ECG Signal Analysis summary

C. Korgialas, Prof. Ioannis Pitas
Aristotle University of Thessaloniki
pitas@csd.auth.gr
www.aiia.csd.auth.gr
Version 3.5.1

Artificial Intelligence &
Information Analysis Lab

Contents

Introduction

The heart is perhaps the most important organ of the body as it supplies blood and oxygen to the other organs. However, sometimes, has problems that make it non-functional.

The heart muscle is checked with an ***Electrocardiogram*** (ECG) which detects heart abnormalities by measuring electrical activity.

Introduction

However, the technology revolution, in addition to recording heart disease, it is also their diagnosis and prediction.

By having plenty of ECGs gave the scientists the opportunity to create machine learning algorithms. These algorithms will be able to categorize any heart disease giving the doctor the opportunity to apply individualized treatment.

Contents

Background Knowledge of ECG Signals

The ECG consists of three different parts:

- ***First part*** concerns the peaks (P, Q, R, S, T and U).
- ***Second part*** includes the intervals (PR, RR, QRS, ST and QT) and finally the third parts contains the segments which are PR and ST.

The amplitude of the ECG signal is calibrated on the vertical axis by taking the 10 mm = 1 mV value [BUG18].

Background Knowledge of ECG Signals

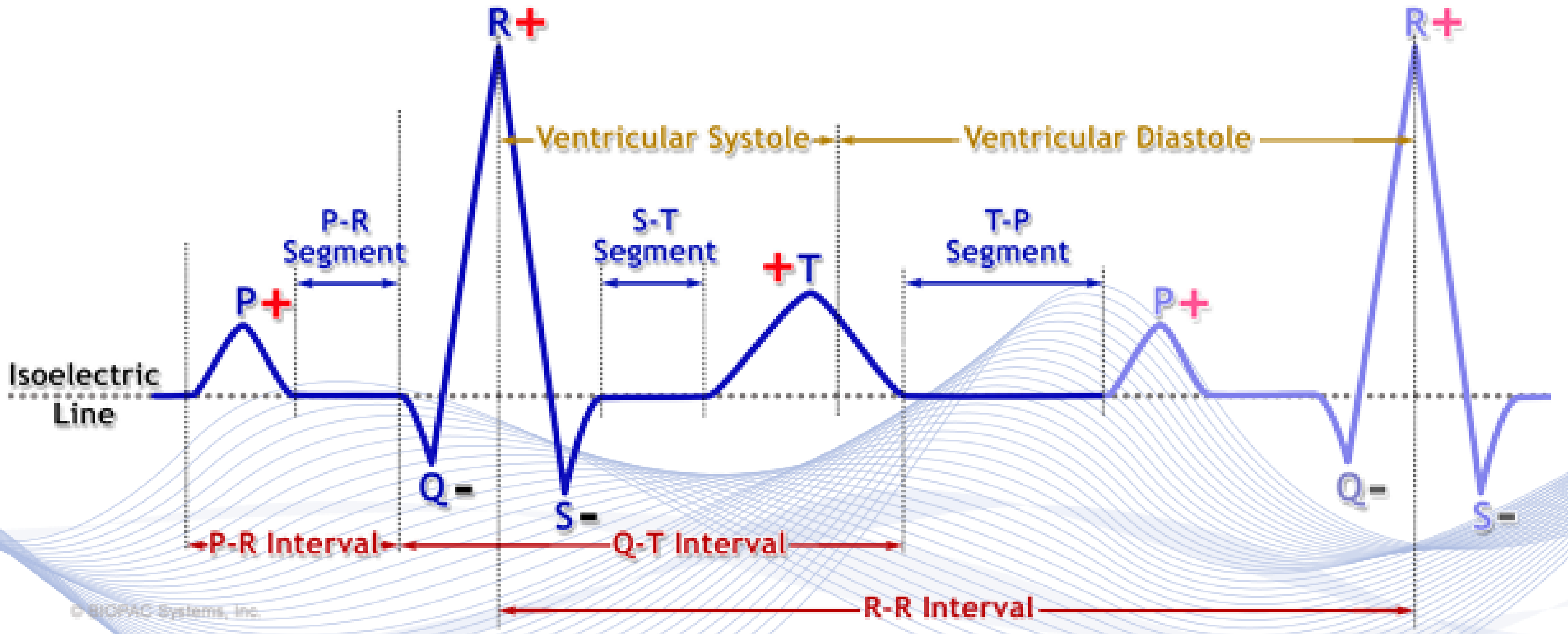


Figure 1: Typical ECG Signal waveform [BIO].

Artificial Intelligence & Information Analysis Lab

Background Knowledge of ECG Signals

The basic building blocks of a heart pulse are the following:

- **P wave,**
- **QRS complex**
- **T wave.**

The ***Atria's*** construction is represented with the ***P wave*** while the ***Depolarization*** of ventricles is represented by ***QRS complex***.

The T wave is caused by the return of ventricles to its resting state.

Contents

Issues in ECG Classification

In this section we will describe some of the most important challenges and problems coming from the classification of ECG signals. This is done for the purpose of every researcher who wants to study the specific field to have knowledge of the specifics of cardiac signals [JPD15].

Contents

Materials and Machine Learning Methods

Machine learning algorithms are the main part of artificial intelligence for the purpose of predicting through education. Machine learning algorithms work by discovering hidden patterns in a set of data.

Then, after training on the existing data, they can discover patterns in new data. This method gives the system a kind of intelligence making it smart and capable of making decisions [AJQ18].

Materials and Machine Learning Methods

An important part of machine learning is the preprocessing of data. Then after the algorithm is trained with healthy and unhealthy data, in this case ECG signals, it receives data from the patient.

The entry is considered a classifier who in the end makes the prediction of a heart disease. In Figure 3 we observe the central idea of a machine learning system [AJQ18].

Materials and Machine Learning Methods

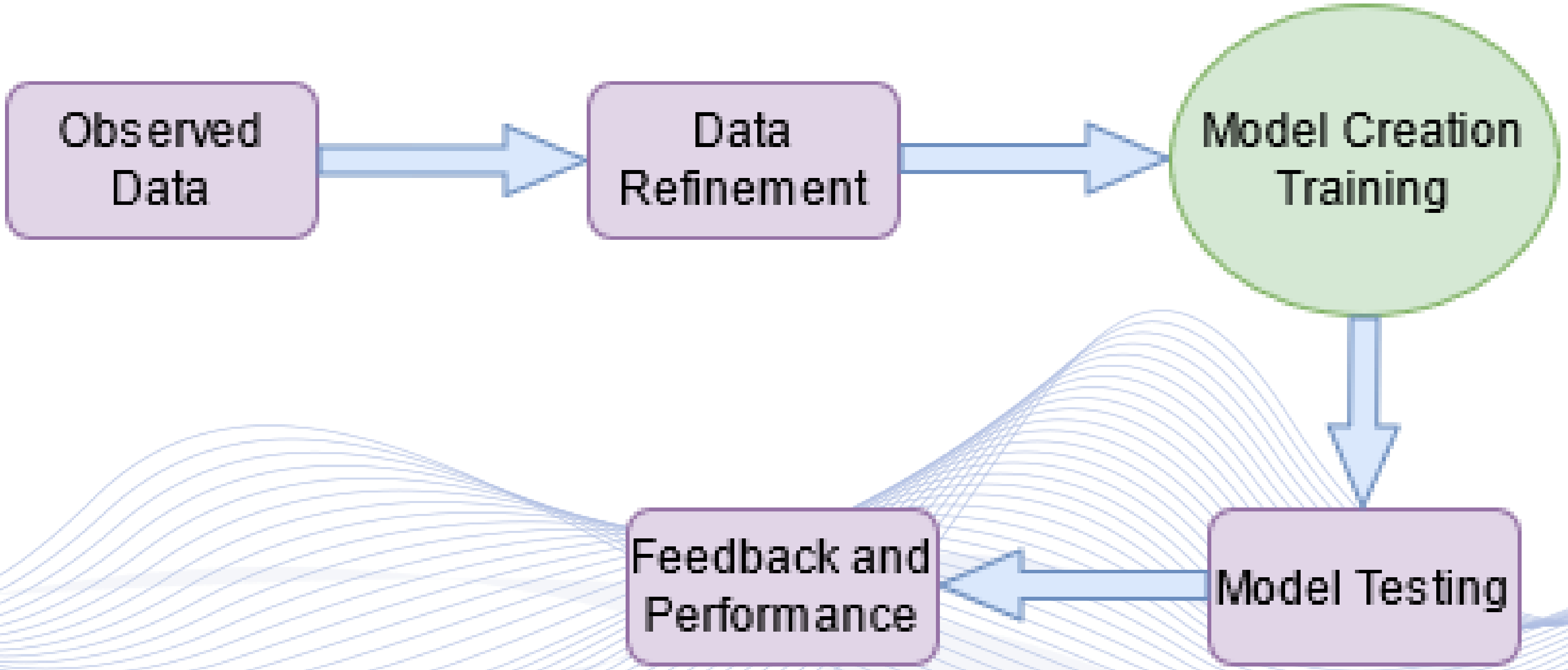


Figure 3: Machine Learning idea representation.

Contents

- Introduction
- Background Knowledge of ECG Signals
- Issues in ECG Classification
- **Materials and Machine Learning Methods**
  - **Dataset**
  - Preprocessing of the Data
  - Feature Selection Algorithms
  - Dimensionality Reduction
  - Machine Learning Classifiers
  - Validation and Evaluation Analysis of Classifiers
- Deep Learning Techniques for ECG Classification
- Heart Disease Classification
  - Cardiac Arrhythmia Classification
  - Heart Failure Classification
  - Cardiac Valve Stenosis Classification
  - Congenital Heart Disease Classification (CHD)
  - Coronary Artery Disease Classification
  - Blood Pressure Hypertension Classification

Dataset

As mentioned in the introduction, the important step in setting up a mechanical learning algorithm is to find the right dataset. It the most works it was used the "Heart Disease Dataset" of the UCI (University of California) Machine Learning Repository [ECG].

Dataset

The dataset which is most referred in literature, has a sample size of 303 patients, 76 features, and some missing values. Some of the most important features are ***Sex, Age, Slope, CP*** and more.

The target label output has two classes and is used to represent a Heart Disease or a normal situation [HLM18].

Contents

Preprocessing of the Data

Data preprocessing is entered into the learning algorithm is important in order to mitigate the noise or errors or missing data and make it as clear and recognizable as possible. This will help the classifier to train properly and effectively.

Preprocessing techniques such as ***removing of missing values***, ***standard scalar***, and ***MinMax Scalar*** have been applied to the dataset for effective use in the classifiers [HLM18].

Contents

Feature Selection Algorithms

Feature selection and ***extraction*** are important in a machine learning model as features irrelevantly often affect its performance.

The choice of features creates better classification, and the algorithm becomes more accurate in terms of implementation time, offering greater accuracy. Some of the most known feature selection models are ones that chose the most important features [HLM18].

Contents

Dimensionality Reduction

Dimensionality reduction transforms the data from high dimension to low dimension. This is achieved by reducing the variables. It is worth noting that the structure is maintained while achieving the extraction of raw data from our data.

The most used dimensionality reduction methods for ECG signals are chi-square test and ***Principal Component Analysis*** (PCA) [GEE20].

Dimensionality Reduction

- ***Chi – Square (CHI)***

Chi-square (CHI) works by sorting the class and filter on which the class tag depends. ***ChiSqSelector*** (CHI) of Apache Spark MLlib is used for feature selection in the model construction. Also, CHI filters the features and sorts them, through repeated iterations, for selection.

Contents

Machine Learning Classifiers

The only way to diagnose a Heart Disease from our data is to perform classification. ***Classification*** is a supervised Machine Learning technique which predicts cases based on a previous dataset.

The main classification techniques for ECG signals will be mentioned below. In Figure 4 we observe the basic idea behind the classification algorithm [TC19].

Machine Learning Classifiers

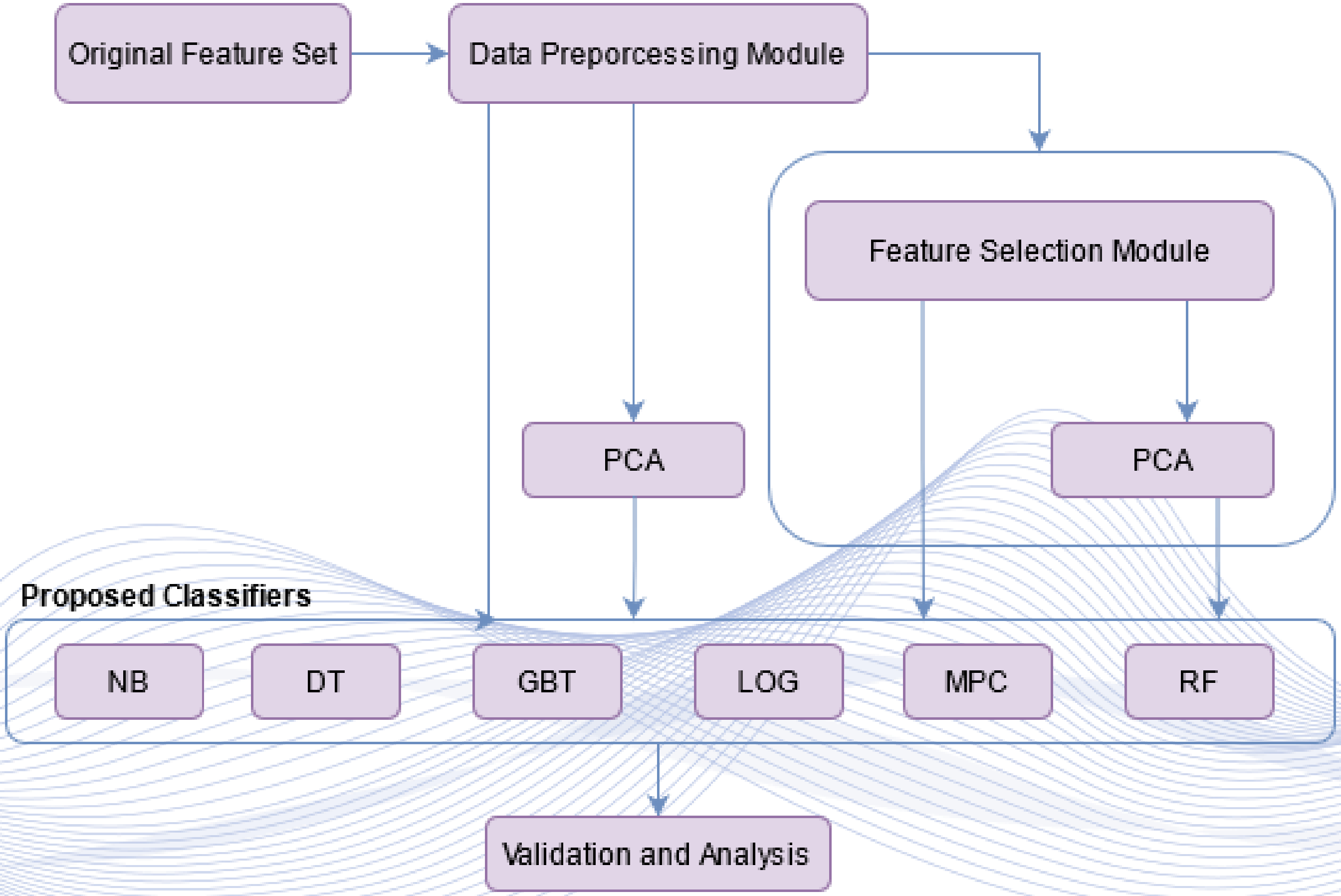


Figure 4: Schematic diagram of Classification process.

Machine Learning Classifiers

- ***Support Vector Machine (SVM)***

A ***Support Vector Machine*** (SVM) consists a hyper plane that creates boundaries between points of data plotted which represent tuples and their feature values.

Also, it has a high computational accuracy. SVM uses a maximum margin strategy that transformed into solving a complex quadratic programming problem.

Contents

Validation and Evaluation Analysis

- ***Validation***

In validation method it can be used the k –fold ***cross-validation*** (CV) method. In this method the dataset is divided into k equal size of parts, in which $k-1$ groups are used to train the classifiers and remaining part is used for checking out performance in each step. The process of validation is repeated k times [HLM18].

The classifier performance is computed based on k results.

Validation and Evaluation Analysis

- ***Evaluation***

The last part of the classification algorithm is the evaluation process. In this method a confusion matrix is used to find the results performance. The confusion matrix elements are [GEE20]:

- True Positive (TP)
- True Negative (TN)
- False Negative (FN)
- False positive (FP)

Validation and Evaluation Analysis

- ***Evaluation***

$$Accuracy = \frac{(TP + TN)}{TP + FN + FP + TN}$$

$$Error = \frac{(FP + FN)}{TP + TN + FP + FN}$$

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

Validation and Evaluation Analysis

- ***Evaluation***

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

$$MCC = \frac{TP \times TN + FP - FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

MCC represents the prediction ability of a classifier with values between $[-1, +1]$. If the value of the MCC classifier is $+1$, the classifier predictions are ideal where -1 indicates that are wrong.

Contents

Deep Learning Techniques for ECG Classification

One of the proposing solution is to use ***Deep Learning*** architectures where first layers of convolutional neurons behave as feature extractors and in the end some fully - connected (FCN) layers are used for making final decision about ECG classes [PYA17].

Deep Learning Techniques for ECG Classification

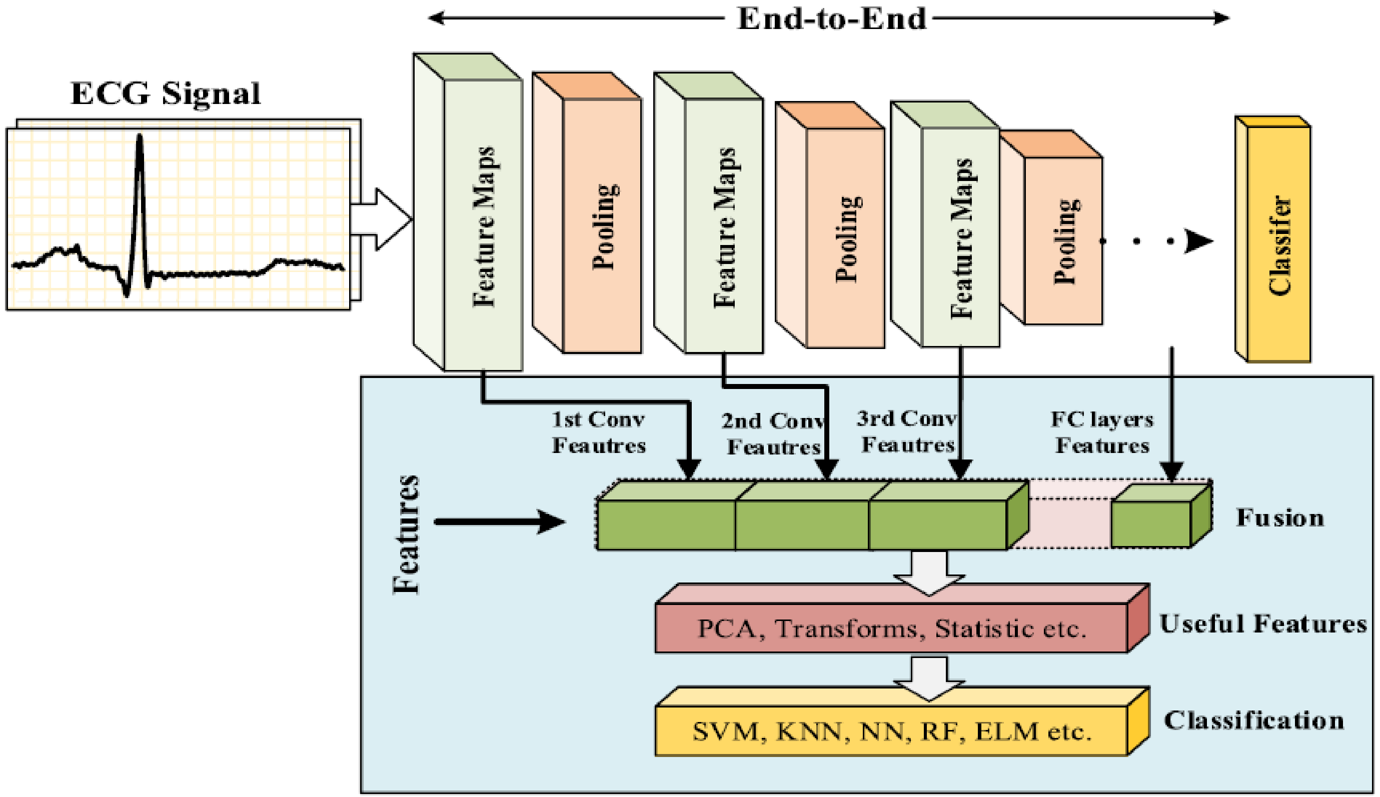


Figure 5: CNN architecture for ECG Signal processing [MUR20].

Artificial Intelligence & Information Analysis Lab

Deep Learning Techniques for ECG Classification

Therefore, many of the traditional machine learning methods are of interest for the feature extraction stage. By using appropriate ***feature selection*** and ***size reduction*** methods on attributes obtained from low-to-high level by convolution, useful input sets for shallow or deep classifiers can be obtained.

In addition, some handcrafted features are added to the feature set in order to improve classification performance. In a scenario as presented in Figure 6 [MUR20].

Deep Learning Techniques for ECG Classification

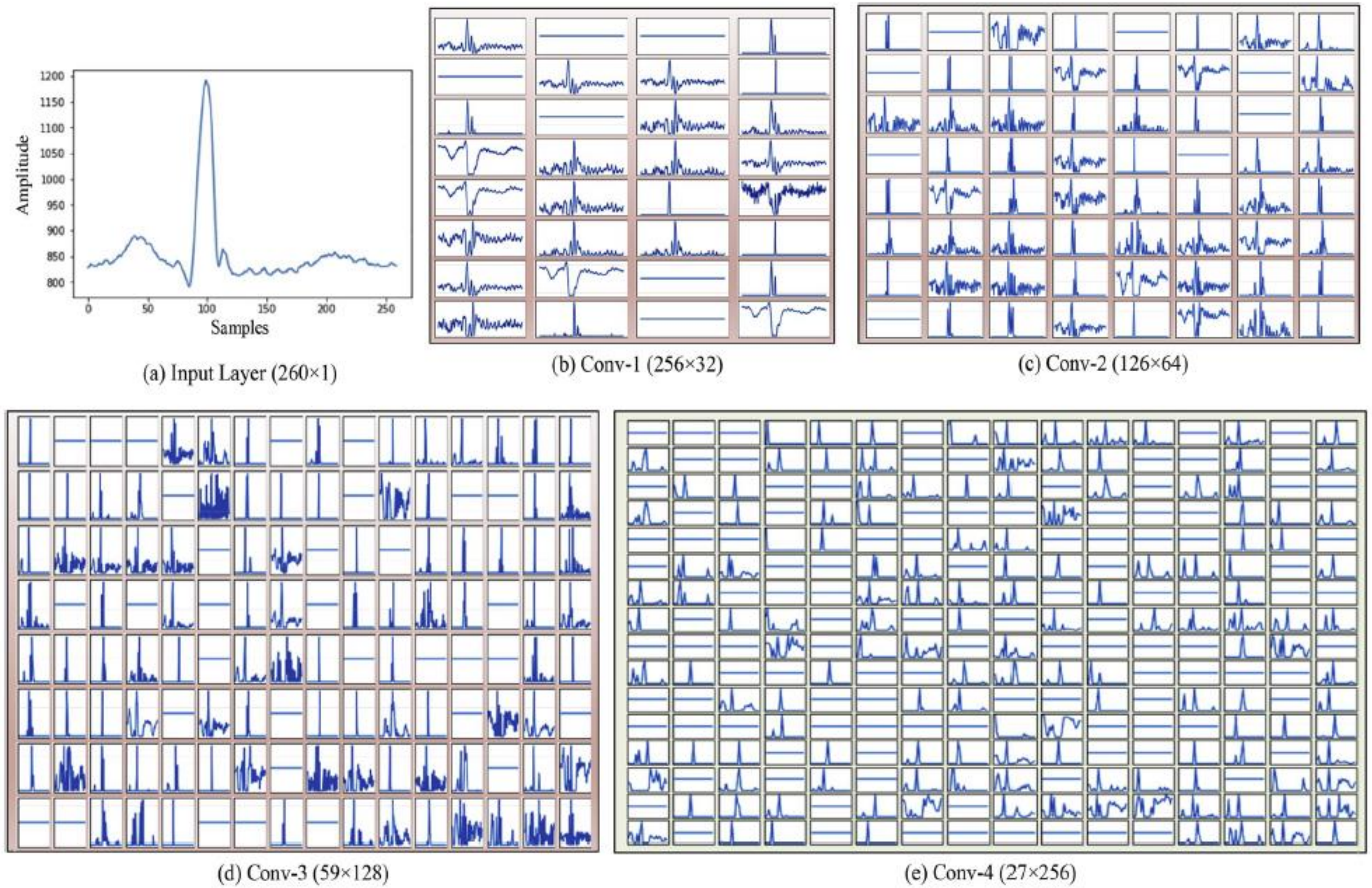


Figure 6: ECG features as they are processed by the CNN [MUR20].

Contents

Heart Disease Classification

In this section we will describe the most common ***Heart Diseases*** and its classification through the ECG datasets.

The aim of these efforts is the early detection of heart disease in general. Early detection can rescue the patient’s life or prevent permanent damage to human organs.

Arrhythmia is the heart problem with the most focus. Apart from arrhythmia, there are also some studies on other disease types [BUG18].

Contents

Cardiac Arrhythmia Classification

Cardiac Arrhythmia consist one of the most common heart disease in the world. Although it is a condition that is not life threatening.

An early diagnosis or prognosis can enable the treating physician to develop the appropriate treatment for the patient.

Cardiac Arrhythmia Classification

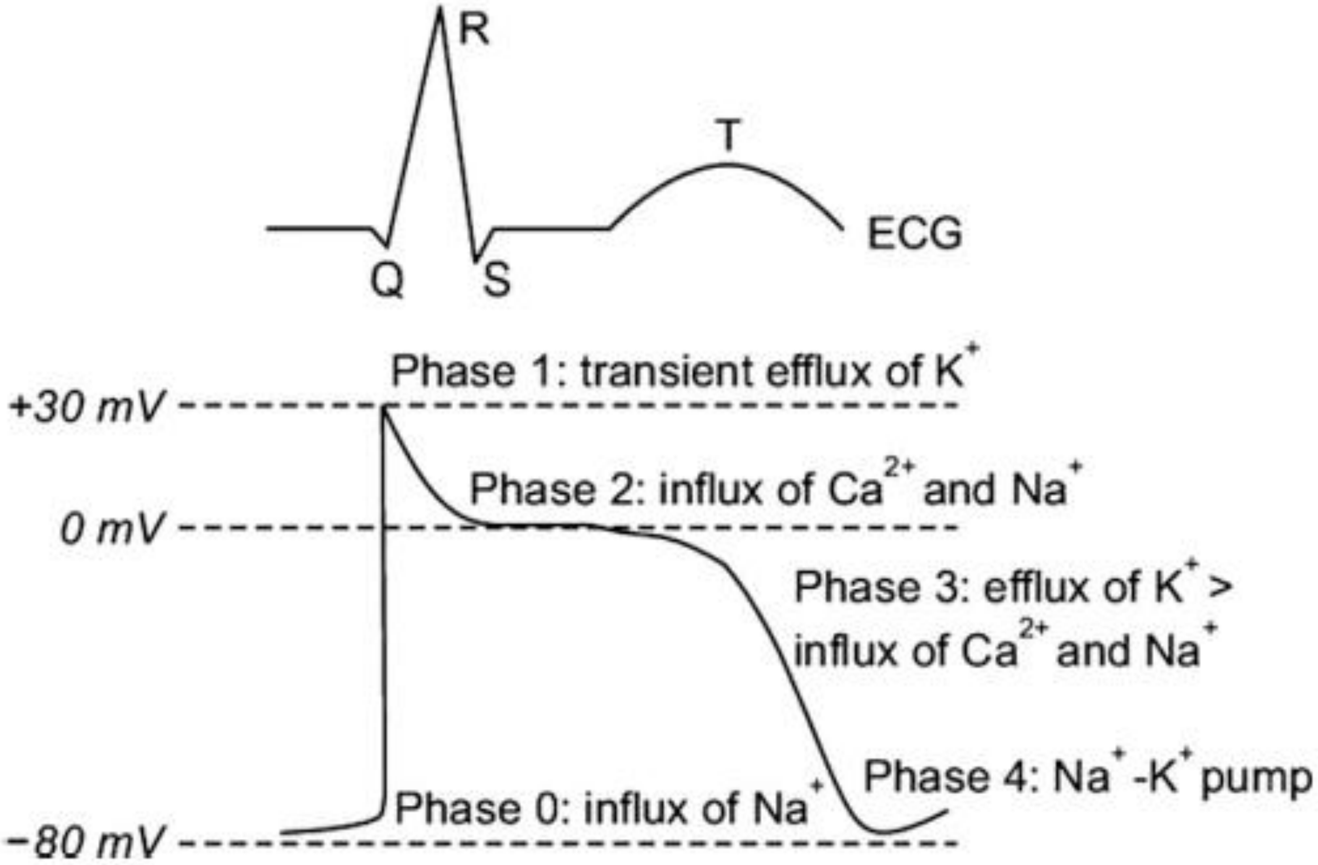


Figure 7: Cardiac actions potentials [YJ21].

Cardiac Arrhythmia Classification

Figure 8: Some of the most important cardiac arrhythmias [ARR].

Cardiac Arrhythmia Classification

- ***RR Interval Approach***

RR intervals are the differences of successive heart beats (R-peaks) in an ECG as the result of an automated heartbeat detection**.**

Currently some conventional mobile monitoring systems can save RR intervals or the heart rate as a processed information [VOL17].

Cardiac Arrhythmia Classification

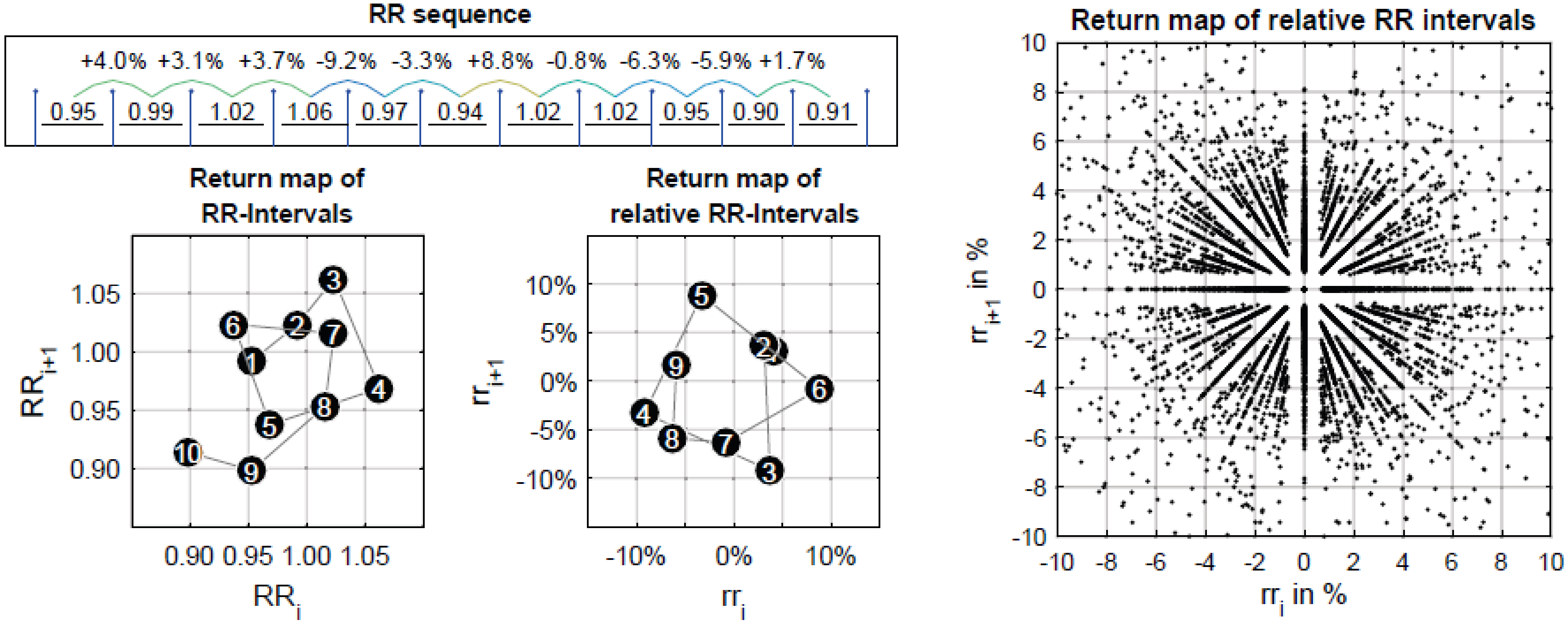


Figure 10: Left: Short sequence of absolute and relative RR intervals and the return maps Right: Coordinates of 10000 random value pairs [VOL17].

Cardiac Arrhythmia Classification

- ***Deep Learning Approach 1/2***

The proposed model consists of two models, a CNN and a LSTM model. Our proposed system classifies the ECG signal of varying lengths if the test segment consists of only a single type of arrhythmia.

This is not always true, as the ECG signal may contain more than one type of arrhythmia [OH18].

Cardiac Arrhythmia Classification

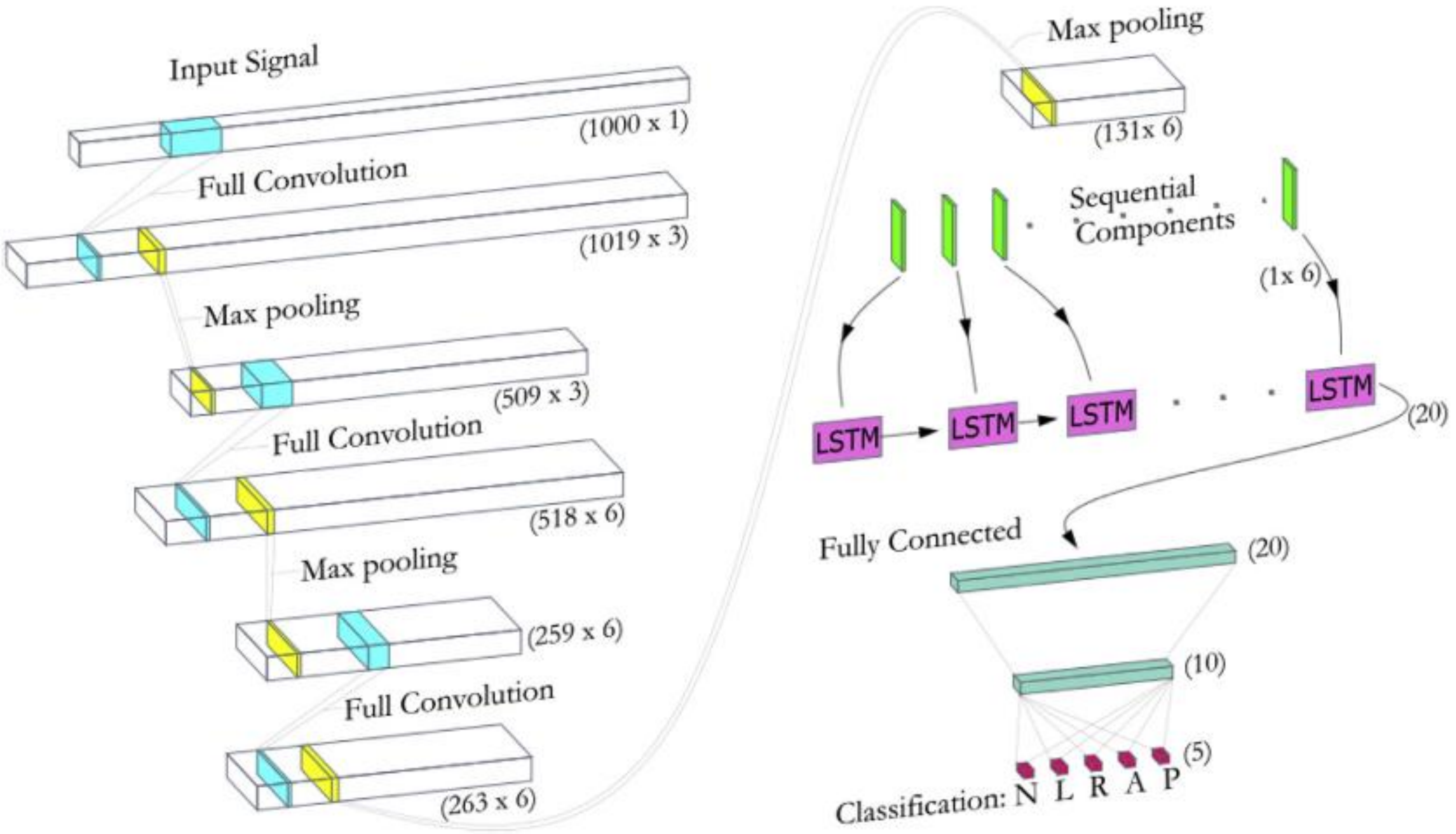


Figure 11: Hybrid CNN - LSTM architecture for arrhythmia classification [OH18].

Cardiac Arrhythmia Classification

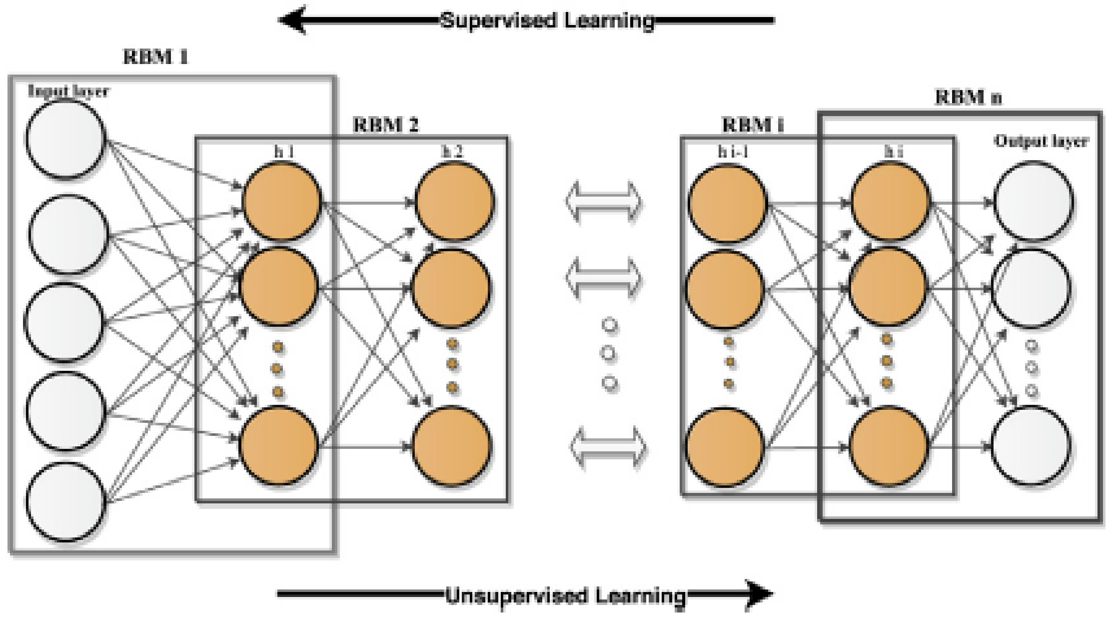


Figure 12: Deep Belief Network (DBN) architecture [EBR20].

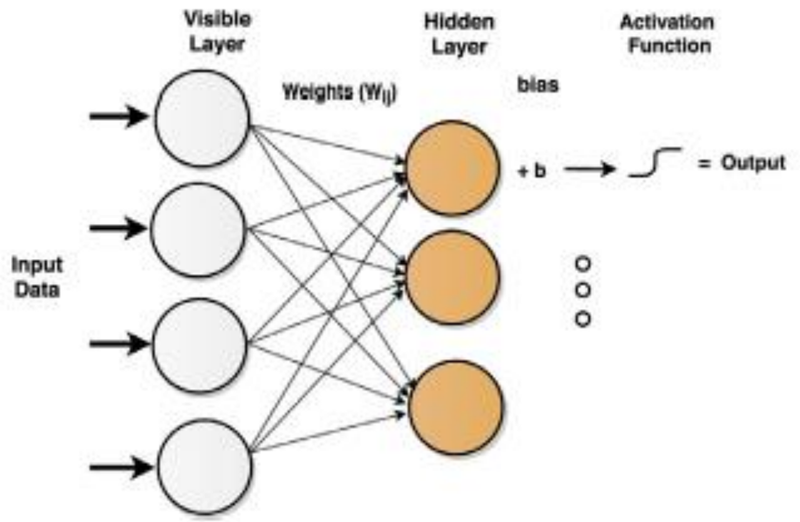


Figure 13: The architecture of RBM model. White nodes are Visible Units and brown nodes are Hidden Units [EBR20].

Contents

Heart Failure Classification

Heart Failure (HF) occurs when the heart cannot pump enough blood to meet the needs of the body. Machine learning predicts patients' survival from their data and can individuate the most important features among those included in their medical records.

ECG is an important way to diagnose heart failure and other heart conditions as already mentioned in this study.

Heart Failure Classification

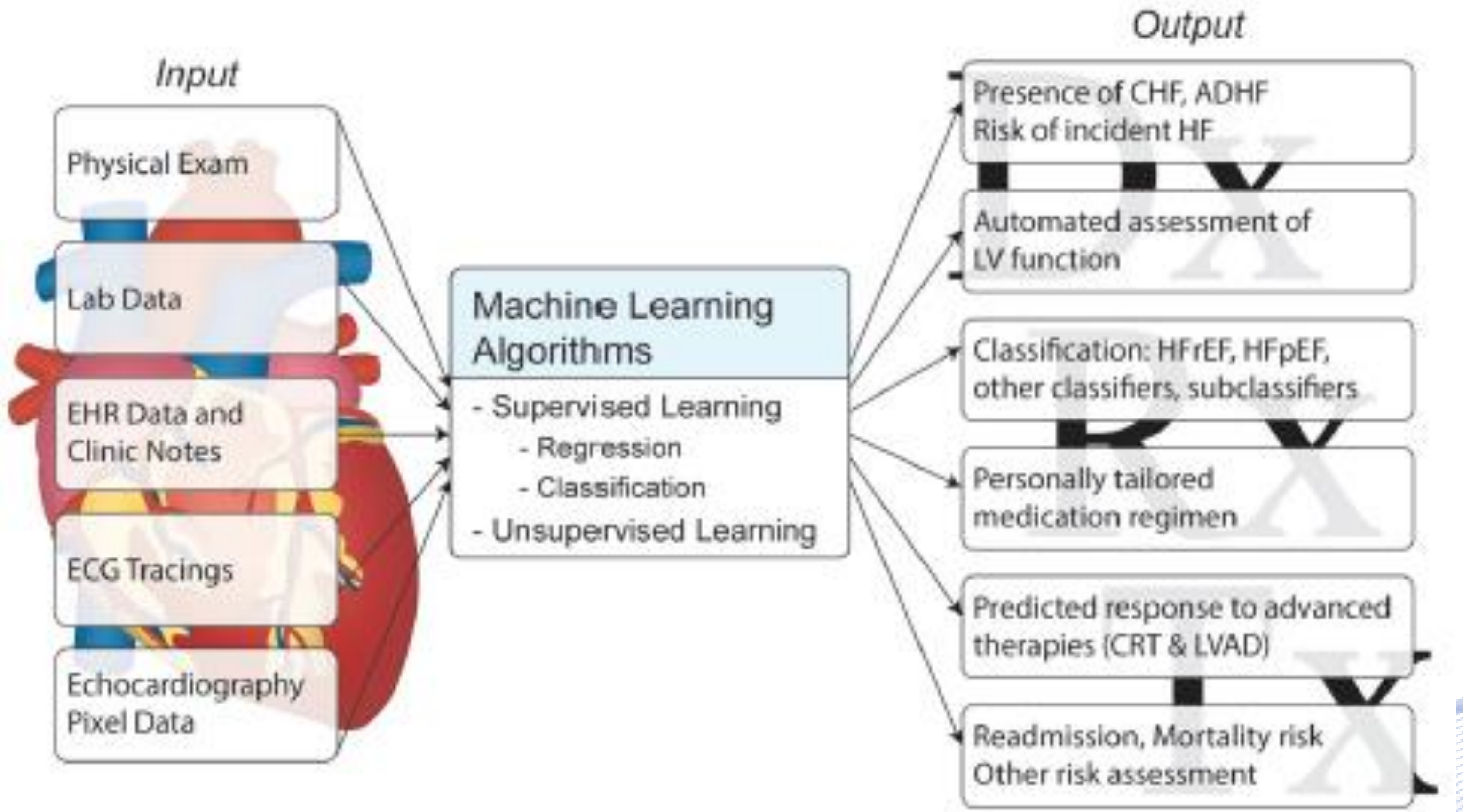


Figure 14: Heart Failure Machine Learning Idea [OMA20].

Heart Failure Classification

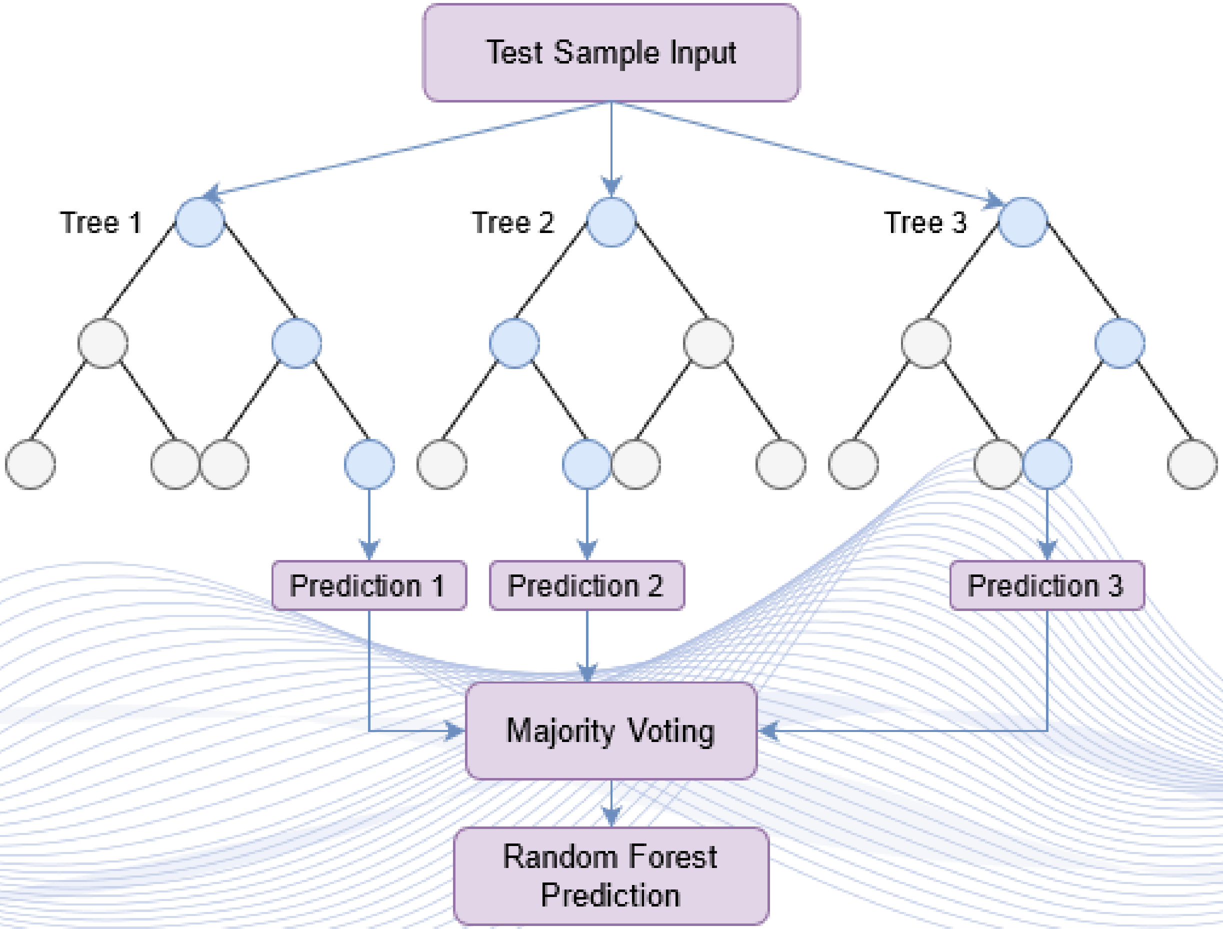


Figure 15: Random Forest algorithm schematic diagram.

Artificial Intelligence & Information Analysis Lab

Heart Failure Classification

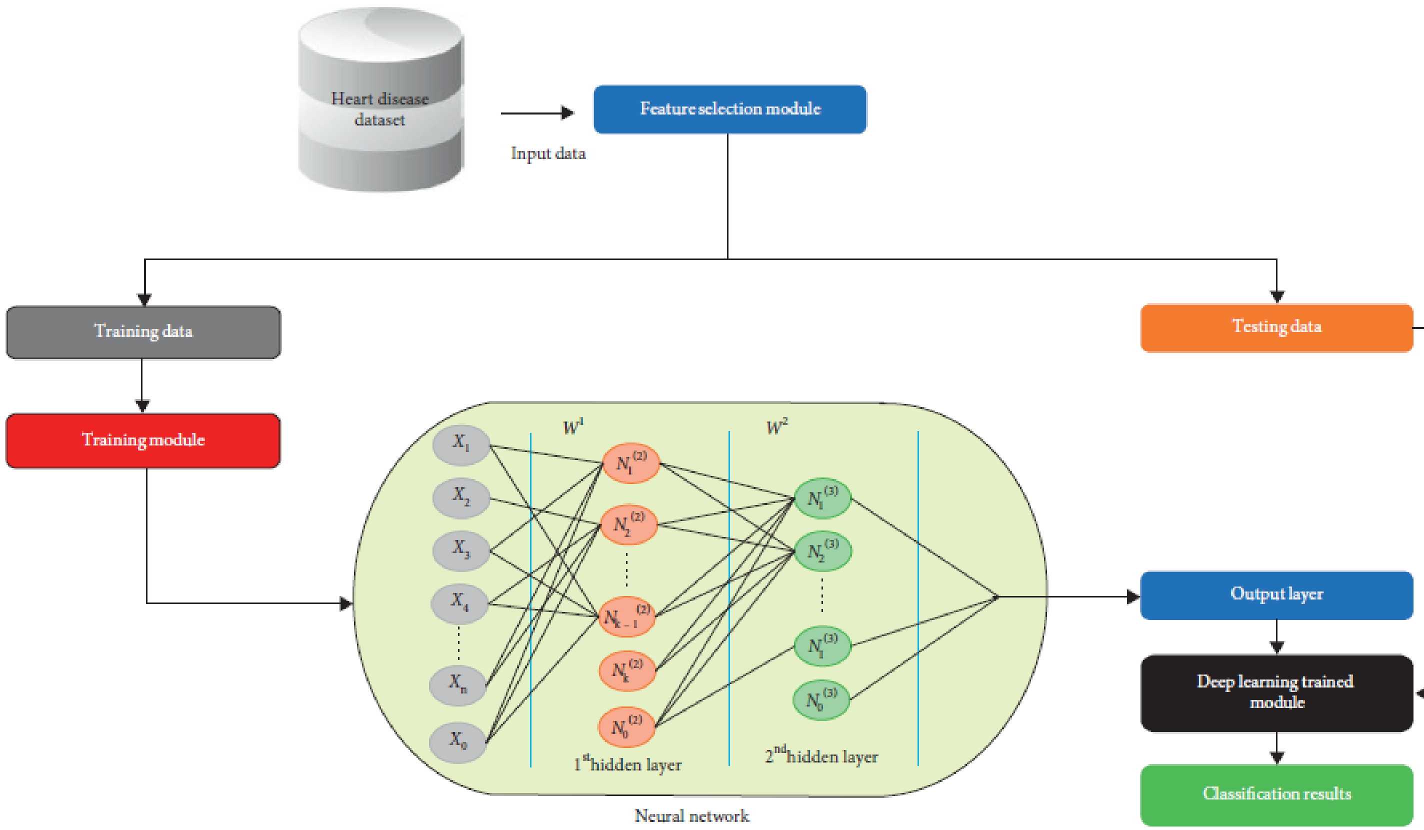


Figure16: Proposed system for Heart Risk Failure detection [JAV20].

Artificial Intelligence & Information Analysis Lab

Heart Failure Classification

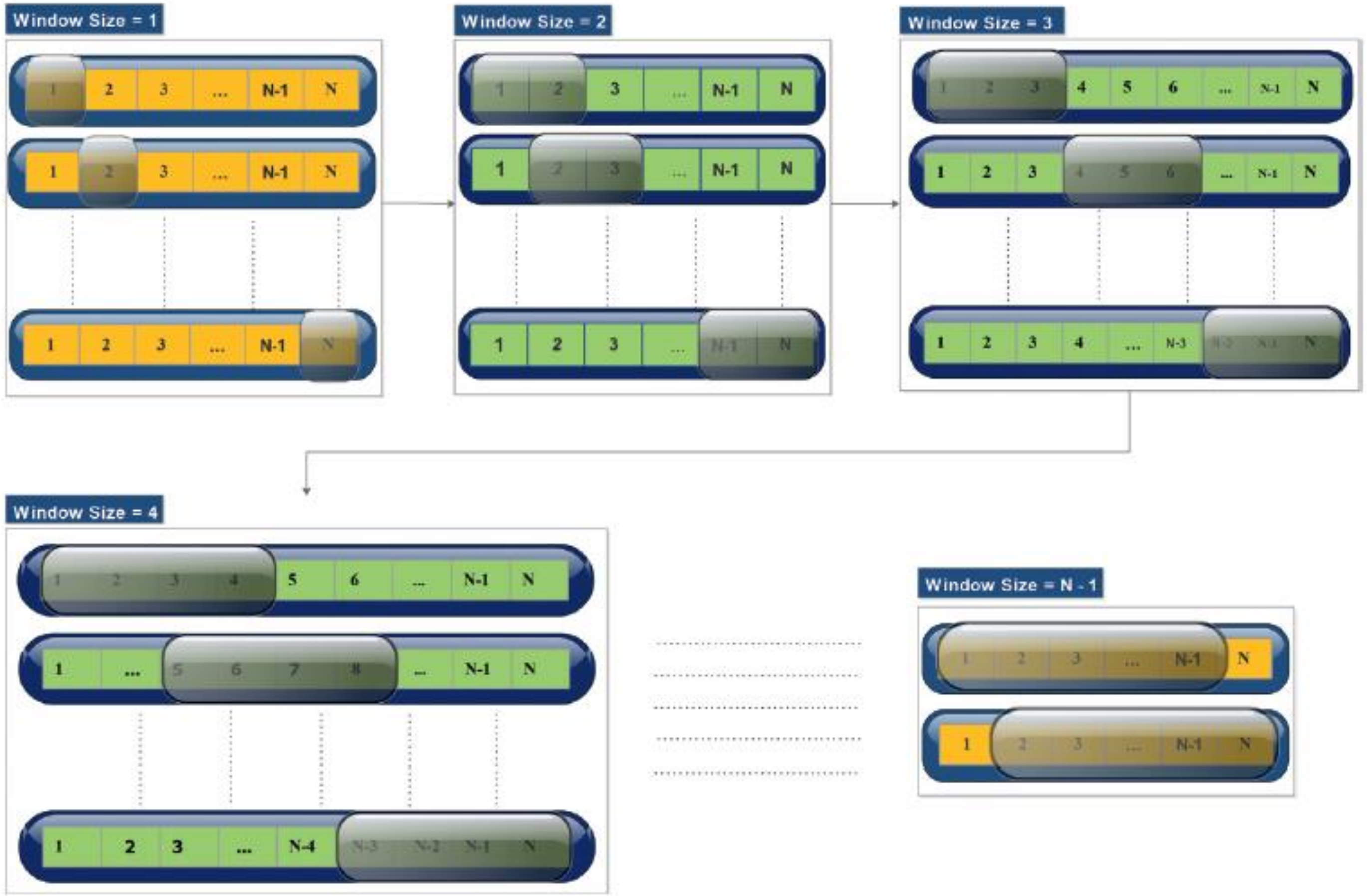


Figure 17: The feature selection algorithm [JAV20].

Heart Failure Classification

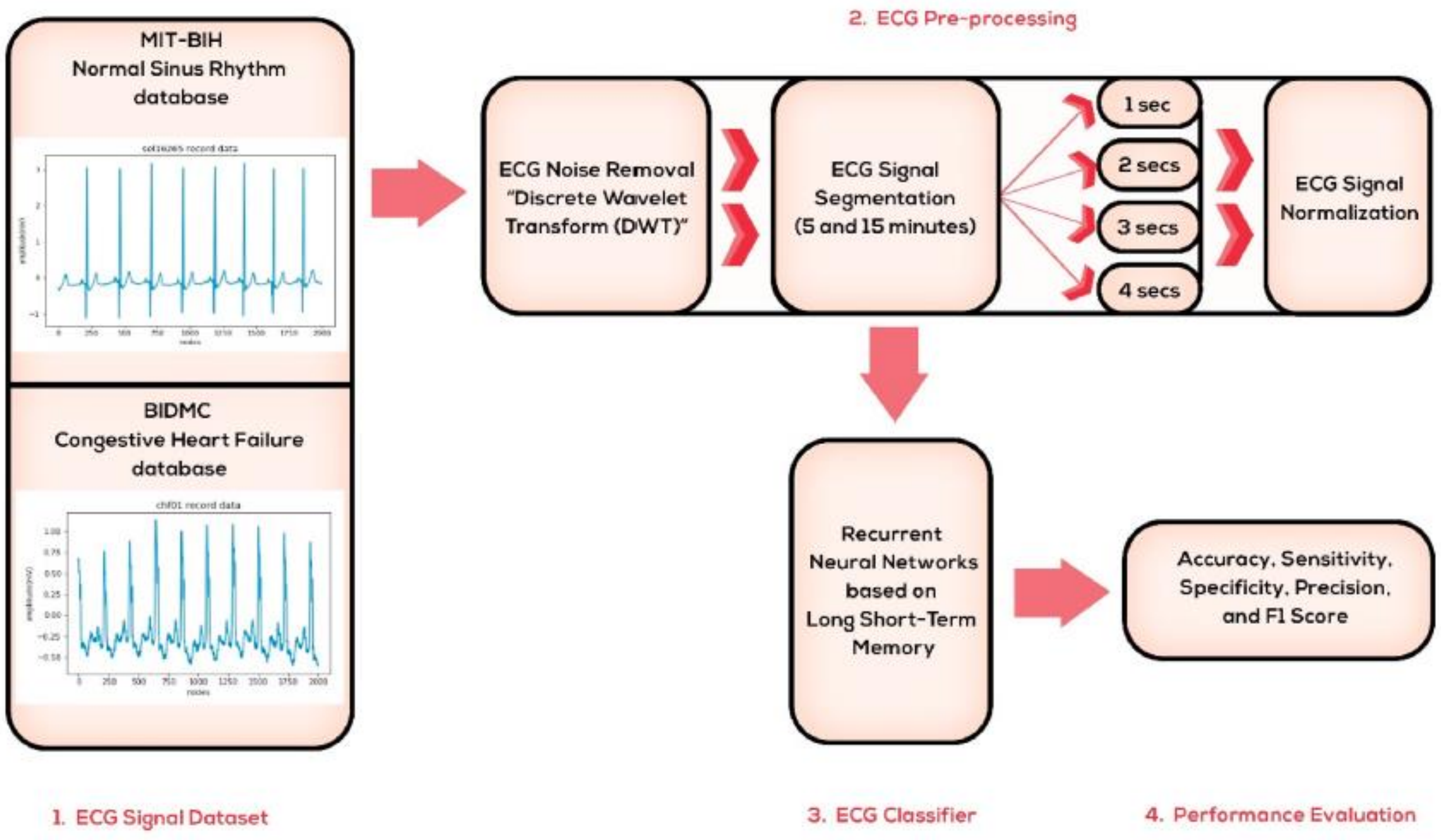


Figure 18: RNN model with Time-Step Analysis [DAR20].

Heart Failure Classification

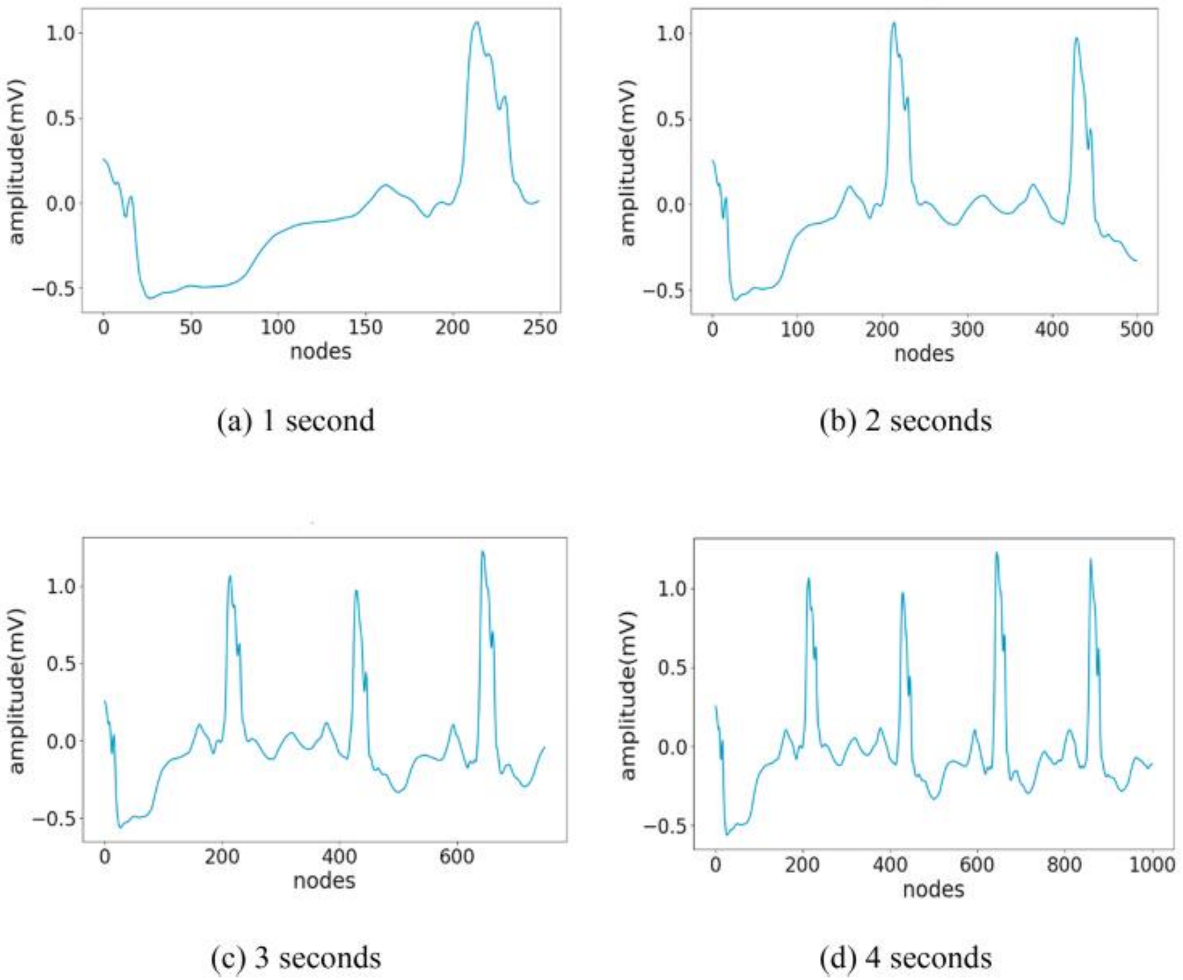


Figure 19: ECG segments from 1 to 4 seconds [DAR20].

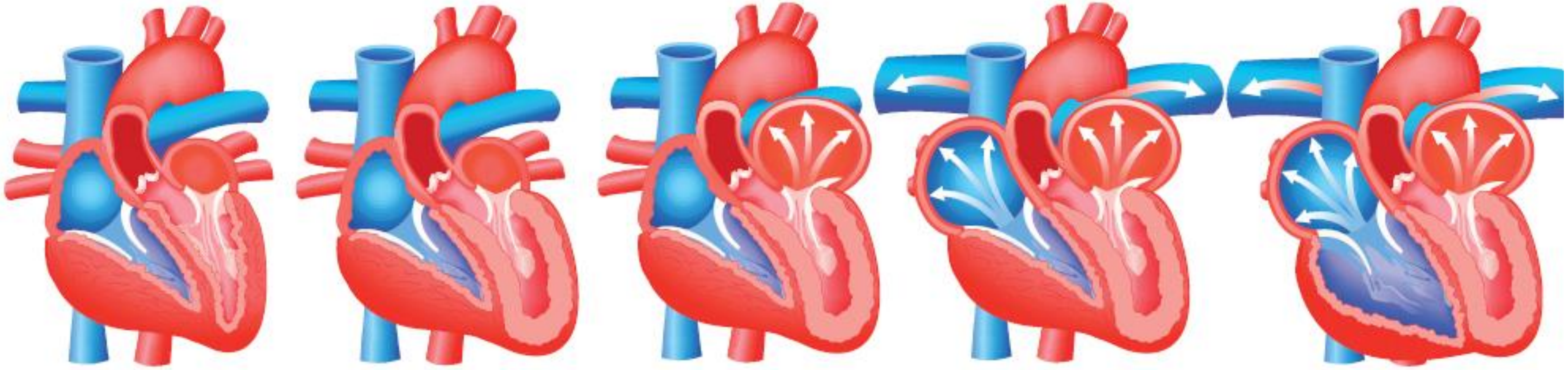
Contents

Cardiac Valve Stenosis Classification

Aortic Stenosis (AS) is the most common valvular heart disease in western world countries. In severe AS, the timing of surgery is vital for preventing subsequent cardiovascular events. AS can be detected from ECG signals is some cases.

Although AS severity can be assessed by accelerometer which measures the ***Seismocardiogram*** (SCG) signal while the gyroscope collects the ***Gyrocardiogram*** (GCG) Signal [YOA20].

Cardiac Valve Stenosis Classification



Stages/Criteria	Stage 0	Stage 1	Stage 2	Stage 3	Stage 4
	No Cardiac Damage	LV Damage	LA or Mitral Damage	Pulmonary Vasculature or Tricuspid Damage	RV Damage

Figure 20: Cardiac stratification of aortic stenosis based on cardiac damage [GPR17].

Cardiac Valve Stenosis Classification

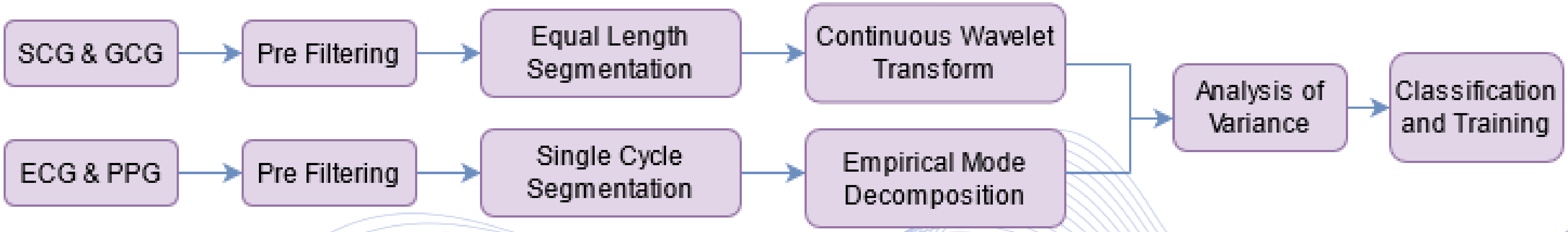


Figure 21: Schematic diagram of the cardiac signal processing.

Cardiac Valve Stenosis Classification

The used system for acquiring the data is shown in Figure 22.

This system is a small machine whine can be used in the laboratory and collect data to be processed in MATLAB [YAGT19].

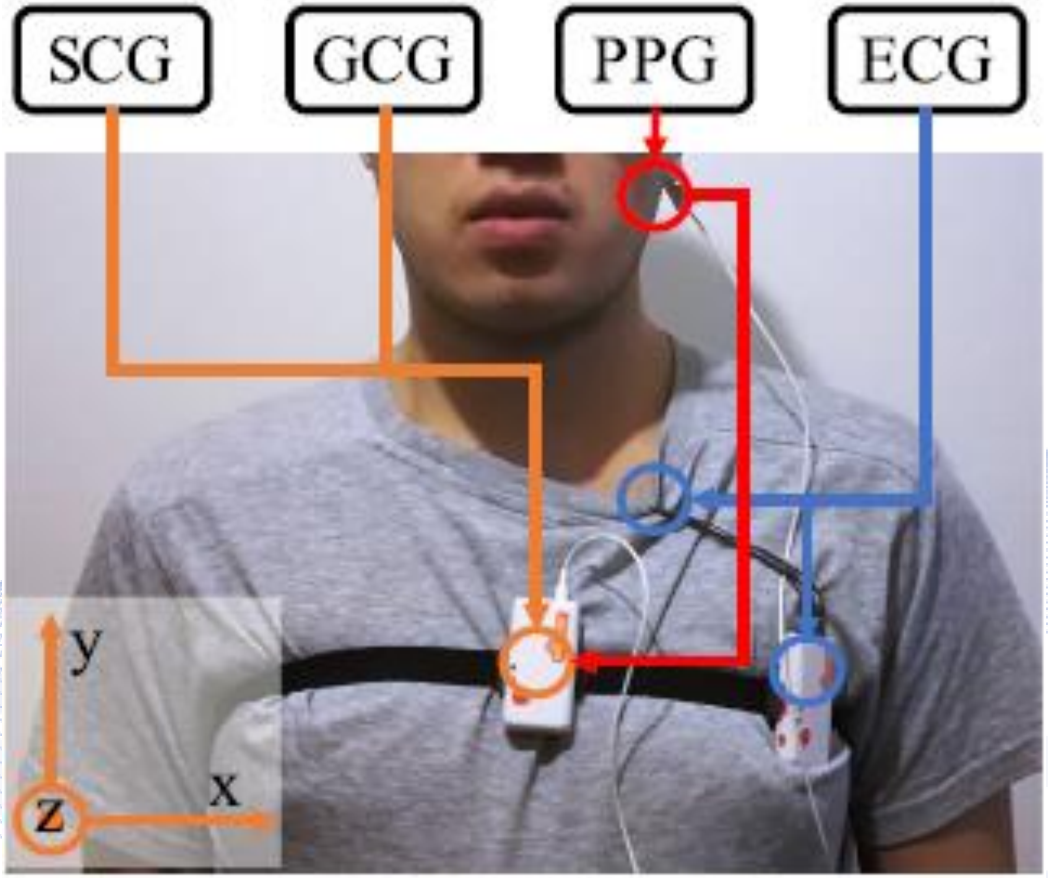


Figure 22: Hardware setup with the SCG, GCG, PPG, and ECG collectors [YAGT19].

Contents

Congenital Heart Disease Classification

Congenital Heart Diseases (CHDs) are the most com-mon congenital anomalies, becoming a major global health problem.

CHD is an abnormality of the early life which in many cases is only detected during the birth and, in the worst cases, much later [CBC21].

Congenital Heart Disease Classification

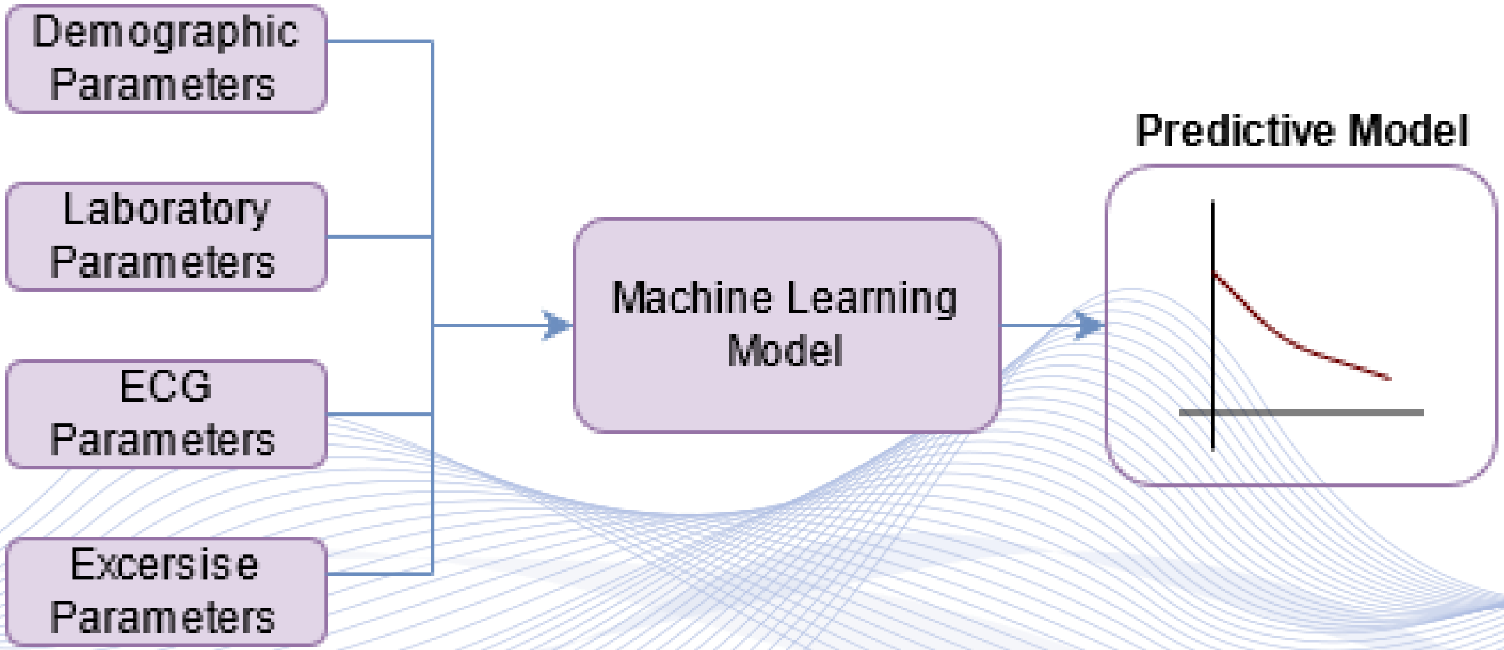


Figure 23: Block diagram of congenital disease prediction.

Contents

Coronary Artery Disease Classification

Coronary Heart Disease is one of them with research showing a rapid increase in the number of diseases.

The highest survival probability depends on early diagnosis which can be made more accurate with machine learning algorithms. Algorithm training can use different types of data including ECG [LGT21].

Coronary Artery Disease Classification

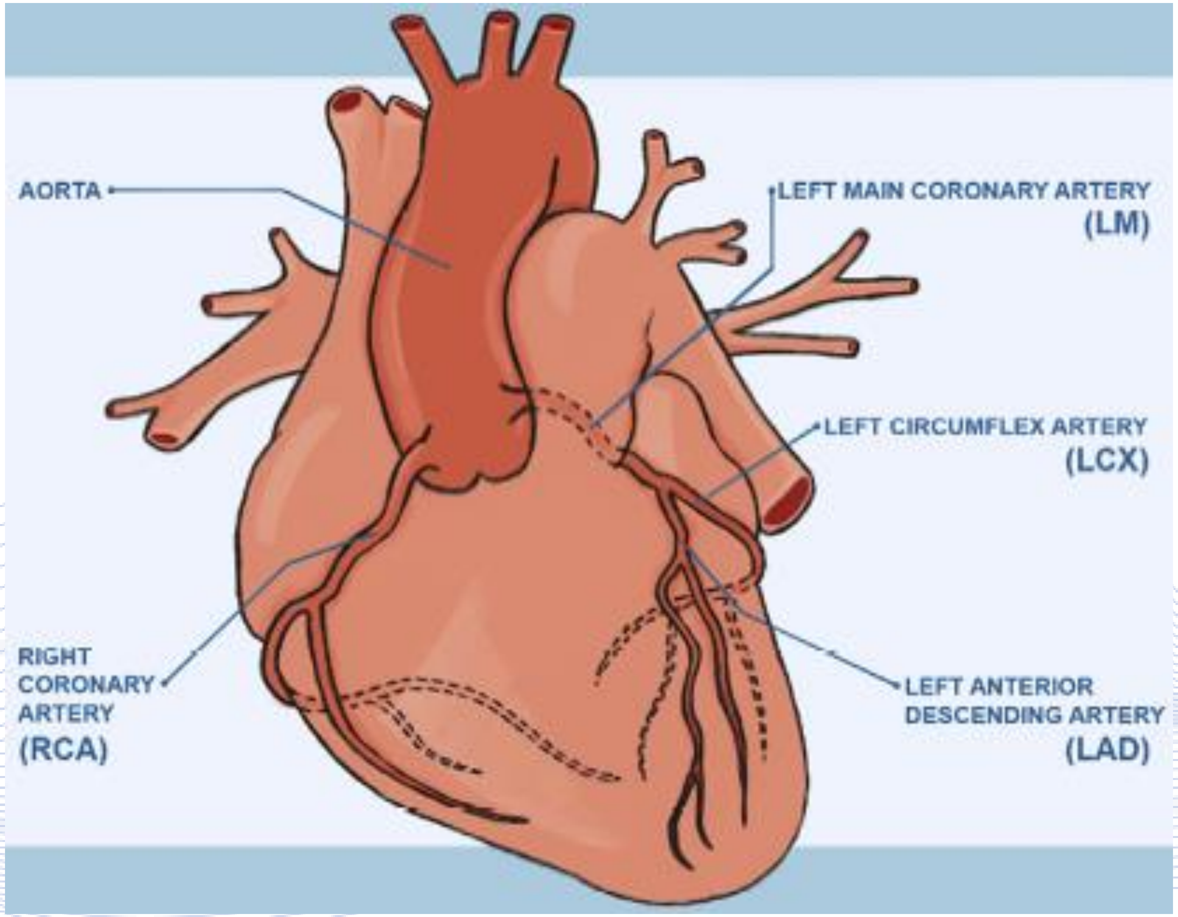


Figure 24: Coronary arteries visualization [AAR19.]

Artificial Intelligence & Information Analysis Lab

Coronary Artery Disease Classification

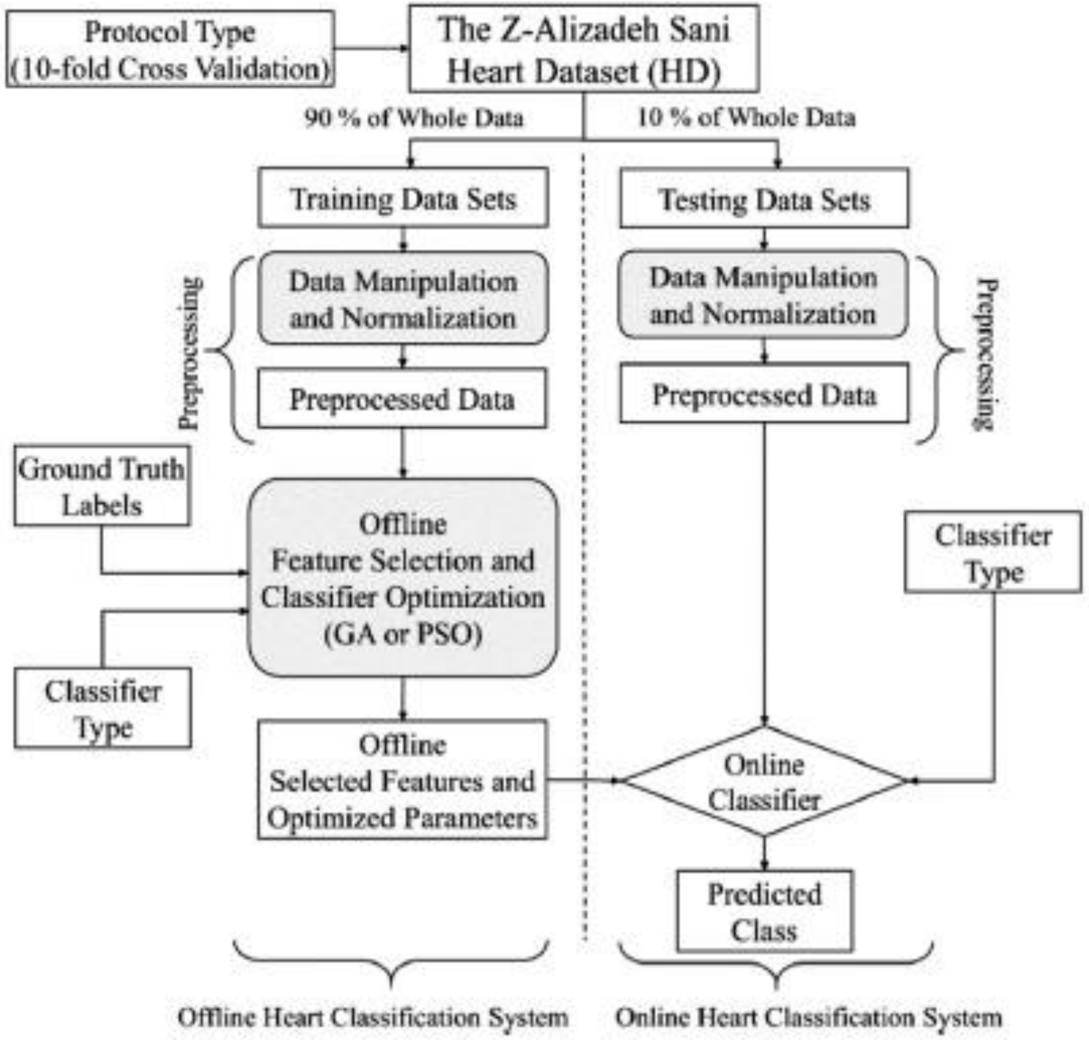


Figure 25: Early diagnosis of Coronary Artery Disease (CAD) diagram [AKA19].

Contents

Blood Pressure Hypertension Classification

Hypertension or high blood pressure is the most common heart diseases that affects many people. However, more than the half of hypertensive patients are indifferent and do not manage their blood pressure.

As hypertensive person has blood pressure readings for systolic BP, above 140 mmHg or diastolic BP, above 90 mmHg [EHK20].

Blood Pressure Hypertension Classification

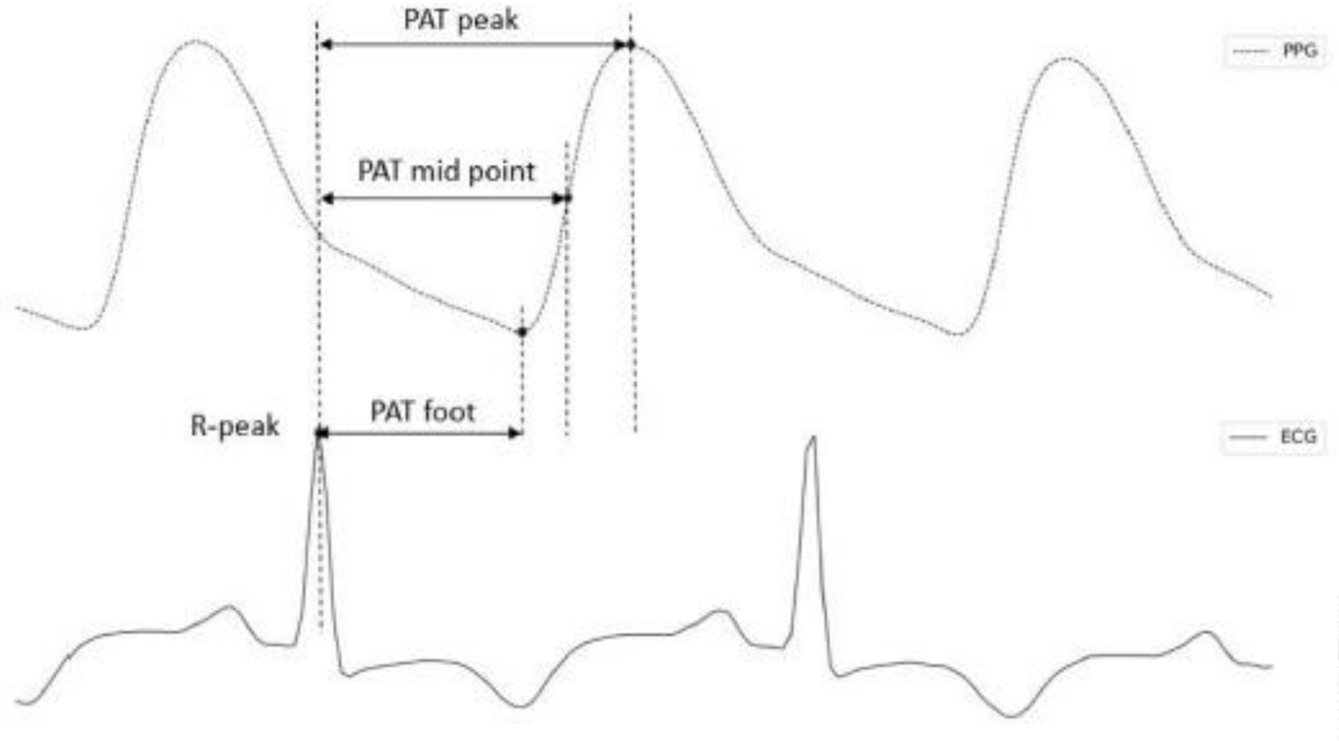


Figure 26: PAT measurement points, from R-peak of the ECG to foot, mid-point and peak of the PPG [EHK20].

Blood Pressure Hypertension Classification

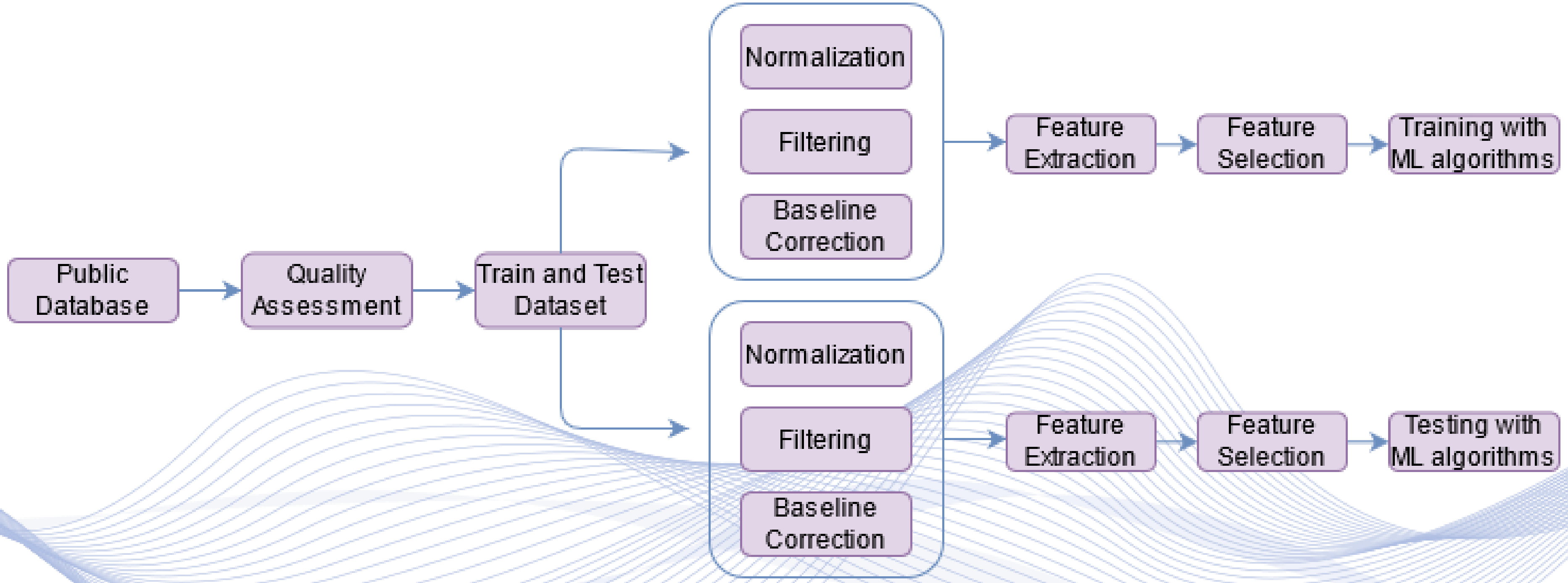


Figure 27: Block diagram of PPG and ECG data process for BP classification.

Bibliography

[OPP2013] A. Oppenheim, A. Willsky, Signals and Systems, Pearson New International, 2013.

[MIT1997] S. K. Mitra, Digital Signal Processing, McGraw-Hill, 1997.

[OPP1999] A.V. Oppenheim, Discrete-time signal processing, Pearson Education India, 1999.

[HAY2007] S. Haykin, B. Van Veen, Signals and systems, John Wiley, 2007.

[LAT2005] B. P. Lathi, Linear Systems and Signals, Oxford University Press, 2005.

[HWE2013] H. Hwei. Schaum's Outline of Signals and Systems, McGraw-Hill, 2013.

[MCC2003] J. McClellan, R. W. Schafer, and M. A. Yoder, Signal Processing, Pearson Education Prentice Hall, 2003.

Bibliography

[PHI2008] C. L. Phillips, J. M. Parr, and E. A. Riskin, Signals, Systems, and Transforms, Pearson Education, 2008.

[PRO2007] J.G. Proakis, D.G. Manolakis, Digital signal processing. PHI Publication, 2007.

[DUT2009] T. Dutoit and F. Marques, Applied Signal Processing. A MATLAB-Based Proof of Concept. New York, N.Y.: Springer, 2009

Bibliography

[PIT2000] I. Pitas, “Digital Image Processing Algorithms and Applications”, J. Wiley, 2000.

[PIT2021] I. Pitas, “Computer vision”, Createspace/Amazon, in press.

[PIT2017] I. Pitas, “Digital video processing and analysis” , China Machine Press, 2017 (in Chinese).

[PIT2013] I. Pitas, “Digital Video and Television” , Createspace/Amazon, 2013.

[NIK2000] N. Nikolaidis and I. Pitas, “3D Image Processing Algorithms”, J. Wiley, 2000.

Q & A

Thank you very much for your attention!

More material in
http://icarus.csd.auth.gr/cvml-web-lecture-series/

Contact: Prof. I. Pitas
pitas@csd.auth.gr