(IJAER) 2019, Vol. No. 17, Issue No. V, May **e-ISSN: 2231-5152, p-ISSN: 2454-1796**

HEART ARRHYTHMIA DETECTION AND CLASSIFICATION FROM ECG SIGNAL USING ARTIFICIAL NEURAL NETWORKS

Gaurav Bindal, Aditi Garg, Jahnvi Singh, and Dr. MM Tripathi

ABSTRACT

Abnormal electrical activity of the heart results in changes in the normal rhythm of heart causing cardiac arrhythmia of various types. The effects of arrhythmia can cause irreparable damage to the heart over long period of time and can be fatal at times. Early detection of arrhythmia can reduce the damage significantly and hence, the detection of arrhythmia from ECG signals is important for the medical world. In this paper we use machine learning algorithms, Neural Network, Random Forest, Logistic Regression, Boosted Trees, SVM, Naive Bayes and Nearest Neighbour. We then compare the efficiency of these algorithms in detecting arrhythmia and classifying it into types.

1. INTRODUCTION

The most common method for studying and detecting heart diseases is using Electrocardiogram (ECG) signals. The components of ECG signal are P wave, T wave and QRS complex. The most common factors which are examined while studying it are shape and correlation of P wave, T wave, and QRS complex, duration and R-R interval. The observations which are irregular cause arrhythmia of various types which has been classified into 1 to 14 classes in this paper.

Many machine learning algorithms have been deployed to detect arrhythmia from ECG signals. In this paper, we have used several algorithms and the performance of their classification is studied using parameters like confusion matrix, confidence interval, precision and accuracy.

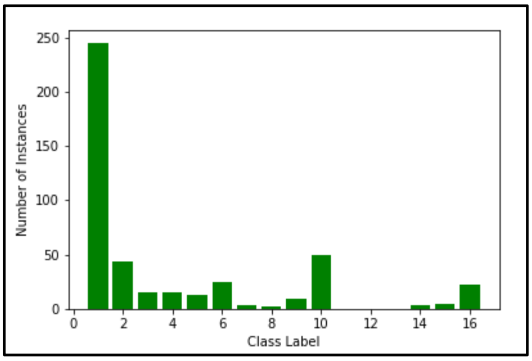
2. LITERATURE REVIEW

Taisser Mohammed et al.[1] used fuzzy logic control to detect the presence of arrhythmia. The proposed algorithm resulted in 100 percent sensitivity. Gaurav Kumar Jaiswal et al. [2] proposed an artificial neural network to classify arrhytmia and identify the normal region of classification of abnormalities in an ECG signal. Sean Franklin et al. [3] filtered the ECG signals using Butterworth Filter and then extracted parameters using Wavelet transform. A neural network was then proposed in the paper to predict cardiac arrhythmia into six classes of arrhythmia including no arrhythmia. Mohd Khalid et al. [4] proposed an artificial neural network to classify arrhythmia into five categories. The MLP neural network was trained with parameters extracted from RR intervals. MaedehKianiSarkaleh et al. [5] classified ECG signals into three categories of Arrhythmia using discrete wavelet transform and implemented it on a Multi- layer Perceptron system. DeshmukhRohan et al. [6] classified abnormal beats using neural network toolbox available on

MATLAB, relying on the Empirical Mode Decomposition methodology. Rahat Abbas et al. [7] assessed the application of direct and iterative method in a neural network model in order to compare the predicted Electrocardiogram signals. Nitish V. Thakor et al. [8]proposed adaptive filter structures for noise cancellation by minimizing the mean-squared error and arrhytmia detection. An adaptive recurrent filter structure is applied to detect P-waves, premature ventricular complexes, and recognition of conduction block, atrial fibrillation, and paced rhythm.

3. DATASET AND PREPROCESSING

We use Arrhythmia dataset found on UC-Irvine archive of Machine Learning repository. The aim is to distinguish between the presence and absence of arrhythmia and also, identify the type of it. The datasetconsists of 452 records and 279 attributes in it. Every record is classified into 1 to 16 classes, where a class label of 1 indicates normal ECG patterns while a class label between 2 to 16 indicates “abnormal ECG patterns” or arrhythmia of varying types. The dataset’s class distribution is shown in the figure. The arrhythmia dataset is run against neural network and various other classification algorithms on jupyter notebook. We selected features from the dataset in groups of 5 and plotted the accuracy of neural network in each case. Based on the observation, it was found that the dataset has been arranged in such a way that the first 100 features played a major role in determining results and hence, the first 100 features were selected and tested against various algorithms.



4. MODELS

We compare our neural network model to several commonly used machine learning classification algorithms:

1. Random Forest
2. Logistic Regression
3. Boosted Trees
4. Support Vector Machine
5. Naive Bayes
6. K Nearest Neighbour

4.1 Neural Network

Neural networks are information processing systems that try to replicate the way of learning and functioning of a human brain. The distinguishing feature of this system is its structure. A large number of highly interconnected information processing paradigms called neurons, work collectively to solve a specific problem. Neural networks like humans, learn by example.

For both binary and multi-class classification, a multi-layer feedforward neural network was employed and trained using backpropagation to minimise the cost function. A large portion of this research was dedicated in analysing the structure of neural network. Different combinations of depth, learning rate, hidden layer size and regularization parameters were explored with the motive of improving accuracy. It was also found that the accuracy improved after preprocessing the dataset using feature extraction as explained in section 4. Based on the results, we use 100 features for our final model and the following hyperparameters :

1. Binary classification
Layer 1: Input Layer, 14 neurons, activation – Relu
Layer 2: Hidden Layer, 8 neurons, activation – Relu
Layer 3: Output Layer, 2 neurons, activation – Sigmoid

2. Multi-class classification
Layer 1: Input Layer, 60 neurons, activation – Relu
Layer 2: Hidden Layer, 50 neurons, activation – Relu
Layer 3: Hidden Layer, 40 neurons, activation – Relu
Layer 4: Hidden Layer, 30 neurons, activation – Relu
Layer 5: Output Layer, 17 neurons, activation – Sigmoid

4.2 Random Forest

Random forest constructs ennumber of decision trees when the data is trained; the output is either the mode of the classes or mean prediction of individual trees. It can be used for many techniques such as classification, regression or prediction. It minimizes the problems related to high variance or high bias by evaluating a balance between the two boundaries using the statistical method of averaging. In case of classification of arrhythmia into the sixteen categories, a margin function is formed using the algorithm. The function calculates the extent by which the average numbers of votes of all the other classes present as a dependent variable are exceed by that of the actual class. In this paper, all the input vectors were pushed down the trees whereas every tree corresponded to one of the type of arrhythmia present in patients. Thus, every tree will vote or favor one of the classes. The class with the highest number of votes will be chosen as the outcome of that particular data entry in the given algorithm.

4.3 Logistic Regression

It aims to minimize the cost function using stochastic gradient descent. The function can take values between 0 and 1.

4.4 Boosted Trees

Boosted trees is a technique used for regression and classification problems which produces a model in the form of an ensemble of weak predicting decision trees. Boosting creates predictors sequentially rather than simultaneously and hence the predictors learn from previous mistakes. The observations have an unequal probability of appearing in different models. Like other boosting methods, it creates the model in a gradual fashion, and it generalizes them by allowing optimization of a random differentiable loss function.

Using a training set { (x(1),y(1)),.....(x(n),y(n)) } of known values of x and corresponding values of y, the goal is to find an approximation function F'(x) to a function F(x) that minimizes the expected value of some specified loss function L(y, F(x)) :

F' = argminExy [L(y, F(x))]

4.5 Support Vector Machine

Support Vector Machine algorithms are used to analyze data and identify patterns. They are supervised learning algorithms which can be used for multivariate purposes, majorly classification and regression. It classifies the data of every new patient based on the set of training data which was fed to the algorithm making it a non-probabilistic linear classifier. The algorithm denotes the training data as points in space such that all the different classes are at a significant distance from each other. The new introduced data points are then plotted on the same space.

4.6 Naïve Bayes

Naive Bayes is a simple technique for constructing classifiers or models that assign class labels to problem instances, represented as vectors of feature values, where the class labels are drawn from some finite set. It works on the Bayes probability theorem

p(Ck|x)= p(Ck) p(x/Ck) / p(x)

It assumes that every pair of feature being classified is independent of every other.

4.7K – nearest neighbour

It is a non-parametric algorithm based on feature similarity. The algorithm classifies any new object based on the majority vote of its neighbours with the case being assigned to the class most common amongst its K nearest neighbours measured by a distance function. If the value of k is kept at unity, then the case is simply assigned to the class of its nearest neighbour.

5. EXPERIMENT

The dataset was separated into training and testing dataset. Out of 452 records, 316 records were randomly chosen for training and rest 136 were held out for testing and validation. Preprocessing, training and testing were done in Python.

Since few records were incomplete, instead of discarding them, imputation was done by using the feature's mean value and it was observed that this improved model's performance.

6. RESULT

Results show that the neural network performs better than other methods on the binary classification task and is outperformed only by Random Forest method. For multi class Arrhythmia classification, our neural network performs well and outperforms Support Vector Machine, Boosted trees and nearest classifier.

Table 1: Comparison of Binary Classification Reports

Algorithms	Precision	Confidence Interval	F1-score	Accuracy
Boosted Trees	0.76	0.76	0.76	76.47%
Logistic Regression	0.75	0.81	0.76	78.54%
Naïve Baye's	0.78	0.69	0.64	69.12%
Nearest Classifier	0.70	0.67	0.63	66.91%
Neural Network	0.82	0.79	0.78	79.41%
Random Forest	0.80	0.80	0.80	80.15%
Support Vector Machine	0.74	0.74	0.74	74.26%

Table 2: Comparison of Multiple Classification Reports

Algorithms	Precision	Confidence Interval	F1-score	Accuracy
Boosted Trees	0.56	0.63	0.59	63.24%
Logistic Regression	0.58	0.68	0.60	68.38%
Naïve Baye's	0.66	0.66	0.63	66.18%
Nearest Classifier	0.50	0.58	0.47	58.09%
Neural Network	0.46	0.65	0.54	65.44%
Random Forest	0.67	0.74	0.68	73.53%
Support Vector Machine	0.51	0.57	0.45	57.35%

7. CONCLUSION

A feedforward neural network was implemented to obtain the results. Recurrent and autoregressive neural networks can also be used because of the time-series nature of ECG signals. Additionally, a combination of algorithms may prove to be better in terms of accuracy. In this research work, Arrhythmia dataset available on UCI Machine Learning Repository was used. Accuracy was plotted against specific features selected in groups of five and based on that, 100 features were selected. Missing value imputation was used to account for irregularities due to missing data. It was found out that neural networks performs the best out of the seven algorithms in case of binary classification. In case of multi-class classification, it is outperformed by random forest. As in case of any learning algorithm, a larger and more accurate dataset collected from patients can prove to be beneficial. Practical implementation of this research work can be the construction of smart health sensors for real time cardiac monitoring.