Research Article

Stacked Bi-directional LSTM Layer Based Model for Prediction of Possible Heart Disease during Lockdown Period of COVID-19

Shawni Dutta[1], Samir Kumar Bandyopadhyay[2]

[1]Lecturer, Department of Computer Science, The Bhawanipur Education Society College, Kolkata, India.
[2]Professor, Department of Computer Science and Engineering, University of Calcutta, Kolkata, India.
DOI: https://doi.org/10.24321/2394.6539.202006

I N F O

Corresponding Author:
Samir Kumar Bandyopadhyay, Department of Computer Science and Engineering, University of Calcutta, Kolkata, India.
E-mail Id:
1954samir@gmail.com
Orcid Id:
https://orcid.org/0000-0001-8557-0376

A B S T R A C T

Cardiovascular Disease (CVD) affects the walls of the arteries that supply the myocardium. CVD is not a single disease. It is a cluster of diseases and injuries that affect the cardiovascular system i.e. the heart and blood vessels. Heart failure is concern in patients with COVID-19. Alternately mental anxiety during lockdown period of COVID-19 may increase CVD. For assisting health care system, an automated tool is proposed in this paper that uses Reccurrent Neural Network (RNN). A stacked bi-directional LSTM layer based model is proposed in this paper that considers interfering factors from past health records while detecting patients with cardiac problems. Experimental results have indicated promising accuracy of 93.22% in predicting patients with cardiac troubles.

Keywords: Mental Anxiety, Cardiac Troubles, Stacked Bi-LSTM, RNN, Deep Learning

Introduction

CVD causes death of human by effecting the heart and blood vessels of body. So it plays significant role in human's life. Early detection of this disease is necessary for saving patients life. CVD is often dependent on mental anxiety, daily lifestyle, working profile of people. A study suggests that symptoms of anxiety and depression (16-28%) and self-reported stress (8%) are common psychological reactions to the COVID-19 pandemic, and may be supplementary with distressed sleep.[1] Due to current situation occurred for COVID-19 lockdown phases, many employees face numerous problems in their companies. In fact, many of them are losing their jobs due to current socio-economic condition. This is affecting their mental status and increasing anxiety as well. Coronary Heart Disease (CHD) and mental illness in COVID-19 have been the leading causes of morbidity and mortality worldwide. More than 300 million people of all ages suffer from depression worldwide and the disease is expected to become the leading cause of worldwide disability by 2030. It seems that anxiety symptoms or disorders may be associated with a mild-to-moderate increased risk of developing CHD.[2] CHD is also called Coronary Artery Disease (CAD).

Using machine learning approach, heart disease detection is focused in this paper. In order to diagonse CVD, it is necessary to extract knowledge from patient's health history

database and identify relationship between interfereing factors and heart disease probability. The system is proposed in this paper automatically captures previous health records of patient and detects the tendency of heart disease. Timely detection and screening play leading role in prevention of heart attacks.

Deep Learning (DL) is an improvement over conventional artificial neural networks since it facilitates the construction of networks by incorporating more than two layers.[3] A deep learning based method is implemented in this paper that is dedicated for improving the efficiency in heart trouble prediction using medical data. A Recurrent Neural Network (RNN) with a feedback loop structure that follows DL techniques and is often helpful in forecasting purpose. This paper proposes RNN-based which assembles multiple Long Short Term Memory (LSTM) layers where LSTM is known to be a variation of RNN. This neural network classifier receives all interfering factors as features and identifies patients with heart disease troubles.

Related Work

K. Gomathi et al in[4] implemented Naïve Bayes (NB) and J48 classifier and predicted heart disease. Their results concluded that Naïve Bayes classifier reaches an accuracy of 79% where J48 classifier reaches an accuracy of 77%. P.Sai Chandrasekhar Reddy et al[5] used ANN while predicting Heart disease. This ANN method is used to predict the condition of the patient considering various parameters like heart beat rate, blood pressure, cholesterol etc.

Different classification techniques such as J48, Decision Tree, KNN, SMO and Naïve Bayes were implemented by Boshra Brahmi et al while evaluating the prediction and diagnosis of heart disease.[6] After comparing these classifiers with respect to evaluation metrics J48 and decision tree have shown the best result for heart disease prediction. J48 classifier reached the best accuracy of 83.732%. A Hybrid Random Forest with Linear Model (HRFLM) is proposed for detecting patients with cardiac disease. This model uses all features without any restrictions of feature selection.[7]

By implementing data mining rules, data related to coronary illness is extracted from a large database. For this purpose, weighted association implemented in.[8] Using rule mining algorithms on patients dataset, heart disease is predicted. Prediction results achieved 61% training accuracy and 53% testing accuracy. Historical medical data is utilised in order to predict Coronary Heart Disease (CHD) using three supervised Machine Learning algorithms such as Naïve Bayes (NB), SVM and Decision Tree (DT).[9] South African Heart Disease dataset of 462 instances is used for prediction purpose. All these algorithms were performed using 10-fold cross validation method. It was concluded that probabilistic NB classifier achieves better performance over other classifiers.[9]

Background

Deep Learning (DL) is often specified as a subfield of ML that presents a multi-layered hierarchical data representation typically in the form of a neural network by assembling more than two layers. Neural network model is built by assembing multiple layers with linear or non-linear activation functions. These layers are trained together for achieving complex problem solving approach. Activation functions are capable of executing diverse computations and produce outputs within a definite range. In other words, activation function is a step that maps input signal into output signal.[10] A RNN follows deep neural network architecture that processes both sequential and parallel information which in turn makes RNN to be superior than other traditional neural network. Similar operations like human brain can be simulated by incorporating memory cells to the neural network. Another RNN called Bidirectional RNN (BRNN) are designed to access input sequences whose starts and ends are known in advance. Because RNN can only take information from the previous context, further improvements can be made using Bi-RNN. The Bi-RNN can handle two sources of information. While considering both past and future context of each sequence element into justification, one RNN processes the sequence from start to end, the other backwards from end to start.[11]

Long Short-Term Memory (LSTM) neural network is a category of RNN that implements context based prediction which is not considered in traditional RNN. LSTM is quite potential to regulate gradient flow as well as better preservation of long-range dependencies. Every cell in LSTM is comprised of input gate, forget gate, and output gate. Use of input gate estimates when to remember input value, and when to remember or forget the value is determined by forget gate. The output gate determines when the unit should output the value in its memory.[12]

While designing the Deep models, dropout layers are included in order to reduce over-fitting problem. Use of Dropout layers randomly deactivate a fraction of the units or connections in a network during each of the training iterations.[3] One the neural model is configured, training process is executed. The training process goes through one cycle known as an epoch where the dataset is partitioned into smaller sections. An iterative process is executed through a couple of batch size that considers subsections of training dataset for completing epoch execution.[13] Since the problem addressed in this paper is binary classification problem, binary cross entropy function is used as training criterion. Binary cross entropy measures the distance from

the true value (which is either 0 or 1) to the prediction for each of the classes and then averages these class-wise errors to obtain the final loss.[14]

In order to justify the performance of a model, performance measure metrics are employed. Use of these metrics will assist in identifying the best problem solving approach. Following are the metrics those are required to justify the performance of the given model. *Accuracy* is a metric that ascertains the ratio of true predictions over the total number of instances considered.[15] However, evaluating a model in terms of accuracy may not be enough since it does not consider wrong predicted cases. For addressing the above mentioned problem, we yield two more metrics known as, Recall and Precision. *Precision* identifies the ratio of correct positive results over the number of positive results predicted by the classifier.[15] *Recall* denotes the number of correct positive results divided by the number of all relevant samples.[15] *F1-Score* or *F-measure* is a parameter that is concerned for both recall and precision and it is calculated as the harmonic mean of precision and recall.[15] *Mean Squared Error* (MSE) is another evaluating metric which is used for measuring absolute differences between the prediction and actual observation of the test samples.[15] A model having higher values of accuracy, F1-Score and lower MSE value indicate best problem-solving technique. *Cohen-Kappa Score* is also taken into consideration as an evaluating metric in this paper. This metric is a statistical measure that finds out inter-rate agreement for qualitative items for classification problem.[16]

Proposed Methodology

The purpose this paper is to detect patients with CVD. The workflow of this methodology starts with dataset collection, pre-process the data, implementing stacked bidirectional LSTM model and finally evaluate the model with respect to accuracy, f1-score, cohen-kappa score and MSE. This section elaborates the workflow of this employed approach.

Data Collection and Preprocessing

In this framework, *Heart disease dataset* from UCI machine learning repository is utilized for predicting cardiac trouble tendency of a patient.[17] The dataset can be formulated as collection of attributes that include several criterions for detecting heart disease tendency such as patient's age, gender, chest pain type, resting electrocardiographic results, maximum heart rate achieved, fasting blood sugar, serum cholesterol, the slope of the peak exercise ST segment, exercise induced angina, old peak (ST depression induced by exercise relative to rest), number of major vessels (0-3) coloured by fluoroscopy, resting blood pressure, target(probability of being heart patient). The attribute 'target' is the dependent variable in classification.

After collecting the dataset some pre-processing techniques such as NaN values handling, scaling and transform some attributes such as age, cholesterol level etc. This will assist the classifier in obtaining better predictive results. This pre-processed data is then partitioned into training and testing dataset with a ratio of 67:33. Training data is fitted into classifier for learning process and later testing dataset is used for obtaining prediction results.
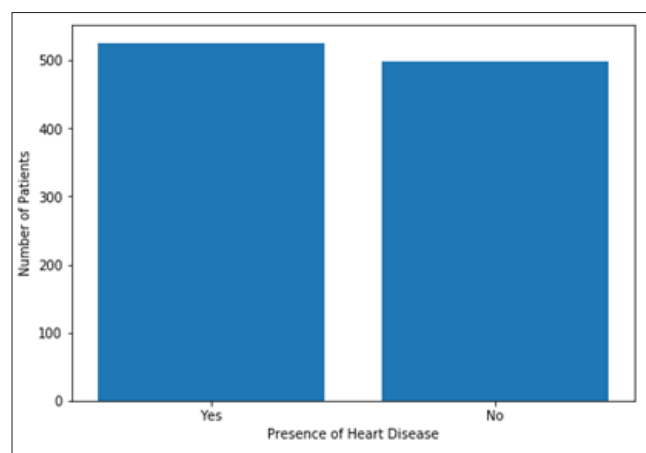

Figure 1.Distribution of target attribution over the collected dataset

Methodology

The objective of any classifier model is to map input variables into target variables considering the training dataset. The proposed classifier employs deep learning techniques in order to identify whether a patient has heart disease or not. The proposed method uses LSTM-BRNN model for such prediction. A stacked LSTM-BRNN model is implemented in this paper that stacks four Bidirectional LSTM layers and four dense layers. This stacked LSTM-BRNN model is built up using 256,128,64,16 number of nodes respectively in each layer. Each of these is followed by dropout regularization of amount 20%. Next, four dense layers are stacked by incorporating 8,4,2,1 number of nodes respectively. The first four LSTM layers and the final dense layers are activated using sigmoid activation function. Finally, the above mentioned layers are compiled using adam optimizer and binary cross entropy loss function. Construction of this model is dependent on 100 epochs and with a batch size of 32. Fine-tuning the hyper-parameters will assist in obtaining best problem-solving approach. Once this model is constructed, training data is fitted into this model. During training, this neural network model accepts a total of 1,367,993 numbers of parameters and trains those parameters in order to obtain prediction. A detailed description in terms of layers, type of layers, activation function used, output shape produced by each layer, number of parameter accepted by each layer is summarised in Table 1.

Table 1.Proposed Stacked-LSTM-BRNN Model Description

| Layer	Number of Nodes/ Percentage Rate	Output Shape	Number of Parameters Received	Activation function Used
Bidirectional LSTM layer	256	(None, 30, 512)	528384	Sigmoid
Dropout Layer	20%	(None, 30, 512)	0	None
Bidirectional LSTM layer	128	(None, 30, 256)	656384	Sigmoid
Dropout Layer	20%	(None, 30, 256)	0	None
Bidirectional LSTM layer	64	(None, 30, 128)	164352	Sigmoid
Dropout Layer	20%	(None, 30, 128)	0	None
Bidirectional LSTM layer	16	(None, 32)	18560	Sigmoid
Dropout Layer	20%	(None, 32)	0	None
Dense layer	8	(None, 8)	264	None
Dense layer	4	(None, 4)	36	None
Dense layer	2	(None, 2)	10	None
Dense layer	1	(None, 1)	1	Sigmoid
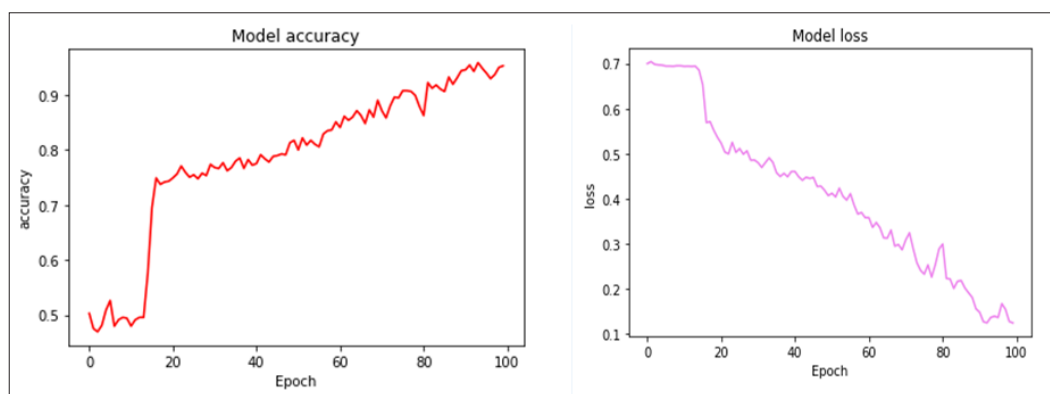
Figure 2.Training loss and Training accuracy with Stacked-Bidirectional LSTM model

Table 2.Performance of Stacked Bidirectional LSTM Model

Performance Measure Metrics	Accuracy	Cohen-Kappa Score	F1-Score	MSE
Stacked-Bidirectional LSTM model	93.22%	0.86	0.93	0.07

Experimental Result

During training process of the model, accuracy and loss are calculated for each epoch as shown in Figure 2. As the number of epochs increases, the accuracy increases gradually and reaches around a value of 0.95. In contrast, the loss gradually decreases and attains lowest value around 0.12. Once the training process is done i.e., after completing 100 epochs, accuracy, f1-score, cohen-kappa score and MSE value for test dataset. Table 2 provides the performance of prediction for the proposed model.

Conclusion

Early prediction of heart disease may decrease life of the heart patient due to anxiety in COVID-19. Considering past health record of a patient, the proposed Stacked Bidirectional LSTM Model predicts cardiac disease probabilities. This will assist the medical care units as well as accompany the doctors so that counter measures such as surgeries, medicines can be suggested. This proposed method reaches a promising and significant result that is dedicated towards heart disease prediction. Experimental results have shown prediction accuracy of 93.22%, F1-score of 0.93, Cohen-kappa score of 0.86 and MSE of 0.07.

Conflicts of Interest: None

References

1. Rajkumar RP. COVID-19 and mental health: A review of the existing literature. *Asian J Psychiatr* 2020; 52:

102066.

2. De Hert M, Detraux J, Vancampfort D. The intriguing relationship between coronary heart disease and mental disorders. 2018; 31-40.

3. Shen SHI, Dinggang Wu Gurrong. Deep Learning in Medical Image Analysis. *Annu Rev Biomed Eng* 2017; 19: 221-248. DOI: 10.1146/annurev-bioeng-071516.

4. Kirmani M. Cardiovascular Disease Prediction using Data Mining Techniques. *Orient J Comput Sci Technol* 2017; 10(2): 520-528. DOI:10.13005/ojcst/10.02.38.

5. Sai PP, Reddy C. International Journal of Computer Science and Mobile Computing Heart Disease Prediction Using Ann Algorithm In Data Mining. *Int J Comput Sci Mob Comput* 2017; 6(4): 168-172.

6. Bahrami B, Hosseini Shirvani M. Prediction and Diagnosis of Heart Disease by Data Mining Techniques. *J Multidiscip Eng Sci Technol* 2015; 2(2): 3159-1340.

7. Mohan S, Thirumalai C, Srivastava G. Effective heart disease prediction using hybrid machine learning techniques. *IEEE Access* 2019; 7: 81542-81554. DOI: 10.1109/ACCESS.2019.2923707.

8. Chauhan A, Jain A, Sharma P, Deep V. Heart Disease Prediction using Evolutionary Rule Learning. *Int Conf Computational Intell Commun Technol CICT* 2018; 1-4. DOI: 10.1109/CIACT.2018.8480271.

9. Gonsalves AH, Thabtah F, Mohammad RMA, Singh G. Prediction of coronary heart disease using machine learning: An experimental analysis. *ACM Int Conf Proceeding Ser* 2019; 51-56. DOI: 10.1145/3342999.3343015.

10. Nwankpa C, Ijomah W, Gachagan A, Marshall S. Activation Functions: Comparison of trends in Practice and Research for Deep Learning. 2018; 1-20.

11. Lipton ZC, Berkowitz J, Elkan C. A Critical Review of Recurrent Neural Networks for Sequence Learning. 2015; 1-38.

12. Soutner D, Müller L. Application of LSTM neural networks in language modelling. *Lect Notes Comput Sci* 2013; 8082: 105-112. DOI: 10.1007/978-3-642-40585-3_14.

13. You Y, Hseu J, Ying C, Demmel J, Keutzer K, CJ Hsieh. Large-batch training for LSTM and beyond. *Int Conf High Perform Comput Networking, Storage Anal SC* 2019; 1-15. DOI:10.1145/3295500.3356137.

14. Janocha K, Czarnecki WM. On loss functions for deep neural networks in classification. *Schedae Informaticae* 2016; 25: 49-59. DOI: 10.4467/20838476SI.16.004.6185.

15. Hossin M, Sulaiman MN. A Review on Evaluation Metrics for Data Classification Evaluations. *Int J Data Min Knowl Manag Process* 2015; 5(2): 1-11. DOI: 10.5121/ijdkp.2015.5201.

16. Vieira SM, Kaymak U, Sousa JMC. Cohen's kappa coefficient as a performance measure for feature selection. *IEEE World Congr Comput Intell WCCI* 2010; DOI: 10.1109/FUZZY.2010.5584447.

17. Blake C, Keogh E, Merz CJ. UCI repository of machine learning databases" [http://www.ics.uci.edu/mlearn/MLRepository.html], Department of Information and Computer Science, University of California, Irvine, CA, 1998.