ECG Signal Classification using Machine Learning Algorithms

[1]Anitha Prasad, [2]Rajshekhargouda Patil, [2]Manjunath Desai, [2]Shivakumar Yankanchi, [2]Md. Juned

[1]Assistant Professor, Dept of ECE, SJCE, India [2]Student, Dept. of ECE, SJCE, India

Abstract: Cardiovascular diseases are one of the world's leading causes of death. Electrocardiography is the most common way to monitor heart activity. Various heart disorders can be detected by analysis of electrocardiogram (ECG) signal abnormalities. Classification of ECG signals into normal and abnormal using machine learning models, through a sufficient amount of training data gives better insight into the patient condition. Machine Learning models are elegant and efficient. Some good models can be developed provided there is a wide range of reliable data available. Scaled Rolling mean is used for R peak detection. Various features of an ECG signal are fed as the input to different classification models and their performance is evaluated. MIT-BIH Arrhythmia Database is considered for evaluation.

***Index Terms*- ECG Analysis, Support Vector Machines, K – Nearest Neighbours**

I. Introduction

The etymology of electrocardiography is derived from Greek. "Electro" means electrical activity, "Cardio" means related to heart. Electrocardiogram is a medical test to ascertain cardiac abnormalities by measuring the electrical activity produced by the heart muscles, using electrodes placed over the skin.

A normal ECG signal has a characteristic shape. Any deviation from this template indicates cardiac abnormalities. A doctor may recommend an ECG for patients who may at risk of heart disease because of family history of heart disease, smoking, overweight, diabetes, high cholesterol or high blood pressure. The heart disorders that can be detected using ECG include abnormal heart rhythms, heart attack, and an enlarged heart.

ECG is the recording of the electrical property of the heartbeats and has become one of the most important tools in the diagnosis of heart diseases. Due to high mortality rate of heart diseases, early detection and precise discrimination of ECG signal is essential for the treatment of patients. Classification of ECG signals using machine learning techniques can provide substantial input to doctors to confirm the diagnosis

Classification of ECG signals is a challenging problem due to issues involved in classification process. Major issues in ECG classification are lack of standardization of ECG features, variability amongst the ECG features, individuality of the ECG patterns, non-existence of optimal classification rules for ECG classification, and variability in ECG waveforms of patients [1,2]. Applications of ECG signal classification are in detecting abnormality type and diagnosing a new patient more precisely than manually. ECG classification includes steps namely preprocessing, feature extraction, feature normalization, and classification. Here, focus is on feature extraction and classification models.

II. Background Knowledge

In a conventional 12-lead ECG, ten electrodes are placed on the patient's limbs and on the surface of the chest. The overall magnitude of the heart's electrical potential is then measured from twelve different angles ("leads") and is recorded over a period of time (usually ten seconds). In this way, the overall magnitude and direction of the heart's electrical depolarization is captured at each moment throughout the cardiac cycle. The graph of voltage versus time produced by this medical procedure is

an electrocardiogram [3]. From Fig 1, it can be noted that various zones in the cardiac cycle are identified and each is assigned a letter as follows:

- P is the atrial systole contraction pulse
- Q is a downward deflection immediately preceding the ventricular contraction
- R is the peak of the ventricular contraction
- S is the downward deflection immediately after the ventricular contraction
- T is the recovery of the ventricles

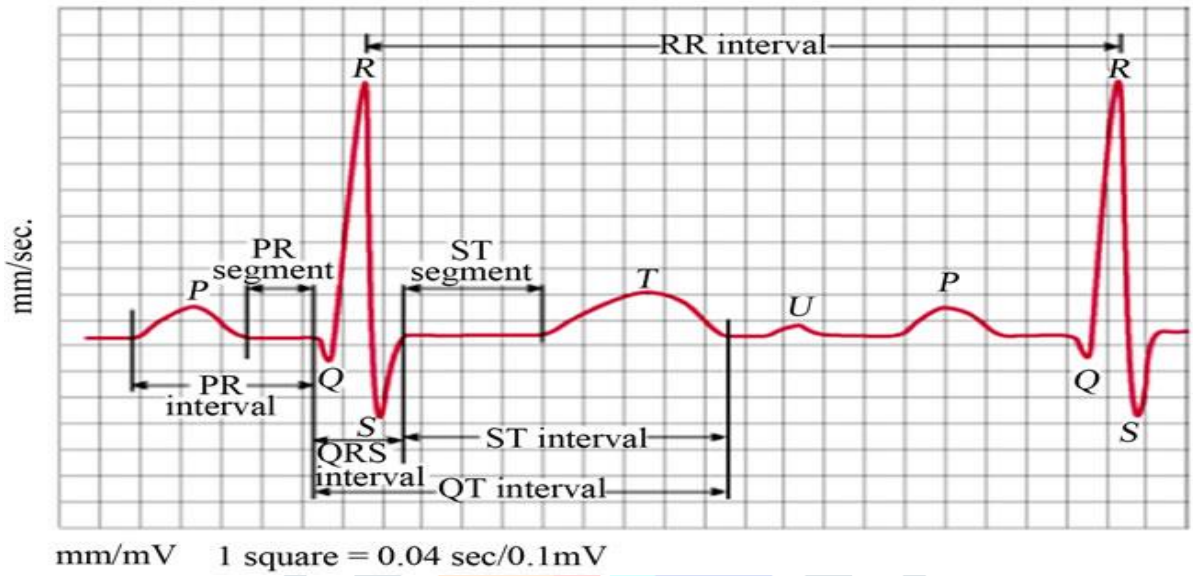

Fig 1: Typical ECG Signal

2.1. Issues in ECG Classification

The Challenges involved in ECG signal classification are enlisted below:

- Lack of standardization of ECG features
- Variability of the ECG features
- Individuality of the ECG patterns
- Non-existence of optimal classification rules for ECG classification
- Patients may have different ECG waveforms
- Beat variations in a single ECG
- Finding out the most appropriate classifier [4].

III. Algorithm for R Peak Detection using Rolling Mean:

QRS complex detection is the major part of feature extraction, many complex algorithms have been proposed previously, here basic principle of moving average is considered.

- Find the Rolling Mean, for a block of 100 datapoints.
- Find the average for the entire dataset and replace the NaN values with it.
- Scale the moving average by a suitable factor, to avoid noise being detected as R Peak.

- Consider two single dimension arrays, one is "Window" and the other is "Peaklist".
- Peaklist is used to store the X- Coordinates of all the R Peaks.
- Window is used to store the amplitude values of the data points, in region of interest.
- Region of interest is the one where, amplitude of dataset, rises above the moving average.
- All the amplitude values of data points in the region of interest are stored in windows list
- The maximum among this list is the R Peak, its X coordinate is stored in peaklist and is retrieved later to find all the RR time interval.

IV. Machine Learning Models, for ECG signal classification:

Here, Support Vector Machines (SVM), and K-Nearest Neighbours (KNN) models are considered. SVM boils down to a constraint optimization problem, while KNN classifies using the k nearest training data points at the nearest cartesian distance from test data point.

Some of the advantages of Machine Learning are:

- It easily identifies trends and patterns
- No human intervention needed
- Continuous improvement
- Handling multi- dimensional and multi-variety data

4.1. Support Vector Machines

Support vector Machines (SVM) are supervised learning algorithms, rose to fame in 90's as they performed better than neural networks in hand written character recognition. It boils down to a constrained optimization problem, where a hyper plane that optimally separates the two groups needs to be found. In other words, it be finding the widest possible street which divides the two groups. If the two groups can be separated by a straight line then, it is called linear SVM. If the two groups are not linearly separable, then a Kernel function is used to project the data points on to another space, where they can be viewed as being linearly separable [5].

Consider a vector, $\vec{w}$ perpendicular to the hyperplane, let $\vec{u}$ be the vector to the unknown datapoint.

$$\vec{w}.\vec{u} \geq c \qquad (1)$$

Dot product of $\vec{w}$ and $\vec{u}$ is the projection of u on w if its greater than some length c then it belongs to the + category. Without loss of generality, let (1) be rewritten as follows,

$$\vec{w}.\vec{u} + b \geq 0 \text{ where } c = -b \qquad (2)$$

Consider the below two constraints to find out the hyperplane,

$$\vec{w}.\vec{x_+} + b \geq 1$$

$$\vec{w}.\vec{x_-} + b \leq -1 \qquad (3)$$

They can be rewritten for mathematical convenience as follows,

Let, $y_i \ni y_i$ = **+1** for + samples

-1 for – samples

$$y_i\,(\vec{x_i}\,\vec{w} + b) \geq 1$$

$$y_i\,(\vec{x_i}\,\vec{w} + b) - 1 \geq 0 \qquad (4)$$

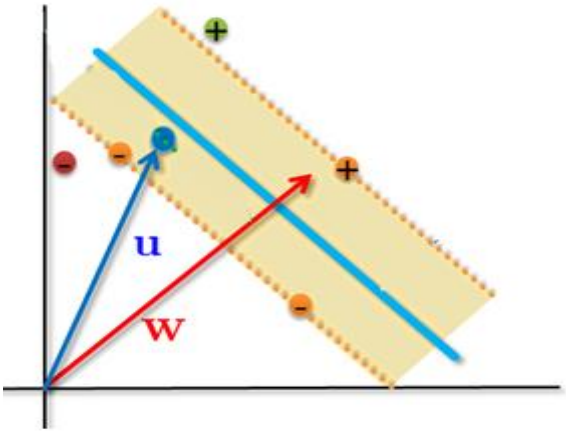

Fig 2: Representation Support Vector Machines

Let $\vec{x_+}$ and $\vec{x_-}$ be the vectors to the positive and negative datapoints of the support vector respectively. Then the width of the margin around the hyperplane is given by:

width = $(\vec{x_+} - \vec{x_-}).\frac{\vec{w}}{||w||}$

substituting for $\vec{x_+}$ and $\vec{x_-}$ from (3), we get,

width = $\frac{2}{||w||}$

The goal is to maximise the width of the margin, while abiding the constraints in (4). Hence this boils down to constraint optimisation problem where, **Max** $\frac{2}{||w||}$ => **Max** $\frac{1}{||w||}$ is to be optimised.

This is equivalent to **Min** $\frac{1}{2}\ ||w||^2$

Employing Lagrange multipliers to solve this constraint optimisation problem we get,

$$L = \frac{1}{2}\ ||\vec{w}||^2 \ - \Sigma\ \alpha_i\,[\,y_i(\vec{x_i}\,\vec{w_i} + b) - 1]$$

Solving this yield,

$$\vec{w} = \ \Sigma \alpha_i y_i \vec{x_i}$$

$$\Sigma \alpha_i y_i = 0$$

$$L = \Sigma \alpha_i - \frac{1}{2}\ \Sigma_i \Sigma_j\ \alpha_i \alpha_j y_i y_j \vec{x_i}\ \vec{x_j}$$

Here it can be noted that the optimisation depends on dot product of the pair of input vectors. This property is exploited in applying kernels and dealing with non-linear data classification.

4.2. K-Nearest Neighbours

In pattern recognition, the k-NN is a non-parametric method used for classification and regression. In both cases, the input consists of the k closest training examples in the feature space. The output depends on whether k-NN is used for classification or regression. It is widely disposable in real-life scenarios since it is non-parametric, meaning, it does not make any underlying assumptions about the distribution of data.

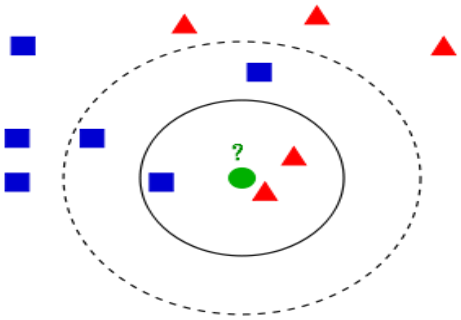

Fig 3: KNN Classification

KNN assumes that the data is in a feature space. More exactly, the data points are in a metric space.

The data can be scalars or possibly even multidimensional vectors. Since the points are in feature space, they have a notion of distance – This need not necessarily be Euclidean distance although it is the one commonly used. There is a single number "k". This number decides how many neighbours (where neighbours is defined based on the distance metric) influence the classification. This is usually an odd number if the number of classes is 2. If k=1, then the algorithm is simply called the nearest neighbour algorithm.

V. Simulation Results

All the simulations were run using Python-3.6.4, standard MIT- BIH Arrythmia database was used. This contained, subjects that were 25 men aged 32 to 89 years, and 22 women aged 23 to 89 years. The data was sampled at 360 Hz, with 11-bit resolution.

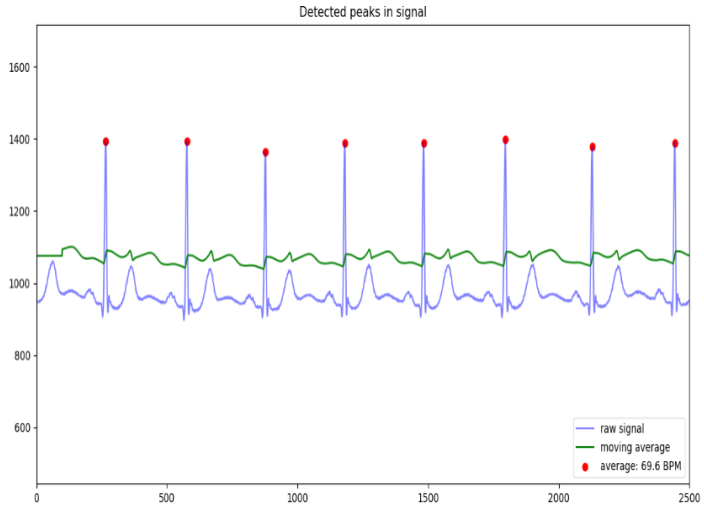

Fig 4: Detected R Peaks

Consider Fig 4, red dots in the graph denote the R peaks detected, green line is the scaled moving average, when ECG signal raises above the green line, it is regarded as the region of interest in which R peak is located.

Table 1. Comparison of the performance of Machine Learning models

Model	Accuracy
Support Vector Machines	90%
K- Nearest Neighbours	96%

From the comparison table above, it can be ascertained that KNN performs better than SVM for the dataset of MIT-BIH Arrythmia database. SVM can predict with 90% certainty about the healthy or unhealthy status of a subject, while KNN can predict with 96% certainty.

VI. Conclusion

As KNN works better than SVM, it indicates that the data set is not easily separable using the decision planes. KNN can generate a highly convoluted decision boundary as it is driven by the raw training data itself. SVM uses a highly restricted parametric approximation of the decision boundary, which is an excellent trade-off for classification performance against data storage space/ processing speed. This is a step towards a future where the diseases can be diagnosed at an earlier stage without human intervention. Machine learning models can be further improved to predict the possibility of a cardiac arrest with some tangible certainty, using the ECG data gathered. Lack of standardisation for the ECG parameters can be overcome by customizing the standardisation for each individual depending on his/her physical stature, food habits and athletic prowess etc. With the IOT enabled, miniaturisation of electronics, data gathering is of little consequence, it is the optimized analysis of collected data, that needs to be emphasised. Data driven analysis powered by machine learning models provide key insights in diagnosing diseases in remote areas where there is shortage of doctors.

References

[1]. Y. N. Singh, S. K. Singh, and A. K. Ray, "Bioelectrical signals as emerging biometrics: Issues and challenges," ISRN Signal Processing, pp. 1-13, 2012.

[2]. M. E. A. Bashir et al., "Highlighting the current issues with pride suggestions for improving the performance of real time cardiac health monitoring," Inform. Technology in Bio-and Medical Informatics, ITBAM, Springer Berlin Heidelberg, pp. 226-233, 2010.

[3]. Vaidehi Arun Dixit, Prof. P. R. Thorat, ECG Detection Using Controller, International Journal of Innovative Research in Science, Engineering and Technology, Vol. 6, Issue 11, November 2017

[4]. Shweta H.Jambukia, Vipul K. Dabhi, Harshadkumar B. Prajapati, Classification of ECG signals using Machine Learning Techniques: A Survey, International Conference on Advances in Computer Engineering and Applications (ICACEA) IMS Engineering College, Ghaziabad, India, 2015

[5]. Patrick Winston. Massachusetts Institute of Technology: MIT OpenCourseWare, 6.034 Artificial Intelligence. Fall 2010.