IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 4, April 2021

DOI 10.17148/IJARCCE.2021.10405

A Model For Electrocardiogram (ECG) Classification Using Convolutional Neural Network (CNN)

J. Palimote[1], F. Gaage[2], K.T. Igulu[3]

Department of Computer Science, Kenule Beeson Saro-wiwa Polytechnic, Bori, River State, Nigeria[1,2,3]

Abstract: Electrocardiogram (ECG) is an intermittent sign, which mirrors the movement of the heart. From ECG a great deal data is gotten for typical and obsessive physiology of heart. The ECG signal is non-fixed in nature and extremely hard to dissect. Clinical perception takes long time and the sign is non-fixed. This paper presents a convolutional neural network algorithm for electrocardiogram signal classification. This system uses a ECG dataset which was downloaded from kaggle.com. The dataset signals were preprocessed to make sure that each segment conforms to a heartbeat. This dataset was read into the jupyter notebook by using pandas.read_csv function. The dataset was made into two, which are the training data and the testing data. After successful reading of the data signal from directory and solving of the in balance problem by means of data augmentation, the proposed model was trained using a convolutional neural network(CNN) algorithm with a total hidden layers of six, hiden size of 128, batch_size of 96, and number of epoch to be 10. After successful training, we had an accuracy of 99% at an epoch level of 10.

Keywords: Deep Learning, Electrocardiogram, ConvolutionalNeural Network, Heart Beat, Signal

1. INTRODUCTION

Diseases influencing heart have gotten normal on daily basis. Diseases influencing heart is expanding as a result of present day feeble way of life increment in ailments like diabetes, hypertension and tobacco smoking. Heart can be influenced because of various conditions. Electrocardiogram (ECG) is one of the basic, effectively accessible, more affordable, effectively feasible, non-obtrusive examinations accessible at all spot incorporating provincial territories with insignificant framework. Specialists at the far off spots with fundamental clinical information may not be all around prepared to decipher an ECG. Appropriately diagnosing cardiovascular (heart related) condition at the most punctual is of most extreme significance. During these ailments, time matters the most. Consistently delay in treating them, trail to more harm in heart muscle bringing about antagonistic result. Consequently, exact distinguishing proof and early diagnosing these heart infections is vital. Anticipating patients into low and high danger is vital. Okay patients can be dealt with locally at a similar clinic with insignificant framework itself. High-hazard patients require early reference to cardiovascular reference clinics. ECG affirms or associates the finding with myocardial localized necrosis, arrhythmias and different conditions. When affirmed, they are treated with prescriptions or methodology (medical procedures) contingent upon the patient subtype and the offices accessible [1].

Electrocardiogram (ECG) is an intermittent sign, which mirrors the movement of the heart. From ECG a great deal data is gotten for typical and obsessive physiology of heart. The ECG signal is non-fixed in nature and extremely hard to dissect. Clinical perception takes long time and the sign is non-fixed. In this way, computer based method is utilized in ECG investigation. The rule of ECG signal is electrical movement. This is sent through the body and pick up on the skin. At last, the ECG machine records the action utilizing terminals and show graphically. The ECG signal comprises of dreary complex waveform with a recurrence estimation of 1 Hz. Each cardiovascular cycle comprises of three waves, for example, P wave QRS wave and T wave. These waves are delivered because of the capacity of atria and ventricular pieces of the heart. Cardiovascular illness is one of the significant reasons for death in the western reality where in excess of 16 million kick the bucket every year [2].

An Electro cardio realistic sign is recorded on a long timescale to recognize discontinuously happening aggravations in the heart musicality. To diminish the cardiovascular illnesses the way of life changes, for example, decreased cholesterol admission and customary exercise are obligatory. Early discovery is an essential advance in the cardiovascular sickness. To gauge the electrical sign from various pieces of heart two sorts of terminals are utilized. They are appendage anodes and chest terminals. The P wave is created by atrial depolarization, the length is under 120 ms, and this can be considered as low recurrence wave. ORS complex wave has span of 70-110 ms and T wave is delivered by the repolarization of ventricles having length more noteworthy than 300 ms [3].

Electrocardiogram (ECG) accounts are frequently debased by a lot of commotion and ancient rarities that can be inside the recurrence band of helpful heart information and can show with comparable morphologies to the ECG waveform itself [4]. Not exclusively does the presence of commotion and antiquity meddle with the right acknowledgment of QRS,

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 4, April 2021

DOI 10.17148/IJARCCE.2021.10405

P and T floods of the ECG, yet in addition builds the pace of bogus cautions for cardiovascular screens [5]. Some prior distributed works have depicted ECG signal quality evaluation and clamor assessment techniques, by and large as a feature of an overall way to deal with ECG examination. Grouchy and Mark utilized the remaining subsequent to projecting a QRS complex onto the initial five head parts (PCs) of an outfit of QRS edifices, or the Karhunen–Loève change [6]. Standard commotion estimation strategies for ECG, which can be utilized as individual sign quality files (SQIs), were audited by Clifford et al., including the root mean square (RMS) power in the isoelectric area, the proportion of R-top to clamor abundancy in the isoelectric district, the Crest factor or top to-RMS proportion and the proportion between in-band (5–40 Hz) and out-of-band otherworldly force.

2. RELATED WORKS

Roopa and Harish [1] Review the different Machine Learning Techniques for diagnosing Myocardial Infarction (respiratory failure), separate Arrhythmias (heart beat variety), Hypertrophy (increment thickness of the heart muscle) and Enlargement of Heart. The methods and algorithms that were broadly utilized are counterfeit neural networks, support vector machines, fuzzy logic, decision trees, rough set hypothesis, genetic and hybrid algorithm. Further, they likewise present different Machine-Learning approaches and contrast various techniques and results utilized to examine the ECG. The current techniques are investigated dependent on subjective and subjective boundaries viz., motivation behind the work, calculations received and results got.

S. Celin and K. Vasanth, [7] proposed a technique that will be used to characterize the ECG signal by utilizing classification method. In the first place, the Input signal is preprocessed by utilizing separating technique like low pass, high pass and spread worth channel to eliminate the high recurrence commotion. Margarine worth channel is to eliminate the overabundance commotion in the sign. In the wake of preprocessing top focuses are identified by utilizing top recognition calculation and concentrate the highlights for the sign are extricated utilizing factual boundaries. At last, extricated highlights are characterized by utilizing SVM, Adaboost, ANN and Naïve Bayes classifier to group the ECG signal data set into typical or strange ECG signal. Test result shows that the precision of the SVM, Adaboost, ANN and Naïve Bayes classifier is 87.5%, 93%, 94 and 99.7%. Contrasted with other classifier gullible bayes classifier precision is high.

Qiao et.al. [8] Outlines a five-level ECG signal quality grouping calculation. A sum of 13 sign quality measurements were gotten from fragments of ECG waveforms, which were marked by specialists. A help vector machine (SVM) was prepared to play out the characterization and tried on a reenacted dataset and was approved utilizing information from the MIT-BIH arrhythmia data set (MITDB). The reenacted preparing and test datasets were made by choosing clean portions of the ECG in the 2011 PhysioNet/Computing in Cardiology Challenge information base, and adding three sorts of genuine ECG commotion at various signalto-clamor proportion (SNR) levels from the MIT-BIH Noise Stress Test Database (NSTDB). The MITDB was re-explained for five degrees of sign quality. Various mixes of the 13 measurements were prepared and tried on the reenacted datasets and the best blend that created the most noteworthy order exactness was chosen and approved on the MITDB. Execution was surveyed utilizing characterization exactness (Ac), and a solitary class cover precision (OAc), which expects that an individual sort arranged into a contiguous class is worthy. An exactness of 80.26% and an OAc of 98.60% on the test set were gotten by choosing 10 measurements while 57.26% (Ac) and 94.23% (OAc) were the numbers for the inconspicuous MITDB approval information without retraining. By playing out the fivefold cross approval, a precision of 88.07 ± 0.32% and OAc of 99.34 ± 0.07% were acquired on the approval overlap of MITDB.

Lyon et.al, [9] recognized an unmistakable Hypertrophic Cardiomyopathy (HCM) aggregates dependent on ECG computational examination, and portray contrasts in clinical danger components and anatomical contrasts utilizing heart attractive reverberation (CMR) imaging. Four HCM aggregates were recognized dependent on QRS morphology and T wave biomarkers utilizing an AI approach. Patients with essential TWI not auxiliary to QRS anomalies had an expanded HCM Risk-SCD score and existing together septal and apical hypertrophy. Their outcomes, and the idea of the fundamental cycles caught by the ECG, recommend that computational ECG phenotyping can possibly be a novel and free factor for hazard definition.

Keshman et.al, [10] we apply AI techniques and calculations to distinguish pressure from electro-cardiogram (ECG) signals in car drivers under various degrees of natural pressure brought about by driving conditions. We locate that high pressure can be effectively distinguished from ECG flags alone with 100% exactness. These outcomes are gotten by taking a gander at the variety in the ECG signal between conditions of rest and high pressure in an individual and not just on estimations dependent on pressure state alone.

Ledezema et.al, [11] utilized ten Tusscher-Panfilov 2006 model and the O'Hara-Rudy model for human myocytes to make two populaces of models that were in concordance with information got from sound people (control populaces) and included between subject inconstancy. The impacts of ischemia were thusly remembered for the control populaces to mimic the impacts of gentle and extreme ischemic occasions on single cells, full ischemic links of cells and links of cells with different sizes of ischemic locales. Activity potential and pseudo-ECG biomarkers were estimated to evaluate how the development of ischemia could be measured. At last, two neural organization classifiers were prepared to distinguish the various levels of ischemia utilizing the pseudo-ECG biomarkers. The control populaces showed activity

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 4, April 2021

DOI 10.17148/IJARCCE.2021.10405

potential and pseudo-ECG biomarkers inside the physiological reaches and the patterns in the biomarkers regularly distinguished in ischemic patients were seen in the ischemic populaces. From one perspective, between subject inconstancy in the ischemic pseudo-ECGs blocked the recognition and grouping of early ischemic occasions utilizing any single biomarker. Then again, the neural organizations showed affectability and positive prescient incentive above 95%. Furthermore, the neural organizations uncovered that the biomarkers that were pertinent for the recognition of ischemia were not quite the same as those significant for its order.

Simjanoska et.al, [12] built up a technique for BP assessment utilizing just electrocardiogram (ECG) signals. Strategies: Raw ECG information are separated and fragmented, and, following this, an intricacy investigation is performed for highlight extraction. At that point, an AI technique is applied, consolidating a stacking-based grouping module and a relapse module for building systolic BP (SBP), diastolic BP (DBP), and mean blood vessel pressure (MAP) prescient models. Moreover, the strategy permits a likelihood dispersion based alignment to adjust the models to a specific client. Results: Using ECG chronicles from 51 unique subjects, 3129 30-s ECG portions are built, and seven highlights are removed. Utilizing a train-approval test assessment, the strategy accomplishes a mean supreme blunder (MAE) of 8.64 mmHg for SBP, 18.20 mmHg for DBP, and 13.52 mmHg for the MAP expectation. At the point when models are aligned, the MAE diminishes to 7.72 mmHg for SBP, 9.45 mmHg for DBP and 8.13 mmHg for MAP.

3. METHODOLOGY DESIGN

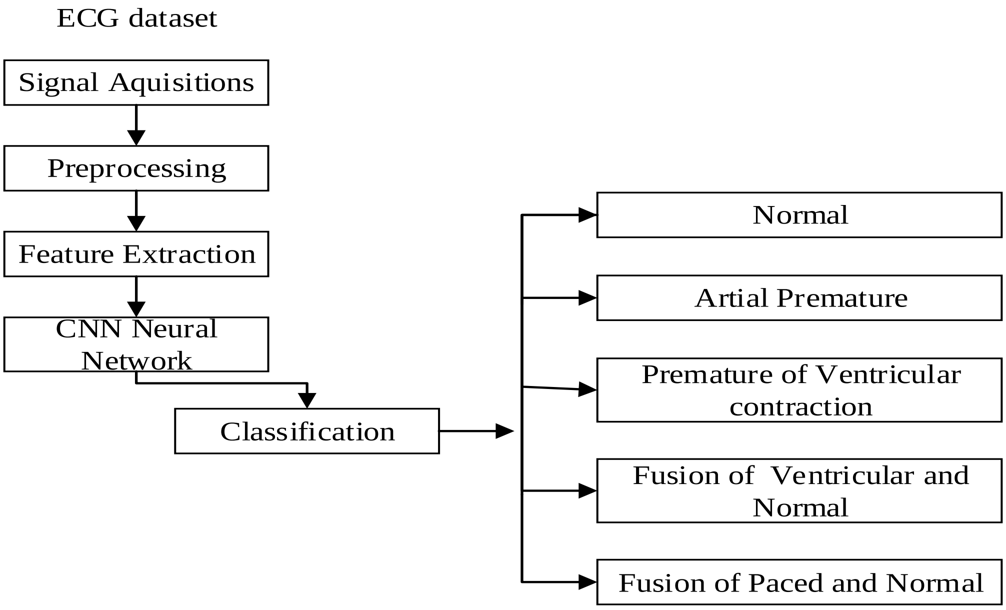

Figure 1: Architecture of the proposed system design

ECG Signal Dataset: The proposed system uses a ECG dataset. The dataset is made up of two categories of heartbeat signals, which was gotten from two popular datasets in heartbeat classification. They: are as follows MIT-BIH Arrhythmia Dataset and the PTB Diagnostic ECG Dataset. The dataset is made up of 10505 rows × 188 columns. The dataset consist of five labels which can be seen in figure 2.The heart beat signals conform to electrocardiogram (ECG) lines of heartbeats for the normal case and the cases affected by different arrhythmias and myocardial infarction. The dataset signals was preprocessed to make sure that each segment conforms to a heartbeat.

Pre-Processing: The dataset was preprocessed by removing some NAN values, infinite values and the conversion of words to 0s and 1s for easy fitting of the deep learning algorithm.

Feature Extraction: This has to do with the removal of unwanted columns/ features from the dataset. Therefore, reducing the dimension of the dataset to a pre-defined function.

Deep Neural Network: We made used of a Convolution Neural network with a total of six layers which comprises of an input layer, 5 hidden layer and an output layer which output the result to be either 0:"Normal",1: "Artial Premature", 2: "Premature ventricular contraction", 3: "Fusion of ventricular and normal", 4: "Fusion of paced and normal"

Classification: After successful training and scoring of the deep learning model, the trained model, we conducted a test using a test dataset in other to check for the performance of our model.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 4, April 2021

DOI 10.17148/IJARCCE.2021.10405

4. EXPERIMENT AND RESULT

This present a deep learning techniques for the classification of the different levels of Electrocardiogram (ECG) using a Convolutional neural network algorithm. First, we downloaded ECG Heartbeat Categorization Dataset. This dataset is made up of two categories of heartbeat signals, which was gotten from two popular datasets in heartbeat classification. They: are as follows MIT-BIH Arrhythmia Dataset and the PTB Diagnostic ECG Dataset. The dataset is made up of 10505 rows × 188 columns. The dataset consists of five labels which can be seen in figure 2. The heart beat signals conform to electrocardiogram (ECG) lines of heartbeats for the normal case and the cases affected by different arrhythmias and myocardial infarction. The dataset signals waspreprocessed to make sure that each segment conforms to a heartbeat. This dataset was read into the jupyter notebook by using pandas.read_csv function. The dataset was made into two, which are the training data and the testing data. After successful reading of the data signal from directory and solving of the in balance problem by means of data augmentation, the proposed model was trained using a convolutional neural network algorithm with a total hidden layers of six, hiden size of 128, batch_size of 96, and number of epoch to be 10. After successful training, we had an accuracy of 99% at an epoch level of 10. This can be seen in figure 8.

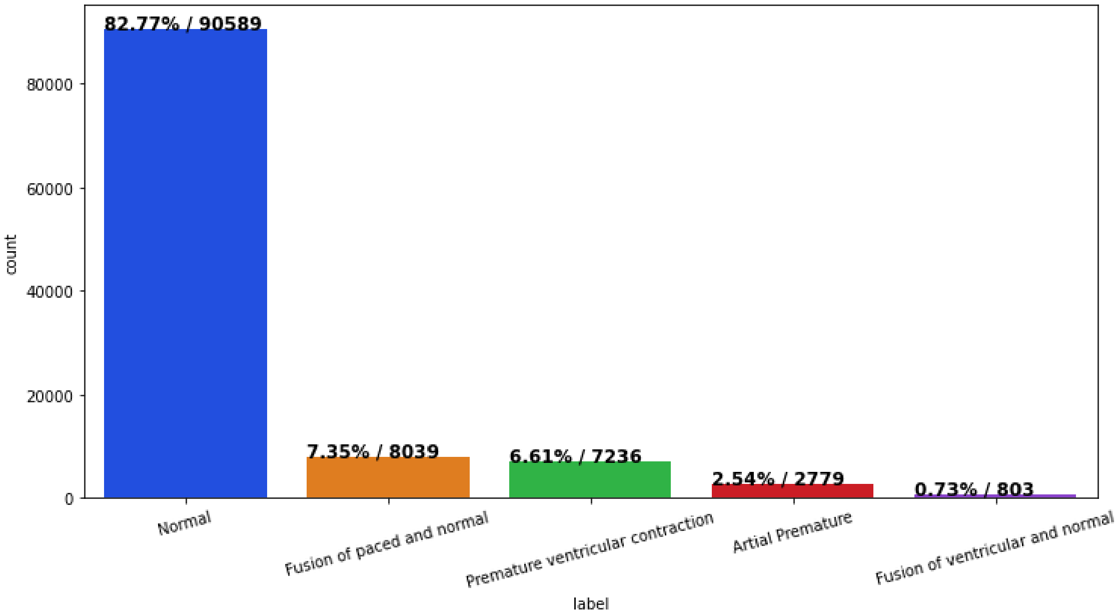

Figure 2: The label columns in the ECG dataset. The label columns can be classified as Normal, Fusion of paced and normal, premature ventricular contraction, Artial Premature and fusion of ventricular and normal.

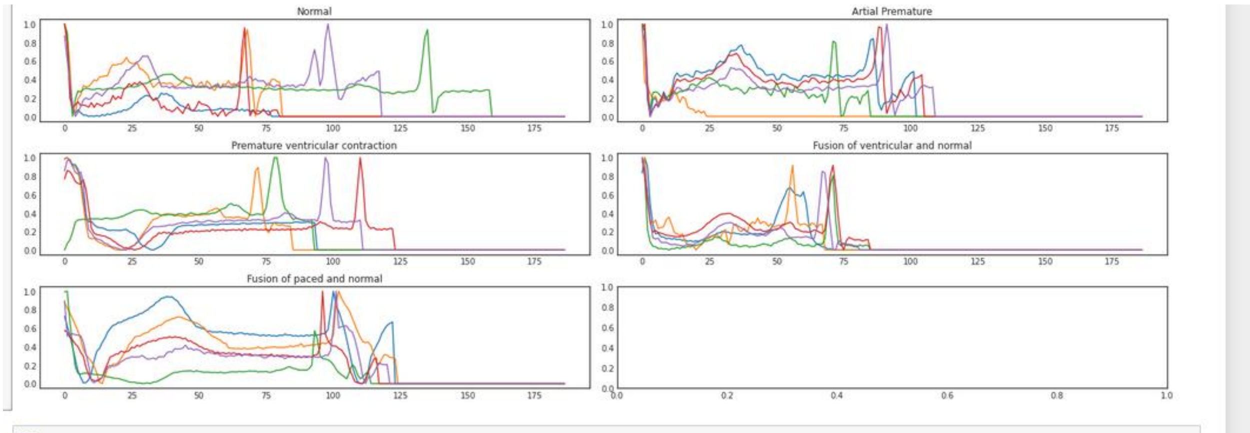

Figure 3: The ECG signal data of the labeled columns

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 4, April 2021

DOI 10.17148/IJARCCE.2021.10405

Out[11]: <AxesSubplot:>

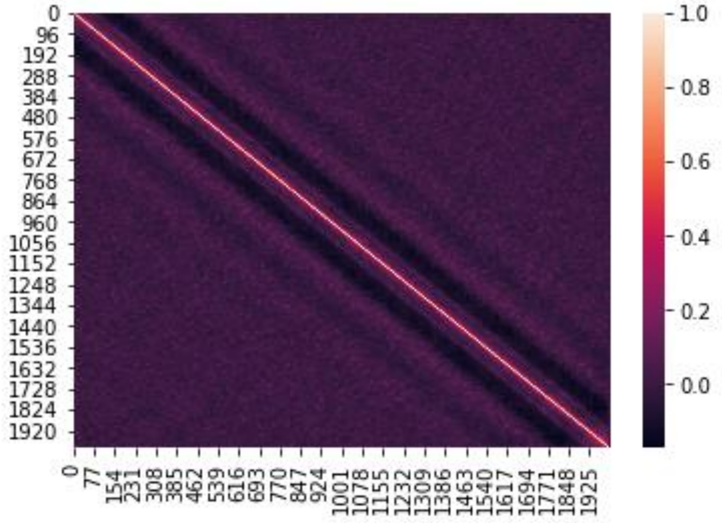

Figure 4: The correlation matrix of the dataset

9893273157039988

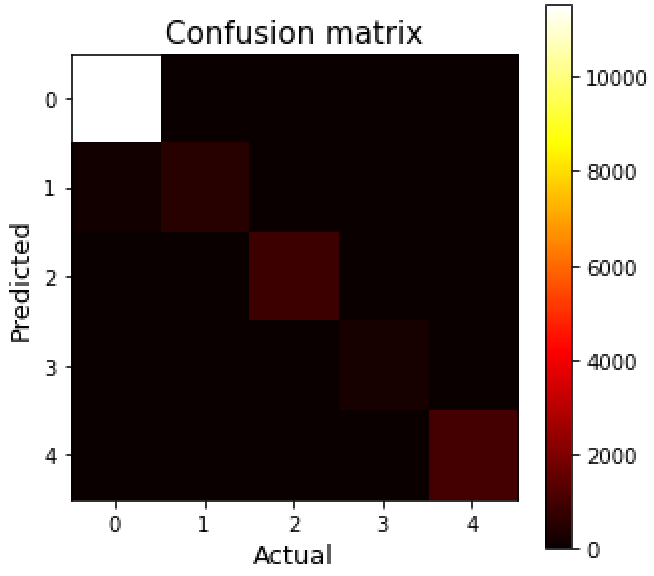

Figure 5: The confusion matrix of the predicted and actual data

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 4, April 2021

DOI 10.17148/IJARCCE.2021.10405

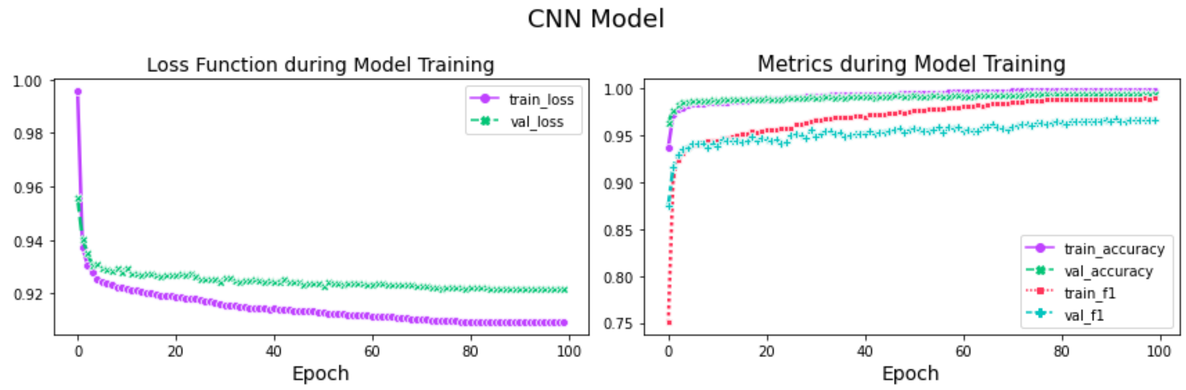

Figure 6: The loss function and the metrics evaluation of the trained model using Convolutional Neural network which had an accuracy of 99%

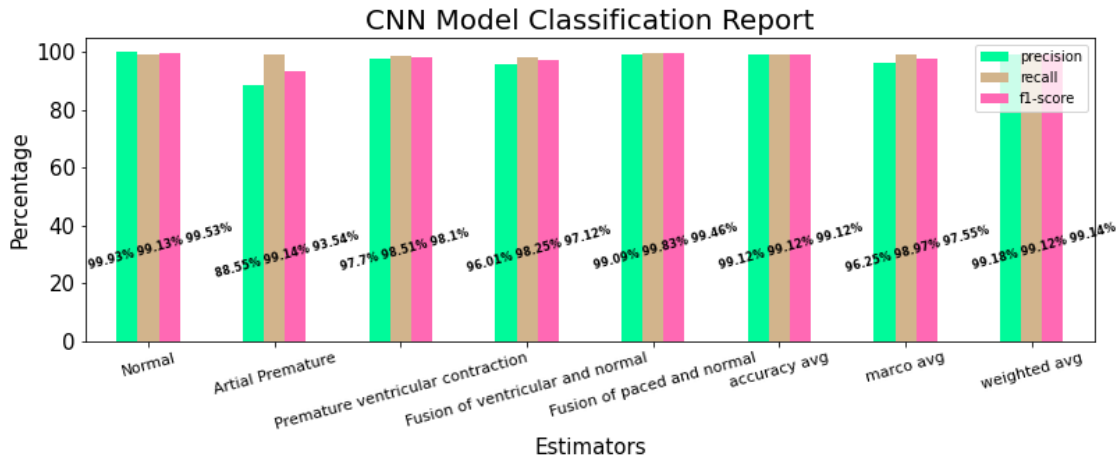

Figure 7: Classification report of the CNN model on each of the labeled classes

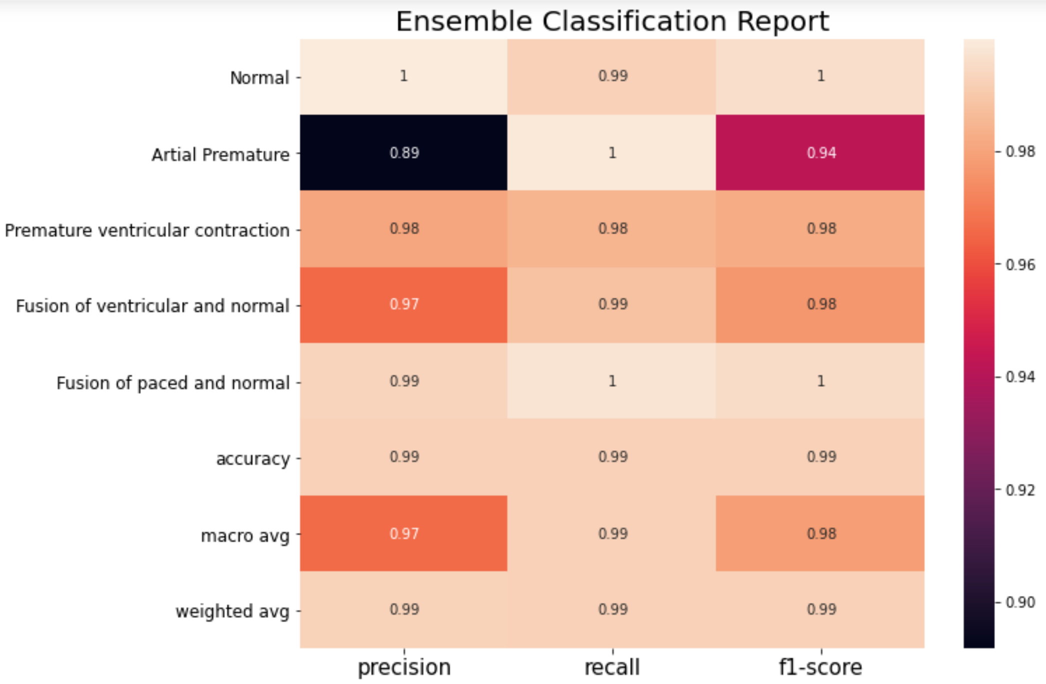

Figure 8: An Ensemble classification report of the trained model

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 4, April 2021

DOI 10.17148/IJARCCE.2021.10405

5. CONCLUSION AND RECOMMENDATION

Electrocardiogram (ECG) is an intermittent sign, which mirrors the movement of the heart. From ECG a great deal data is gotten for typical and obsessive physiology of heart. The ECG signal is non-fixed in nature and extremely hard to dissect. Clinical perception takes long time and the sign is non-fixed. This paper presents a convolutional neural network algorithm for electrocardiogram signal classification. This system uses a ECG dataset which was downloaded from kaggle.com. The dataset signals was preprocessed to make sure that each segment conforms to a heartbeat. This dataset was read into the jupyter notebook by using pandas.read_csv function. The dataset was made into two, which are the training data and the testing data. After successful reading of the data signal from directory and solving of the in balance problem by means of data augmentation, the proposed model was trained using a convolutional neural network algorithm with a total hidden layers of six, hiden size of 128, batch_size of 96, and number of epoch to be 10. After successful training, we had an accuracy of 99% at an epoch level of 10. This work can further be extended by building a real time system for Electrocardiogram classification, so that individuals can have access to it and make classification themselves.

REFERENCES

[1]. C. K. Roopa, B. S. Harish, “A Survey on various Machine Learning Approaches for ECG Analysis”, International Journal of Computer Applications Volume 163 – No 9,pp.25-33 2017.
[2]. Zhao, Q., and Zhan, L., ECG Feature Extraction and Classification Using Wavelet Transform and Support Vector Machines. International Conference on Neural Networks and Brain, ICNN&B 2:1089–1092, 2005.
[3]. El-Khafif, S. H., and El-Brawany, M. A., Artificial Neural Network-Based Automated ECG Signal Classifier. ISRN Biomedical Engineering 2013:1–6, 2013.
[4] G.D. Clifford, F. Azuaje, P.E. McSharry, Advanced Methods and Tools for ECG Data Analysis, Artech House, Norwood, MA, 2006, pp. 55–70.
[5] Q. Li, G.D. Clifford, Signal quality and data fusion for false alarm reduction in the intensive care unit, J. Electrocardiol. 45 (2012) 596–603.
[6] G.B. Moody, R.G. Mark, QRS morphology representation and noise estimation using the Karhunen–Loève transform, Comput. Cardiol. 16 (1989) 269–272.
[7]. S. Celin & K. Vasanth “ECG Signal Classification Using Various Machine Learning Techniques” Journal of Medical Systems (2018) 42: 241.
[8]. Q. Li, C. Rajagopalanc , G.D. Clifford “A machine learning approach to multi-level ECG signal quality classification” Computer Methods and Programs Biomedicine, vol.117, pp.435-447, 2014.
[9]. A. Lyon, R. A.riga, A. Mincholé , M. Mahmod, E. Ormondroy, P. Laguna, N. d. Freitas, S. Neubauer, H. Watkins B. Rodriguez, “Distinct ECG Phenotypes Identified in Hypertrophic Cardiomyopathy Using Machine Learning Associate With Arrhythmic Risk Markers”, Frontiers in Physiology vol.9, PP.1-10 2018.
[10]. N. Keshan, P. V. Parimi, I. Bichindaritz “Machine Learning for Stress Detection from ECG Signals in Automobile Drivers” 2015 IEEE International Conference on Big Data.
[11]. C A. Ledezma , X. Zhou, B. R. guez, P. J. Tan, V. D. Zuccarini “A modeling and machine learning approach to ECG feature engineering for the detection of ischemia using pseudo-ECG” PLOS|ONE vol.14. issue.18, pp.1-21 2019.
[12]. M. Simjanoska , M. Gjoreski, M. Gams, A. M. Bogdanova “Non-Invasive Blood Pressure Estimation from ECG Using Machine Learning Techniques” Sensors, pp.2-20, 2018.