Myocardial Infarction Detection using Genetic Algorithm and Convolution Neural Network

Shreya Kumari[1], Rajesh Kumar Sharma2
[1]MTech Scholar, [2]Assistant Professor,
Department of CSE, RKDF University, Bhopal, M.P, India

ABSTRACT

The anomalies in observed data from an ECG equipment are used to diagnose heart illness. Because these signals aren't visible, they can't be recognized directly. As a result, a suitable diagnostics model is intended for real anomaly detection. As a result, the benefits of machine learning, such as convolutional neural networks (CNNs), were employed in this study to identify anomalies. The results were also compared to other machine learning approaches that and showed up its efficiency.

Keywords: Myocardial infarction, Electrocardiogram, Machine Learning, Convolutional Neural Networks (CNN).

1. Introduction

A growing body of evidence suggests that machine learning can perform as well as or better than humans in crucial healthcare activities like disease diagnosis. Algorithms are already surpassing radiologists in terms of detecting dangerous tumors and advising researchers on how to build cohorts for expensive clinical trials. For decades, researchers in both electrical engineering and computer science have been working on a set of algorithms that can quickly uncover statistical links between occurrences and results. These are mathematical tools that are capable of "learning" correlations between input and output and of creating an approximate solution to define their linkage; they integrate with established information and could either locate set of inputs or analytical relationships between the input and output data [1][2].

Machine learning is a family of techniques for mathematical and statistical simulation that utilizes a variation of strategies to naturally study and enhance target state prediction without explicit programming. Various methods such as the artificial neural networks, random forests and bayesian networks, all employ different assumptions and mathematical paradigms to determine how data is ingested and how knowledge takes place within an algorithm. Although many users of these techniques distinguish them from popularly defined machine learning methods, regression analyses, such as linear and logistic regression, are also considered machine learning methods (e.g., Bayesian Networks [BNs], random forests, etc.) [3][4]. Large corporations use the term “machine learning,” while “ML” is more commonly used for marketing purposes. In the vast majority of circumstances, the term “machine learning” is more applicable. Machine learning algorithms can be divided into subcategories based on how they understand inferences from information. Unsupervised learning, supervised learning, and reinforcement learning are the three subcategories [5][6].

2. Literature Review

Nowadays health is an important factor and a major consideration for an individual especially after this pandemic of COVID-19. Doctors and specialist are trying their level best for medication and treatment of various diseases. Various clinical tests have already been implemented and are very successful in the field of health care for the diagnosis. Few of the popular diagnostic procedures in healthcare include MRI, CT scan, X-ray, Ultrasound, Laproscopy, ECG, NCS, and NVS and so on. Among these ECGare the most common means of diagnosis for various abnormalities. Scholars have done numerous studies for the detection of various abnormalities in heart and lungs. Few of the works done in the same field are mentioned as follows:

Sahu et.al [7] proposed solution uses cloud computing technologies to create an efficient internet of things (IoT) enabled real-time ECG monitoring system. The paper describes a cloud-based method for remote CVD monitoring. ECG data is collected and sent to an Amazon web service (AWS) S3 bucket over a mobile gateway. To enable data visualization, quick response, and long-lived connections to devices and users, the AWS cloud employs HTTP and MQTT servers. Bluetooth low energy (BLE 4.0) is a low-power data transmission technology used between a device and a mobile gateway.Filtering methods are used in the proposed system to eliminate distractions, external noise, and motion artefacts. It analyses ECG signals to detect factors including heartbeat, PQRST wave, and QRS complex intervals, as well as breathing rate. The suggested system prototype has been thoroughly tested and certified for dependable ECG monitoring in real time.

A unique ensemble classification algorithm based on ECG morphological data is presented by **Yang et al [8]** as a new strategy for accurate identification of cardiac ventricular and atrial abnormalities.

Tong et al. [9] proposed to construct a unique multi-instance neural network (MINN) model on MIT-BIH arrhythmia database and the CMUH database. The AUC and sensitivity of MINN's greatest performance in detecting ECG segments are 0.9922 and 0.9809, respectively, while its best performance in detecting irregular heartbeats is 0.9473. The results of the experiments suggest that our system is capable of providing categorization and location messages, which could be valuable in the analysis of long-term ECG records.

Venkateshan et.al [10] distinguished between normal and atypical participants, ECG signal preprocessing and support vector machine-based arrhythmic beat categorization are used. A delayed error normalised LMS adaptive filter is utilised in ECG signal preprocessing to achieve high speed and low latency design with fewer computing parts. White noise removal is the major emphasis of the signal processing approach, which was created for remote healthcare systems. For HRV feature extraction, a discrete wavelet transform is used to the preprocessed data, and machine learning algorithms are utilised for arrhythmic beat classification.

Venkatesan et.al [11] distinguished between normal and pathological participants, ECG signal preprocessing and KNN-based arrhythmic beat categorization are used. Based on the LMS Adaptive filters are employed in the preprocessing of ECG signals; however they take longer because of the extensive critical path, processing takes a long time. To deal with this issue, a new adaptive filter with Error with a delay To achieve fast speed and low latency, the normalised LMS algorithm is used. This design uses the pipelining notion in the error to obtain a low-power design. Path of feedback the preprocessed signal is used to perform R-peak detection, which is done using wavelets. On the HRV feature extracted signal, a KNN classifier is used to classify arrhythmic beats. The suggested DWT with KNN classifier has a classification accuracy of 97.5 percent, which is higher than that of conventional machine learning techniques.

3. Proposed methodology

The proposed study works in three major steps: (1) Acquisition of ECG signal (2) Pre-processing on the signal so that accurate and better performance of proposed methodology can be achieved (3) Last and the final step is the bypass of filtered ECG data through CNN architecture which ultimately results in the detection of abnormalities if any. Fig. 1 represents the flow of steps.

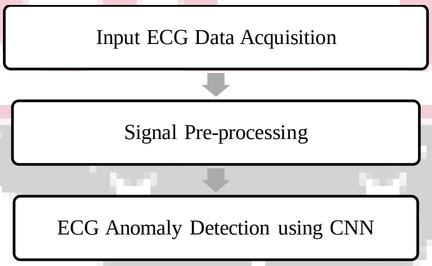

Figure 1: Steps of Proposed Work

Figure 2 depicts the suggested technique, which is explained in further depth in the following sections.

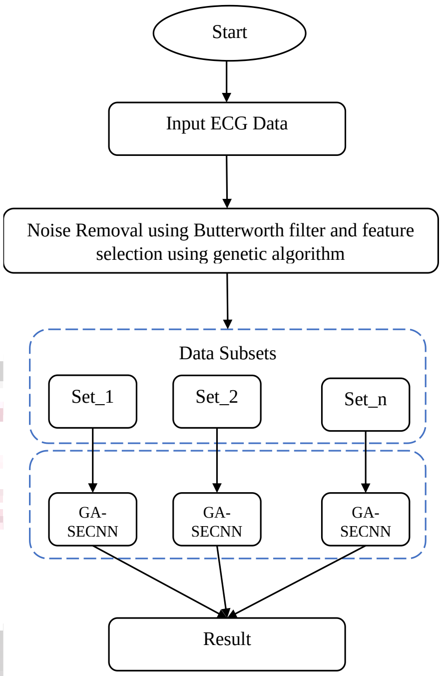

Figure 2: Flowchart of Work

ECG Data Acquisition: The Massachusetts Institute of Technology's arrhythmia data collection was utilised to investigate arrhythmia in this study**.**

Signal Pre-processing: The distorted signal or any unwanted sound considered as noise and it is present in all type of signal, it will often contaminate the signal, resulting in fewer effective feature extraction and classification. Preprocessing is the process of removing noise from an signal obtained from ECG. Filtering a signal removes any noise or distortion that may be present. The Butterworth filter technique is used for filtration.

Feature Selection using Genetic Algorithm: The genetic algorithm is a stochastic method of optimizing functions based on the mechanics of natural genetics and biological evolution. In nature, organisms' genes tend to evolve over successive generations to better adapt to the environment. The genetic algorithm is a heuristic optimization method that is inspired by the processes of natural evolution. Genetic algorithms work with a population of individuals to produce better approximations. The algorithm creates a new population with each generation by selecting individuals based on their fitness level in the problem area. These individuals are then recombined with each other using operators borrowed from natural genetics. The offspring can also mutate. The figure below shows a state diagram for the trait selection process with the genetic algorithm.

This process leads to the development of populations more suited to their environment than the individuals who produced them.

In our case, each individual in the population represents a neural network.

The genes here are binary values and represent inclusion or non-compliance with certain model features. The number of genes is the total number of input variables in the data set.

For each question the number of individuals or the size of the population must be chosen. It is usually set to 10N, where N is the number of functions.

ECG Anomaly Detection using SECNN: In this section, a model of CNN based anomaly detection is presented. In this work, a voting strategy is used to reach the final decision. The feature sets are individually analyzed by CNN block and their aggregate decision is taken as final decision. The comprehensive system is created using a convolutional neural network (CNN). Residual learning is the suggested technique throughout this study. The following is the residual learning loss function:

$$C(\Theta) = \frac{1}{2Ni}\sum_{i=1}^{N} ||R(y_i;\Theta), y_i - x_i||_F^2 \qquad (1)$$

Where,

R= Residual learning function

N_i= number of required training data input.

y_i = Output target
x_i = Input data
Ө re= Convolution neural network Input values
Proposed CNN Design: The CNN layers in this study are as follows:

- The suggested CNN layers are as follows: Convolutional and Exponential Linear Module levels are combined throughout the suggested Convolution network (SELU). The layer contains 64 feature maps with a convolution filter of size 33.
- The SELU activation function is utilised for non-linear functions.

4. Result and Discussions

This section comprises a mathematical modelling and explanation of the proposed method for ECG anomaly classification, as well as a simulation of the suggested algorithm's effectiveness.
The suggested algorithm method is modeled in the given arrangement in assessing its effectiveness and performance:
Software used:-
64-bit Windows OS
Python-3.8
Hardware Requirement
1Tera-Byte Hard Disk
Intel Core i5-3210M CPU @ 2.50GHz
4Gega-Byte RAM
The result is analyzed using parameters such as accuracy, precision, recall and f_measure. These are illustrated as below:
Accuracy: This equation is given by 2 and represents the recognition accuracy in % for every given test input to the entire training data:

$$\text{Accuracy} = (\text{TP} + \text{TN})/\text{TP} + \text{TN} + \text{FP} + \text{FN} \quad (2)$$

Precision: It's calculated as the ratio of properly identified movements for each class to the total number of classes, and it's given by:

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP}) \quad (3)$$

Recall: Recall is the ratio of true positive (TP) to them sum of false negative and true positive.

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN}) \quad (4)$$

F-Measure: The F-Measure is the result of combining precision and recall rates. This factor was used to determine the total performance of the system in terms of accurate outcomes (i.e., without taking into account incorrect recognition observations), and it was calculated as follows:

$$\text{F-Measure} = 2 * [(\text{precision} * \text{recall}) / (\text{precision} + \text{recall})] \quad (5)$$

Where,
The condition becomes true positive (TP) when actual and predicted values of anomaly are matched i.e., 1.
The condition becomes true negative (TN) when actual and predicted values of non- anomaly (normal) are matched i.e. 0.
The condition becomes false negative (FN) when actual and predicted values of anomaly and non-anomaly(normal) do not match and confusion creates i.e., 1 and 0. This situation is undecided.
The condition becomes false positive (FP) when actual is anomaly and predicted values is non-anomaly(normal) i.e., 1 and 0 respectively. This situation is undecided.

The result analysis on testsets are presented in figure 3. In this figure it was observed that proposed model achieved 98.43% of accuracy, 95.30% of precision, 88.34% of recall and 91.45% of f_measure. This result was compared with some other existing models. The comparative analysis was presented in fig 4. The comparative analysis shows the efficacy of proposed model.

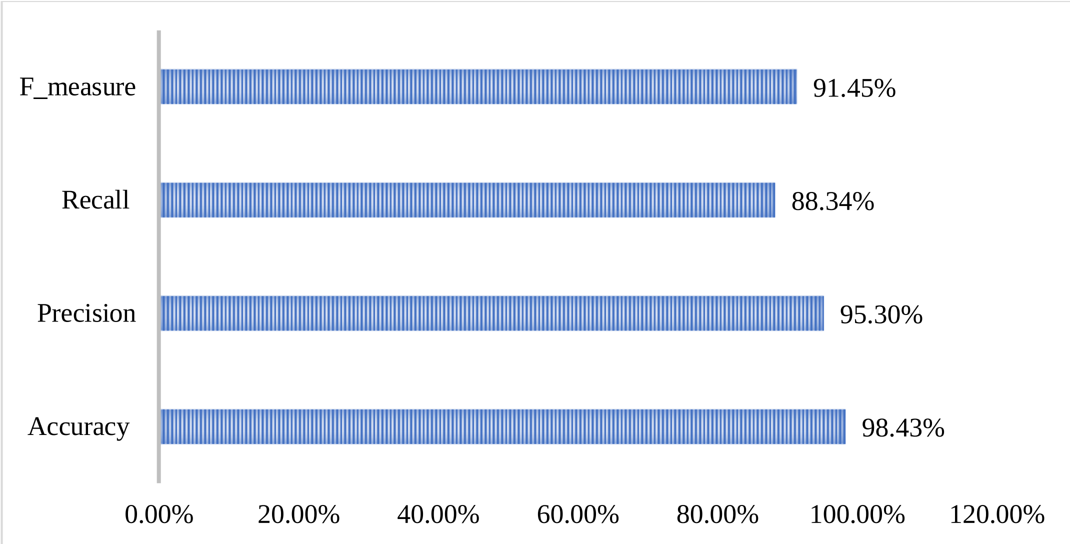

Figure 3: Performance Evaluation of Proposed Model

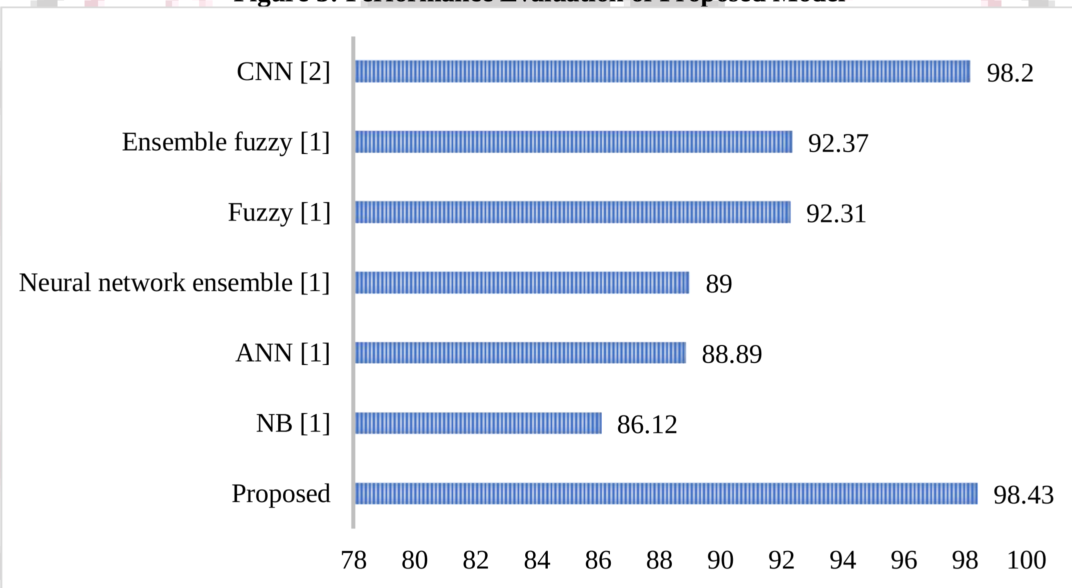

Figure 4: Comparative Performance Evaluation

5. Conclusion

In medical terms, heart attack is termed as myocardial infarction that is considered to be one of the most dangerous disease for human being. For proper diagnosis of such coronary disease, Electrocardiogram(ECG) device is used to identify and control the heart disease. Heart disease is identified by analyzing the abnormalities in recorded signals from ECG device. As these signals cannot be identified visually. So, there is requirement of proper diagnosis system for actual abnormality identification. Therefore, this research work used the advantages of machine learning such as convolutional neural network (CNN) for actual identification of abnormalities. In this work, an GA-SECNN is designed with dense convolution-SELU layers that identified these signals with very low error rate and high accuracy. The result was also compared with other machine learning approach that was presented in other existing work and showed improvement over them. This shows the efficacy of the model that can be further applied on other types of disease analysis as a future work.

References

[1] J. P. Li, A. U. Haq, S. U. Din, J. Khan, A. Khan and A. Saboor, "Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare," in IEEE Access, vol. 8, pp. 107562-107582, 2020, doi: 10.1109/ACCESS.2020.3001149.

[2] M. A. Khan, "An IoT Framework for Heart Disease Prediction Based on MDCNN Classifier," in IEEE Access, vol. 8, pp. 34717-34727, 2020, doi: 10.1109/ACCESS.2020.2974687.

[3] O. R. Shishvan, D. S. Zois, and T. Soyata, "Machine Intelligence in Healthcare and Medical Cyber Physical Systems: A Survey," *IEEE Access*, vol. 6, pp. 46419–46494, Aug. 2018, doi: 10.1109/ACCESS.2018.2866049.
[4] J. Amann, A. Blasimme, E. Vayena, D. Frey, and V. I. Madai, "Explainability for artificial intelligence in healthcare: a multidisciplinary perspective," BMC Med. Informatics Decis. Mak. 2020 201, vol. 20, no. 1, pp. 1–9, Nov. 2020, doi: 10.1186/S12911-020-01332-6.
[5] G. Rong, A. Mendez, E. Bou Assi, B. Zhao, and M. Sawan, "Artificial Intelligence in Healthcare: Review and Prediction Case Studies," Engineering, vol. 6, no. 3, pp. 291–301, Mar. 2020, doi: 10.1016/J.ENG.2019.08.015.
[6] M. A. Shareef, V. Kumar, Y. K. Dwivedi, U. Kumar, M. S. Akram, and R. Raman, "A new health care system enabled by machine intelligence: Elderly people's trust or losing self control," Technol. Forecast. Soc. Change, vol. 162, p. 120334, Jan. 2021, doi: 10.1016/J.TECHFORE.2020.120334.
[7] S. ML, A. M, A. MK, and A. A, "IoT-enabled cloud-based real-time remote ECG monitoring system," J. Med. Eng. Technol., vol. 45, no. 6, pp. 473–485, 2021, doi: 10.1080/03091902.2021.1921870.
[8] H. Yang and Z. Wei, "A novel approach for heart ventricular and atrial abnormalities detection via an ensemble classification algorithm based on ecg morphological features," IEEE Access, vol. 9, pp. 54757–54774, 2021, doi: 10.1109/ACCESS.2021.3071273.
[9] Y. Tong et al., "Locating abnormal heartbeats in ECG segments based on deep weakly supervised learning," Biomed. Signal Process. Control, vol. 68, p. 102674, Jul. 2021, doi: 10.1016/J.BSPC.2021.102674.
[10] C. Venkatesan, P. Karthigaikumar, A. Paul, S. Satheeskumaran, and R. Kumar, "ECG Signal Preprocessing and SVM Classifier-Based Abnormality Detection in Remote Healthcare Applications," IEEE Access, vol. 6, pp. 9767–9773, Jan. 2018, doi: 10.1109/ACCESS.2018.2794346.
[11] VenkatesanC., KarthigaikumarP., and VaratharajanR., "A novel LMS algorithm for ECG signal preprocessing and KNN classifier based abnormality detection," Multimed. Tools Appl., vol. 77, no. 8, pp. 10365–10374, Apr. 2018, doi: 10.1007/S11042-018-5762-6.