International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-2, December 2019

Deep Learning Technique to Predict Heart Disease using IoT Based ECG Data

S. Clement Virgeniya, E. Ramaraj

Abstract: Scientific Knowledge and Electronic devices are growing day by day. In this aspect, many expert systems are involved in the healthcare industry using machine learning algorithms. Deep neural networks beat the machine learning techniques and often take raw data i.e., unrefined data to calculate the target output. Deep learning or feature learning is used to focus on features which is very important and gives a complete understanding of the model generated. Existing methodology used data mining technique like rule based classification algorithm and machine learning algorithm like hybrid logistic regression algorithm to preprocess data and extract meaningful insights of data. This is, however a supervised data. The proposed work is based on unsupervised data that is there is no labelled data and deep neural techniques is deployed to get the target output. Machine learning algorithms are compared with proposed deep learning techniques using TensorFlow and Keras in the aspect of accuracy. Deep learning methodology outfits the existing rule based classification and hybrid logistic regression algorithm in terms of accuracy. The designed methodology is tested on the public MIT-BIH arrhythmia database, classifying four kinds of abnormal beats. The proposed approach based on deep learning technique offered a better performance, improving the results when compared to machine learning approaches of the state-of-the-art.

Keywords: ECG classification, Deep learning, Healthcare, Tensor Flow, Keras

I. INTRODUCTION

Technology and technical way of viewing a disease now-a-days has changed a lot. The statistics of death in 20th and 21st century have a great insight. In 20th century deaths are more, at the same time birth rate is also very high. But in 21st century death rate has decreased and birth rate is moderate. Deaths due to improper health care system are a major concern in 1940's. But now-a-days timely response in healthcare is a challenge in this ever growing era.

In 21st century Artificial Intelligence took over the entire world. Machine learning and deep neural networks are subsets of AI. Artificial Intelligence is a subset of Computer Science making machines think like humans.

Revised Manuscript Received on December 30, 2019.
***** Correspondence Author
S.Clement Virgeniya*,Department of Computer Science, Alagappa University, Karaikudi, India. Email: clementvirgeniya@rediffmail.com
Dr.E.Ramaraj, Professor & Head, Department of Computer Science, Alagappa University, Karaikudi, India. Email: eramaraj@rediffmail.com

It uses computers to implement logics known as computational logics, different techniques and algorithms that work quiet differently. This is similar to our nervous system how we think, react, sense, reason out and take appropriate action [1]. Though the rate of progress in AI is less, now in the 21st century, it acts as a constellation of many technologies and its impact is seen in every individual life.

AI technologies, like targeted treatment, self -driving cars, weather forecasting are growing huge in AI. The proposed methodology focuses on deep learning techniques which is a subset of AI.

Deep Learning is a division of machine learning which is based on the layered depiction of variables known as neural networks. It often takes raw data [2] as input and undergoes many random transformations to calculate the targeted output. Deep learning also excels with unsupervised feature extraction. Where Feature extraction algorithm automatically constructs meaningful features to be further learnt. This is easier with deep learning because the existing machine learning approach uses feature selection and engineering, which is a tedious job and a waste of time.

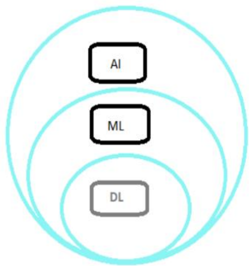

Fig. 1. Diagram representing the three broad techniques in Computational intelligence

Deep learning or feature learning or representation, learning teaches which attribute to focus on machine learning environment. The machine learning algorithm chooses the best attribute to create their own model. Whereas deep learning allows the machine to both learn for a high objective using a well-defined set of features and learn attributes themselves. Deep learning algorithms depend more on the optimal model selection and optimization through model tuning. They are more apt in an environment where knowledge of attributes in not available.

Here, Hybrid Logistic regression algorithm and rule based classification algorithm is compared with deep learning technique in terms of accuracy. It is proven that deep learning techniques excels the machine learning and data mining algorithms.

The organization of this paper is as follows. Section II, presents the related works in heart disease prediction. Section III, presents Proposed Methodology. Section IV, presents Experimental Results of the proposed methodology. And Section V, ends with Conclusion.

II. RELATED STUDY

Research shows various techniques regarding detecting cardiovascular disease in advance. Proper usage of ECG sensors is still a wandering area where proper and skilled person needs to use ECG sensors. Only technically skilled individuals and educated patients can use ECG sensors connected through raspberry pi, Arduino etc., Elisa Spano et al. [3] developed an ECG model for remote Monitoring. The author designed the prototype so that it is used by non-technical users. It is a low cost prototype with seamless integration with other smart home systems. Udit Satija et al. [4] developed an IoT based ECG monitoring system. The author validates ECG signals taken from the MIT-BIH arrhythmia and Physionet Challenge databases and the real-time recorded ECG signals. Experimental results show that real time ECG data outperforms the above two and moreover battery lifetime is also improved. Muhammad Yasin, et al. [5] developed an ultra- low power IoT platform for finding ventricular arrhythmia. The system predicts cardiovascular disease three hours in advance with less power consumption. It consumes only about 62%. Mohsen Hooshmand et al. [6] proposed Unsupervised ECG Signal Compression algorithm for Wearable fitness bands. The chief theme of the author is to develop a specialized irreversible data compression algorithm using the signal derived from the wearable technology to lower the power consumption required for wireless transmission and for more battery lifetime. Christopher Beach et al. [7] presents a wrist worn ECG sensor for healthcare monitoring combining the person and ambient sensors. This device consumes ultra- low power with the ability to convert raw data into processed and stable data for prior decision making. Ayten Ozge Akmando et al [8] analyzed the effectiveness of IoT sensor with six different IoT applications. The authors developed electrocardiogram (ECG) sensor for detecting arrhythmia

III. PROPOSED METHODOLOGY

A. Dataset Description

MIT-BIH-test arrythymia dataset is collected from physionet.org. Testing dataset consists of 188 attributes and 21892 rows and training dataset consists of 188 attributes and 87554 rows. Initially the dataset is subject to analysis. The training dataset is applied to set of rules to classify the data and hybrid logistic regression algorithm is generated. The model proved to be little less efficient compared to the proposed deep neural work. Out of 188 columns last attribute is unlabelled attribute representing the following values.
The attributes [0-4] defined in column 188 are
'N': 0, 'S': 1, 'V': 2, 'F': 3, 'Q': 4
These represent

N: Normal beat
S: Supraventricular premature beat
V: Ventricular premature beat
F: Fusion of ventricular and normal beat
Q: Indeterminate beat

Creating train and test split is an important step to evaluate the performance of the algorithm. Test data is new and need to predict the output variable. Comparing the outcomes of train dataset help us to compute the performance measure for the model on test dataset. Apart from train, test spilt, data Scaling is also an important step in Pre-processing when doing any Deep Learning project. Data scaling can be done by either normalizing or by standardizing. There are two types of data scaling methods which is already done in the MIT-BIH ECG dataset. The dataset is standardized and normalized suitable to undergo deep learning technique.

B. Architecture

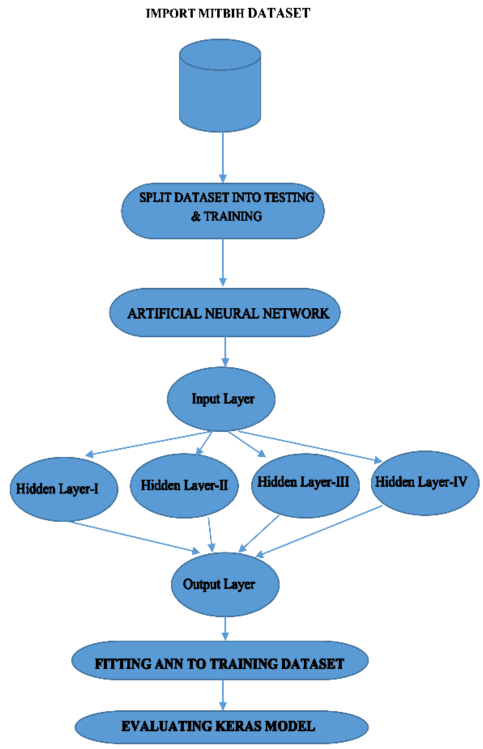

Fig. 2. Architecture of proposed methodology

C Artificial Neural Networks

Deep Learning models use artificial neural networks to train the models. There are three classes of Artificial Neural Networks. They are:

- Multilayer Perceptrons (MLPs)
- Convolutional Neural Networks (CNNs)
- Recurrent Neural Networks (RNNs)

A Multilayer Perceptrons is a typical type of neural network, which is composed of one or more layers of neurons. They are suitable for regression prediction problem (inputs are without a label) and classification prediction problem (inputs are assigned a label). MLP's deals with text data, image data, time series data and other types of data.
Convolutional Neural Networks (CNNs) map image data to an output variable. They are used with image data, regression and classification prediction problems.

Generally, CNNs work well with data that has a spatial relationship.

Recurrent Neural Networks, or RNNs, works with sequence prediction problems. They are not suitable for tabular data and image data. Traditional RNN is difficult to train. But with advent of Long short term memory (LSTM), RNN proved to be a successful neural network model.

D. Keras for Deep Learning

Keras is open source python library. It is used for creating deep learning models quickly and easily [9]. It runs on the top TensorFlow backend. Sequential API creates a model layer by layer and does create models with multiple inputs and outputs. On the other hand, functional API creates more complex models [10]. Rule based Classification algorithms are suitable for problems with the least number of attributes. They provide logical decisions. Here hybrid model is compared with deep learning technique using Keras and TensorFlow as backend. The algorithmic representation is given below.

Step 1: Import Public MIT-BIH Dataset
Step 2: For ∀ attributes apply Rule Based Classification
Step 3: Generate the results using Hybrid Logistic Regression Algorithm.
Step 4: Initializing the ANN
Step 5: Compiling ANN using sparse_categorical_cross Entropy
Step 6: Fitting ANN to the training dataset
Step 7: Evaluating through Keras model
Step 8: Predicting output on test dataset

IV. EXPERIMENTAL RESULTS

The Experimental results are visualized below. The work is carried using Python in jupyter notebook using Keras API and TensorFlow as backend. Deep Learning Technique is deployed in the proposed work. A typical single beat of ECG waveform is shown in the following Fig.3. Heart Beat is classified into five different classes representing one normal and four abnormal beats.

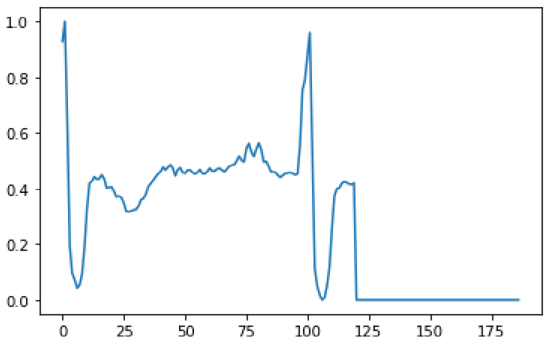

Fig. 3. Typical ECG Waveform

The Table I and Fig.4. shows the classification of ECG signal of five different classes.

Table- I: ECG Signal of Five Different Classes

Classification of Beats	Absence of Heart Disease	Occurrence of Heart Disease
Normal Beat(0)	72471	0
Supraventricular premature beat(1)	0	2223
Ventricular premature beat(2)	0	5788
Fusion of ventricular and normal beat(3)	0	641
Indeterminate beat(4)	0	6431

The Different values of ECG signal given above are represented diagrammatically below.

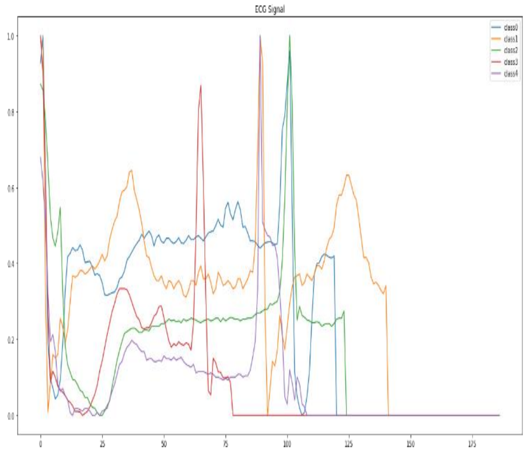

Fig. 4. Visualize one signal from each class in the same graph

MIT-BIH Dataset is subject to Rule based Classification Algorithm and Hybrid Logistic Regression Algorithm. The results show an accuracy of 85% using the so said algorithm. The confusion matrix and classification report are shown in Table II and Table III.

Table- II: Confusion Matrix with Accuracy Score of MIT-BIH Training Dataset

Confusion Matrix					
	0	1	2	3	4
0	72471	0	0	0	0
1	0	2223	0	0	0
2	0	5788	0	0	0
3	0	641	0	0	0
4	0	6431	0	0	0
Accuracy Score(in %):85.31192					

Table- III: Classification Report of MIT-BIH Training Dataset

	Precision	Recall	F1-score	Support
0	1	1	1	72471
1	0.15	1	0.26	2223
2	0	0	0	5788
3	0	0	0	641
4	0	0	0	6431
Accuracy			0.85	87554
Macro Avg	0.23	0.4	0.25	87554
Weighted Avg	0.83	0.85	0.83	87554

Retrieval Number: B7166129219/2019©BEIESP
DOI: 10.35940/ijitee.B7166.129219
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

A Deep learning technique using Artificial Neural Network for evaluating the training dataset is performed using Keras with TensorFlow Backend. An Epoch of 10 iterations is used for 87554 rows of data. Through epoch the entire dataset is lapsed forward and backward through the neural network completely. Since it takes much time for an epoch to pass an entire dataset with huge number, they are divided into a number of batches.

Here batch size is 1000 since 87554 rows of data are present. An Epoch of 10 with Accuracy and loss is presented in Fig. 5. A good model always contains high accuracy and minimal loss. Our model is good since our loss is below zero and accuracy is 94%.

The Accuracy Report of Existing Machine Learning and proposed methodology using Deep Neural Network is given in Table-IV. The Results of Machine Learning and Deep Neural Network are tabulated.

```
Epoch 1/10
87554/87554 [==============================] - 3s 32us/step - loss: 1.0
521 - accuracy: 0.8189
Epoch 2/10
87554/87554 [==============================] - 1s 16us/step - loss: 0.5
635 - accuracy: 0.8277
Epoch 3/10
87554/87554 [==============================] - 1s 17us/step - loss: 0.4
846 - accuracy: 0.8334 0s - loss: 0.4873 - accu
Epoch 4/10
87554/87554 [==============================] - 2s 18us/step - loss: 0.3
856 - accuracy: 0.8984
Epoch 5/10
87554/87554 [==============================] - 1s 15us/step - loss: 0.3
056 - accuracy: 0.9177
Epoch 6/10
87554/87554 [==============================] - 1s 17us/step - loss: 0.2
650 - accuracy: 0.9275
Epoch 7/10
87554/87554 [==============================] - 2s 18us/step - loss: 0.2
394 - accuracy: 0.9339
Epoch 8/10
87554/87554 [==============================] - 2s 20us/step - loss: 0.2
215 - accuracy: 0.9386
Epoch 9/10
87554/87554 [==============================] - 2s 17us/step - loss: 0.2
108 - accuracy: 0.9407
Epoch 10/10
87554/87554 [==============================] - 1s 15us/step - loss: 0.2
019 - accuracy: 0.9424
                                                                  Out[5]:
<keras.callbacks.callbacks.History at 0x24f22508dd8>
```

Fig. 5. Fitting the ANN to the Training Dataset

Table- IV: Accuracy report of Algorithms

ALGORITHM	ACCURACY (in%)
Machine Learning Algorithm	85
Deep Neural Network	94

Predictions were made with a dense of 5. Dense here represents the number of outputs a test dataset has. Here it produces an output of five predictions. The predictions report for a single attribute is shown below.

Predicted Output on test Dataset= [0.9870674 0.00645401 0.00115941 0.00273011 0.00258912]

V. CONCLUSION

Prediction in advance is a key goal of any Healthcare Industry. In this aspect each and every healthcare industry is working on it to achieve this ever greening goal. Though technology and technical way of viewing a disease has changed a lot in this era, timely response and accurate prediction alone saves a life. In this aspect Deep learning techniques provide an accuracy of 94%, which is a quite a satisfactory approach in predicting a disease in advance. However, in the aspect of time, Deep learning models are quite slow in training a model using the CPU. Moreover, this limitation can be swamped when models are trained using external GPU or in Cloud.

ACKNOWLEDGMENT

This Paper has been written with the financial support of RUSA - Phase 2.0 grant sanctioned vide Letter No. F. 24-51/2014-U, Policy(TNMulti-Gen), Dept. of Edn. Govt. of India, Dt.09.10.2018

REFERENCES

1. Artificial Intelligence and Life in 2030. Available Online:https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai_100_report_0906fnlc_single.pdf.
2. https://www.innoarchitech.com/blog/artificial-intelligence-deep-learning-neural-networks.
3. Elisa Spanò, Stefano Di Pascoli, and Giuseppe Iannaccone, "*Low-Power Wearable ECG Monitoring System for Multiple-Patient Remote Monitoring*", IEEE-2016
4. Udit Satija, Barathram Ramkumar, M Sabarimalai Manikandan, "*Real-time signal quality-aware ECG telemetry system for IoT-based health care monitoring*", IEEE Internet of Things Journal.
5. Muhammad Yasin, Student Member, IEEE, Temesghen Tekeste, Student Member, IEEE, Hani Saleh, Senior Member, IEEE, Baker Mohammad, Senior Member, IEEE, Ozgur Sinanoglu, Senior Member, IEEE, and Mohammed Ismail, Fellow, IEEE, "*Ultra-Low Power, Secure IoT Platform for Predicting Cardiovascular Diseases*", IEEE-2017.
6. Mohsen Hooshmand, Davide Zordan, Tommaso Melodia,"*SURF: Subject-Adaptive Unsupervised ECG Signal Compression for Wearable Fitness Monitors*", IEEE-2017.
7. Christopher Beach, Student Member, IEEE, Sammy Krachunov, James Pope, Xenofon Fafoutis, Member, Robert J. Piechocki, Ian Craddock, Fellow and Alexander J. Casson, "*An ultra-low power personalizable wrist worn ECG monitor integrated with IoT infrastructure*", IEEE-2018.
8. Ayten Ozge Akmandor, Student Member, IEEE, Hongxu Yin, Student Member, IEEE, and Niraj K. Jha, Fellow, IEEE, "*Smart, Secure, yet Energy-efficient,Internet-of-Things Sensors*", IEEE-2018
9. https://www.edureka.co/blog/tensorflow-tutorial/
10. https://machinelearningmastery.com/how-to-make-classification-and-regression-predictions-for-deep-learning-models-in-keras/

AUTHORS PROFILE

S. Clement Virgeniya is currently pursuing Ph.D in the Department of Computer Science, Alagappa University, Karaikudi, India. Her current area of interest includes Machine Learning, Deep neural Networks and IoT. She is currently working on IoT Based Healthcare System to get ECG data from Sensors for decision making. Her Areas of interest extends widely and working more on deep learning, machine leaning and how to integrate these techniques using Cloud so that anyone can access data from anywhere in future.

Dr. E. Ramaraj is working as the Professor and Head of the Department of Computer Science, Alagappa University, Karaikudi. He has a sound knowledge in many research fields especially in Data Mining, Network Security, Remote Sensing and Big Data & Analytics. He has published more than 100 international journals and presented more than 90 papers in many National and International Conferences.