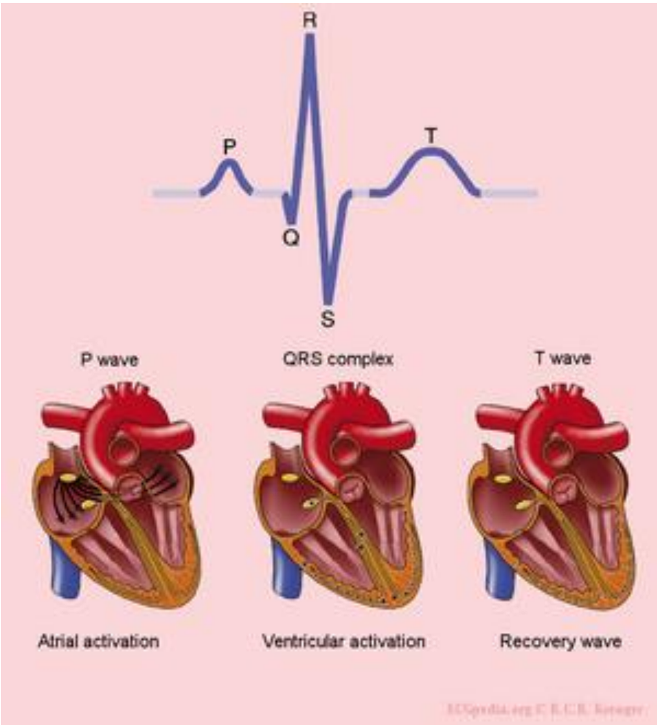
Spike Sorting:

- Correct spike detection
- Compare normal spike with the considered one
- Abnormality detection by the use of the comparison

Machine Learning:

Supervised machine learning algorithms:

The classification of the ECG signal is a very important and challenging task. It can provide substantial information about the CVDs of a patient without the involvement of a cardiologist. Only a technician is required to attach the probes, and the machine learning based solution can automatically diagnose the CVDs of the patient. This technique can immediately prioritize the patients that need urgent medical attention. In this work, the SVM and MLP supervised learning algorithms were used for classification.



What do the segments of the ECG represent?

- P-wave: Atrial contraction
- PR interval: Represents the time taken for excitation to spread from the sino-atrial (SA) node across the atrium and down to the ventricular muscle via the bundle of His.
- QRS: Ventricular contraction
- ST segment: Ventricular relaxation
- T-wave: Ventricular repolarization

Normal duration of ECG segments:

- PR interval: 0.12 – 0.2 secs (3-5 small squares)
- QRS: <0.12 secs (3 small squares)
- QTc: 0.38 – 0.42 secs

https://oxfordmedicaleducation.com/ecgs/ecg-interpretation/

Support vector machine classifier:

The SVM algorithm can be used in classification and regression problems. In SVM, data is plotted in an l- dimensional space, where l denotes the number of features. After plotting the data, classification is performed by finding a hyperplane that differentiates between different classes. The maximization of the margin optimizes the hyperplane. Then, the hyperplane, that is at a higher distance from the closest data points among other hyperplanes, is chosen.

Multi-layer perceptron classifier:

Artificial-neural-network (ANN) algorithms classify regions-of-interest using a methodology that performs functions similar to those of the human brain, such as understanding, learning, solving problems, and making decisions. The ANN architecture consists of three layers. The first layer is the input layer, and the input parameters determine the number of neurons in this layer. The last layer is the output layer, and the number of neurons in this layer represents the number of output classes. The layers between the input and output layers are called the hidden layers. MLP was used in this work, and it is a subclass of the feed-forward ANN.

https://www.nature.com/articles/s41598-021-97118-5

SE (Sensitivity): The sensitivity of a clinical test refers to the ability of the test to correctly identify those patients with the disease.

SP (Specificity): The specificity of a clinical test refers to the ability of the test to correctly identify those patients without the disease.

To calculate SE and SP, these terms should be defined as follows:

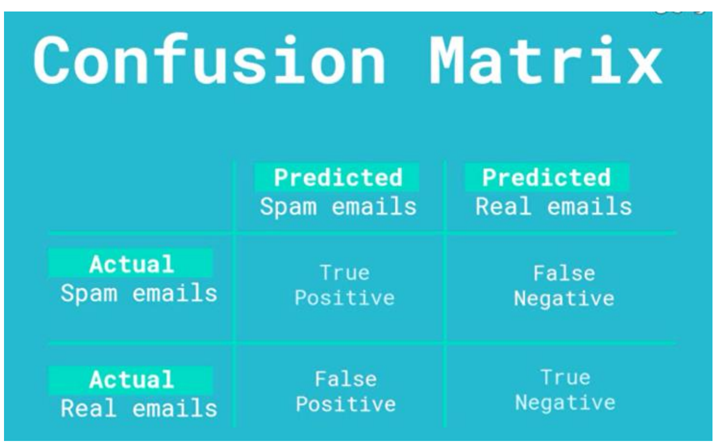
- TP True positive: the patient has the disease and the test is positive.
- FP False positive: the patient does not have the disease but the test is positive.
- TN True negative: the patient does not have the disease and the test is negative.
- FN False negative: the patient has the disease but the test is negative.

And equations of SE and SP:

$$SE = \frac{TP}{TP + FN} * 100 \quad SP = \frac{TN}{TN + FP} * 100$$

CC: Correct Classification is computed as below:

$$CC = \frac{TP + TN}{TP + TN + FP + FN} * 100$$



https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0244-x

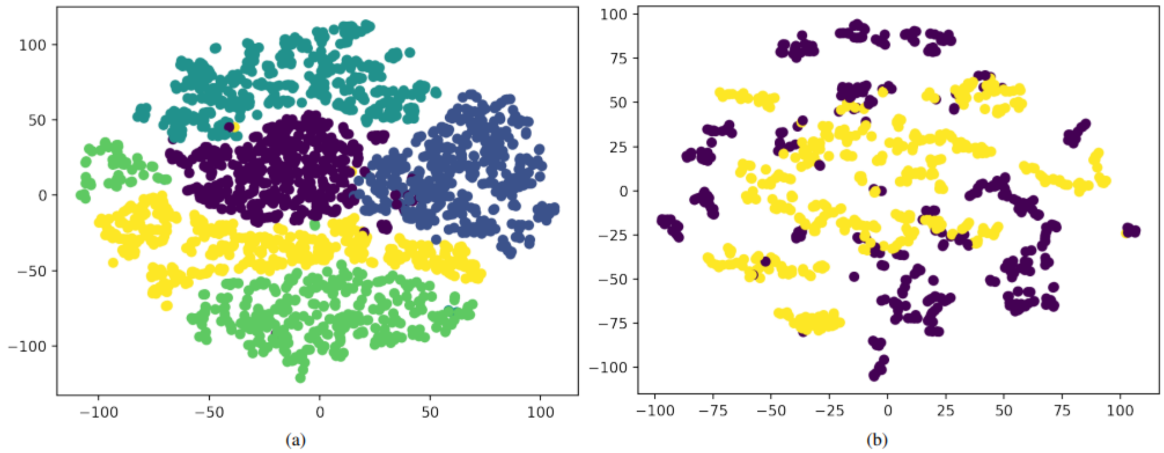


Fig. 4: t-SNE visualization of the learned representation: (a) samples from MIT-BIH for ECG beat classification (b) samples from PTB dataset for MI classification. Labels for each task are indicated with colors (best viewed in color).

The ECG provides information about both the presence and localization of MI. MI characteristics include ST-segment elevation, abnormal Q wave appearance, and T-wave inversion. These are commonly used for classification of feature vectors

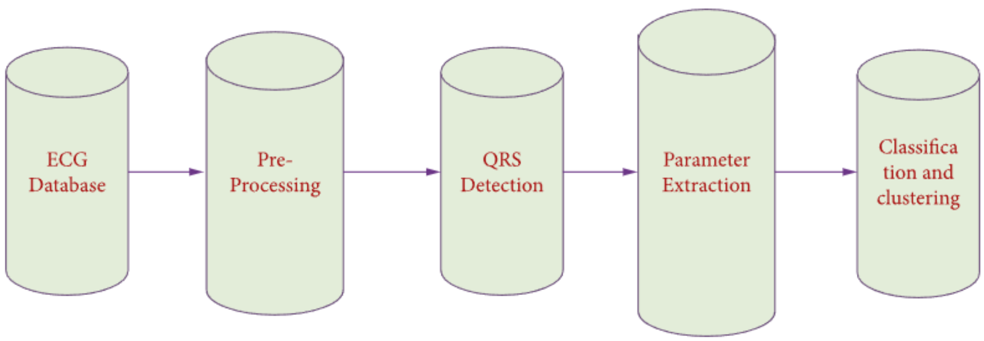
Myocardial infarction (MI), commonly known as heart attack.

From previous studies, KNN classifier has been successfully applied for the detection of MI. Principal Component Analysis (PCA) and polynomial approximation have been combined in the feature extraction phase of MI classification to improve the classification performance. Optimization algorithms can be used to improve the performance of the feature extraction phase.

Some articles used Support Vector Machine (SVM) classifier to develop a model that did not need signal preprocessing with digital filters, band-pass filter, filter-banks, and wavelet transform. Separately, He developed an unsupervised classification scheme using wavelet tensor decomposition and two- dimensional Gaussian spectral clustering. They used decision tree (DT) and KNN classifiers at the classification stage to obtain a single diagnostic method which can identify different cardiac abnormalities such as coronary artery disease, congestive heart failure and MI at early stages. In another study, Fujita et al. proposed an automated classification that can detect cardiac abnormalities as four different classes instead of just classifying them in two classes as normal and abnormal.

https://www.sciencedirect.com/science/article/abs/pii/S016786551930056X#:~:text=The%20ECG%20provides%20information%20about,6%5D%2C%20%5B7%5D https://arxiv.org/pdf/1805.00794.pdf

It is possible to classify arrhythmias with greater accuracy and a shorter SVM classifier with discrete wavelet transform (DWT) is the machine learning technique used in this study. From the MIT-BIH and BIDMC databases, seventy-three percent of the composed signals are divided into training and testing sets, with 70 : 30. DWT was used to extract a total of 190 features. Due to its flexibility to alter the window size based on frequency, DWT as a solution SVM classifier was used to classify the retrieved characteristics.



$$\text{positive predictivity } (+\text{PP}) = \frac{\text{TP}}{\text{FP} + \text{TP}},$$

$$\text{failure rate } (\text{Frr}) = \frac{\text{TP} + \text{FP}}{\text{TP}},$$

$$\text{sensitivity } (\text{ST}) = \frac{\text{TP}}{\text{FN} + \text{TP}},$$

$$\text{overall accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{TP} + \text{FN} + \text{FP}},$$

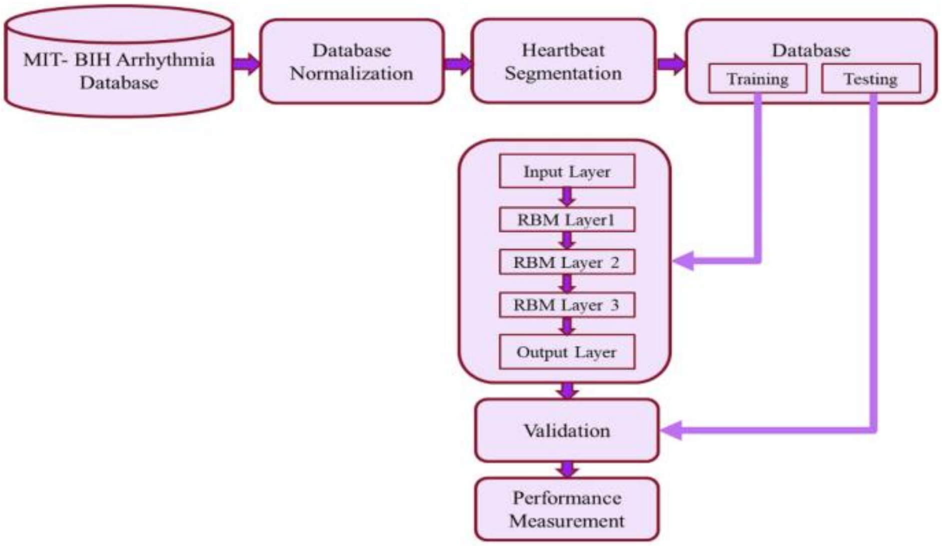
$$\text{recall} = \frac{\text{TP}}{\text{FP} + \text{TP}},$$

$$\text{precision} = \frac{\text{TP}}{\text{FN} + \text{TP}},$$

$$f_1 - \text{score} = 2.\,\frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}},$$

https://www.hindawi.com/journals/cin/2021/7677568/

Restricted Boltzmann machine (RBM) model:



https://link.springer.com/article/10.1007/s42452-021-04621-5

https://link.springer.com/article/10.1007/s00034-015-0068-7

LDA Distance Classification:

The distance in R space can be represented as the similarity between feature vectors x^p and x^q in the Euclidean metric system by:

$$d(x^p, x^q) = \sqrt{\sum_{i=1}^{R} (x_i^p - x_i^q)^2}$$

However, in the feature space, not all the features are equally weighted. So, this relation can be adjusted by adding a weight vector $w = [w_1, w_2, \dots, w_R]$.

$$d(x^p, x^q) = \sqrt{\sum_{i=1}^{R} w_i (x_i^p - x_i^q)^2}$$

The smaller the value of $d(x^p, x^q)$ the closer the distance between vector x^p and x^q. And the distance between two classes, called S_L and S_K, can be described by:

$$D(S_L, S_K) = \frac{1}{m_L \cdot m_K} \sum_{x^p \in S_L} \sum_{x^q \in S_K} d(x^p, x^q)$$

Where m_L and m_K are the numbers of feature vector in S_L and S_K.

file:///C:/Users/Asus/Desktop/Implementation_of_a_one-lead_ECG_human_identificat.pdf

Methods:

PCA

CNN

SVM

MLP

Decision Tree

KNN