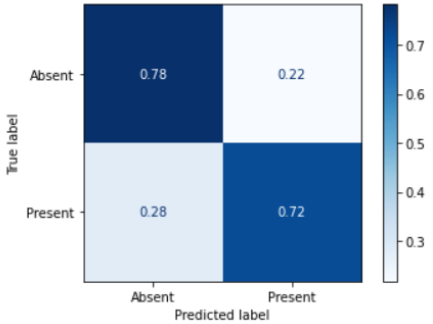
MLP

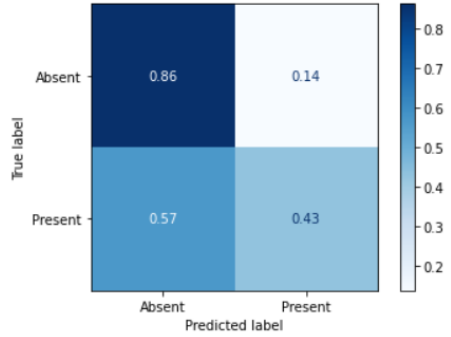


RF

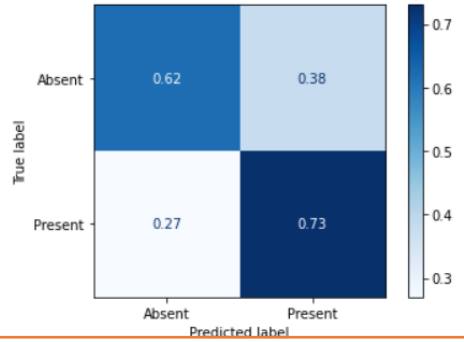


LR

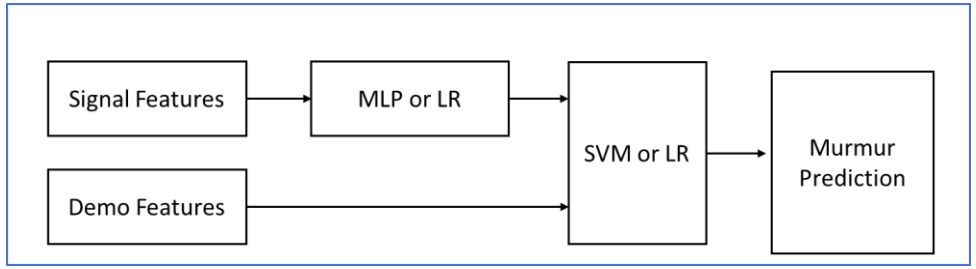


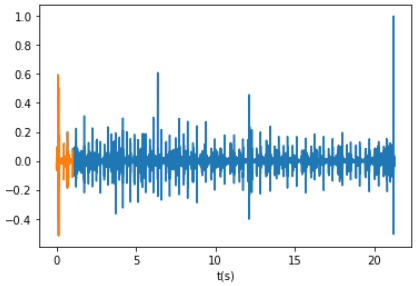


KNN



Final Model



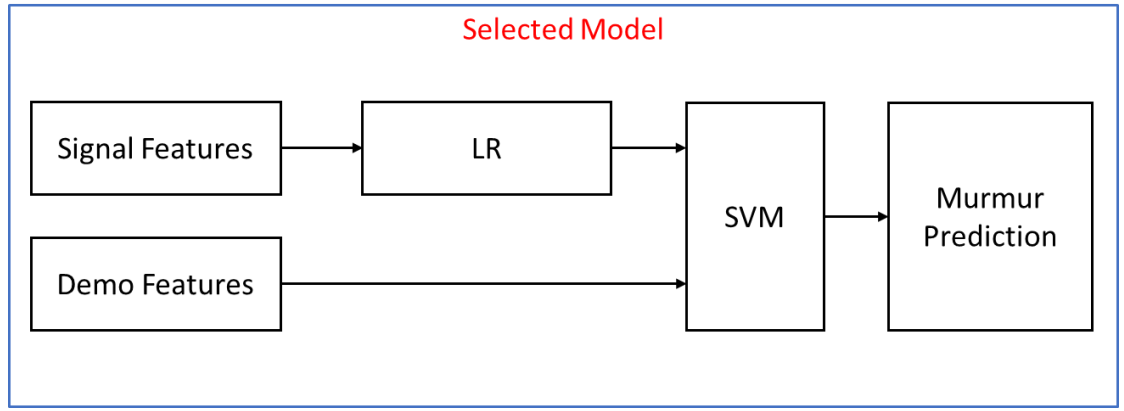


2530 - Notepad

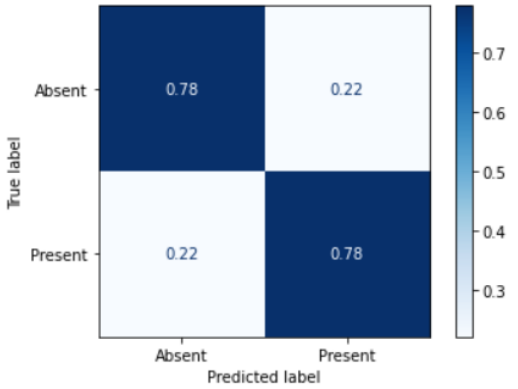
File Edit Format View Help

```
2530 4 4000
AV 2530_AV.hea 2530_AV.wav 2530_AV.tsv
PV 2530_PV.hea 2530_PV.wav 2530_PV.tsv
TV 2530_TV.hea 2530_TV.wav 2530_TV.tsv
MV 2530_MV.hea 2530_MV.wav 2530_MV.tsv
#Age: Child
#Sex: Female
#Height: 98.0
#Weight: 15.9
#Pregnancy status: False
#Murmur: Absent
#Murmur locations: nan
#Most audible location: nan
```

Selected Model



	precision	recall	f1-score	support
Absent	0.94	0.78	0.85	153
Present	0.45	0.78	0.57	36
accuracy			0.78	189
macro avg	0.69	0.78	0.71	189
weighted avg	0.84	0.78	0.80	189

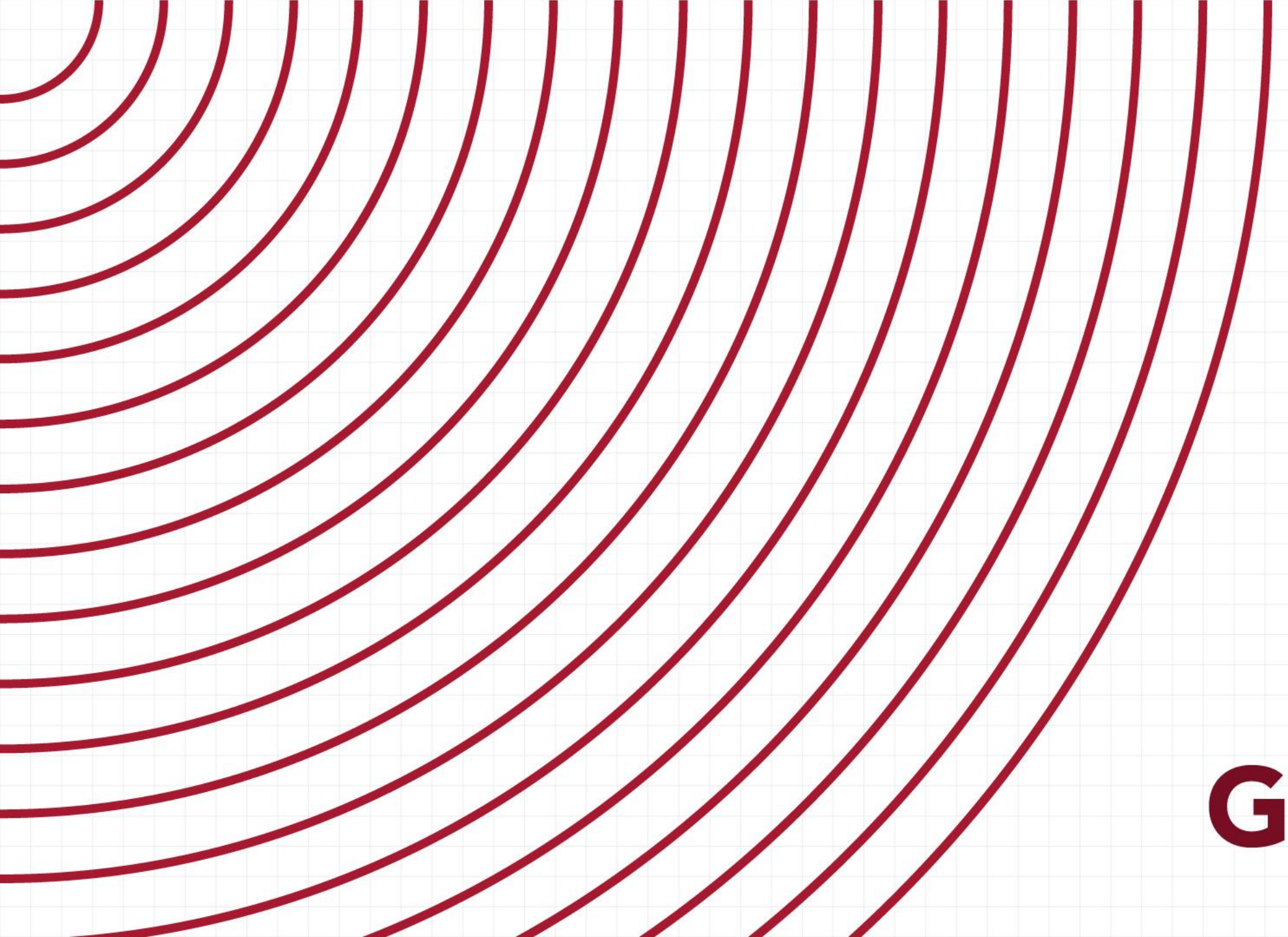


CONCLUSIONS

- Development of a model for detecting heart murmurs in patients.
- 78% of balanced accuracy.
- Patient screening tool.
- Importance for low-income countries.
- Use of deep learning can improve results.

Base de Datos

https://moody-challenge.physionet.org/2022/



ESCUELA COLOMBIANA DE INGENIERÍA JULIO GARAVITO

UNIVERSIDAD

Gracias