IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Applications of Deep Learning in Cardiovascular Disease (CVD) diagnosis and prognosis: A Review

Kumar Rajnish, Dr. Nalin Choudhary

Vivekananda Global University, Jaipur, India

Abstract

Diagnosis and prognosis of cardiovascular diseases (CVDs) are becoming a tough task in the health care system. Hospitals are providing various expensive treatments to treat heart-related diseases. As per WHO due to CVD the death toll is approximately 18 million per year throughout the world. So the early detection of CVD can cause the degradation in death toll by providing earlier treatment. Machine learning plays a vital role in providing automated diagnosis methods through which the disease can be detected in the early stage. Several state-of-the-art machine learning algorithms such as Support Vector Machine (SVM), Naïve Bayes (NB), Logistic Regression (LR), Decision Tree (DT), kNN, and Neural Network (NN) are present which helps in providing the automated model. Nature-inspired optimization algorithms and dimensionality reduction techniques help in increasing the efficiency of the model while dealing with the microarray data. But the machine learning model accuracy downgrades with image data. As in CVDs, the image data are the main input through which the disease can be detected more accurately so the researchers move towards the deep learning approach. Some of the deep learning approaches such as the ANN, RNN, CNN, etc. play an important role while dealing with some image data and the accuracy is shown as high as compared to the ensemble machine learning models in various research work. In this current article, a brief literature survey has been done with machine learning and deep learning approaches to deal with CVDs.

Keyword: CVD, Machine Learning, Deep Learning, ANN, CNN, RNN.

1. Introduction

In the current situation, CVD is one of the leading reasons for increasing the death toll and more than a 17millin people die due to CVD as per the WHO. CVD is the combination of coronary artery disease (CAD) and heart failure (HF). Various factors such as high cholesterol, smoking, high BP can raise CVD. Due to CVD, the death toll is continuously in hike mode and the statistics show that in 2040 the death toll will be approximately 40 million per year due to this disease. Hence there is a need for the proper diagnosis and prognosis model to deal with the CVD and AI gives the way in terms of machine learning (ML) and deep learning (DL) to do so.

Machine Learning (ML) is the subset of Artificial Intelligence that provides computer systems the ability to simulate human intelligence. ML teaches machines to handle data efficiently. Mainly ML is used to interpret the pattern present in the data and to extract information from it, where the data set is very large. It is based upon a model of brain cells. In 1949, Donald Hebb created this model which is written in a book titled “The Organization of Behaviour “. But actually, the name machine learning was proposed by Arthur Samuel, 10 years later in 1959. Since that day many researchers have done tremendous work regarding how to make machines learn by themselves. It enables computer systems for searching and identifying hidden

information or patterns, without being programmed explicitly, when exposed to new datasets. The demand for ML is increasing day by day because of the presence of an abundant amount of data all over the world. In this regard, many algorithms are developed by many researchers now. According to M. Welling, the algorithms in machine learning can be categorized into supervised, semi-supervised, unsupervised, reinforcement, neural network, ensemble, and multitask learning.

There are two kinds of data present such as the biopsy data and microarray data. The biopsy data are lab test results that can't be used for machine learning to obtain a model with high accuracy. But the microarray data contain the genetic information of a patient that can be used for building a better machine learning model. ML with nature-inspired optimization algorithm and dimensionality reduction technique provides an efficient model for any kind of disease. But the difficulty comes when the input data is in image format as the ML does provide any method to deal with these kinds of data. So, the researchers move towards the DL approach is the extension Neural Network method which can be used to deal with the image data. DL is able to draw the outcomes from the complex inputs because of its ability to learn in a multi-layered representation schema. ML and DL differ from each other in terms of data set size and its category. The main objective of this research article is summarized below.

- To study the basic algorithms of ML and DL
- To provide a brief survey based on some existing ML and DL methodology to deal with the CVDs.

This article is a structure in the following way. In section 2 background study has been done while in section 3 the brief survey work has been performed. Finally, the section 4 represents the conclusion of the research work.

2. Background Study

This section will show the short introduction of various machine learning and deep learning algorithms. The ML algorithms such as the SVM, DT, kNN, RF, and LR has been focused and in the case of DL the ANN, CNN, etc. algorithms are discussed.

- SVM: Support Vector Machine is a supervised learning algorithm which can be used for classification purpose. It is used to classify the binary class dataset. The main objective of the SVM is to draw a hyperplane that can differentiate different samples or support vectors for that corresponding dataset. The performance of the SVM can be enhanced by using the hyperparameters such as the kernel trick, cost and gamma function. There are four kinds of kernel present in SVM such as the RBF, Linear, Polynomial and Sigmoid.
- DT: Decision tree examination is a prescient displaying instrument that can be applied across numerous spaces. Choice trees can be built by an algorithmic methodology that can part the dataset in various manners dependent on various conditions. Decision trees are the most impressive calculations that fall under the classification of administered calculations. They can be utilized for both classification and regression. The decision tree has three kinds of nodes such as the root node, internal node and leaf node. The leaf node is assigned with a class level and other nodes contain the test condition to make the separation. The output is obtained at the leaf node.
- kNN: It is the simplest supervised machine learning algorithm that can be used for both classification and regression. It utilizes the feature similarity index for classifying the dataset. It is a valuable procedure which can be to allot weights to the commitments of the neighbours so that the closer neighbours contribute more to the normal than the more removed ones. The neighbours are taken from a set of objects for which the class is known. This may be thought of as the preparing set for the calculation, in spite of the fact that no unequivocal preparing step is required. An idiosyncrasy of the k-NN calculation is that it is delicate to the neighbourhood structure of the information.

- RF: It is the first ensemble supervised machine learning technique that makes the utilization of more than one decision tree on different subsets of the given dataset and the average is taken for improvising the classification and predictive accuracy of that corresponding dataset. Increasing the number of trees also increases the accuracy level of the classification and
- LR: It is a supervised learning calculation used to foresee the likelihood of an objective variable. The idea of target or ward variable is dichotomous, which implies there would be just two potential classes. In basic words, the reliant variable is twofold in nature having information coded as one or the other 1 or 0. Numerically, a strategic relapse model predicts P(Y=1) as a component of X.
- CNN: It stands for a convolutional neural network which is an extended version of the neural network which is specialized in handling image data. It has three components such as the input layer, hidden layer and latent layer. The hidden layer is further categorized into the fully connected layer, pooling layer and convolutional layer. Each convolutional contains the filter for better feature extraction. One CNN model can have more than one convolutional layer. The max-pooling layer produces various grids from the output obtained from the convolutional layer. Fully connected network is almost the output of CNN which is obtained by applying the dense function.
- ANN: It is based on the collection of some connected neurons or nodes. It comprises three different layers - the input layer, hidden layer and output layer. The input layer is responsible for taking the inputs in the form of different parameters. The hidden layer is wholly responsible for performing the calculation and the output layer is responsible for providing the output based on the input.

3. Related Work

The author in [7] had proposed a model based on the fuzzy interference system for heart disease diagnosis and the model shows the sensitivity and specificity as 98% and 99% with RMS as 0.01. The author in [8] makes the utilization of DNN with $\chi 2$ statistical approach for predicting the CVD and the model is compared with various state of the art machine learning algorithms. The proposed model shows the accuracy level as 93%.

Poplin et al in [9] had proposed a deep learning model based on the deep neural network (DNN) for heart disease prediction and shown the accuracy level of the proposed model approximately 92% and the AUC is 0.71. The author in [10] had used RNN for developing one model for heart disease prediction and the accuracy is shown as 94% by using the benchmark dataset.

The author in [11] had proposed a model based on CNN for heart disease prediction, and compared the model with the LSTM model and shown that the accuracy level of the developed model is around 96% while the accuracy level using LSTM is 89%.

The author in [12] had used the CNN model for HF prediction and compared the model with different machine learning models. The proposed model shows the highest accuracy as 97% in comparison to the other ML models.

Khan et al in [13] had proposed an IoT based framework on based on the Modified Deep Convolutional Neural Network MDCNN for heart disease prediction which takes the ECG signal as the input and shown that the proposed model gives the highest accuracy level as 96.2%.

The author in [14] had proposed an ensemble deep learning method based on the DNN and feature fusion for heart disease prediction and shown that the highest achieved accuracy is 98.2%. The proposed model is being compared with other state-of-the-art machine learning algorithms.

Baccouche et al in [15] had developed an ensemble-based deep learning model in order to classify heart disease. The ensemble model consists of BiLSTM and BiGRU model with CNN. The performance is calculated on the basis of accuracy and F1 score. The accuracy obtained by using the developed model is 91% whereas the F1 score of the model is 96%. The main benefit of this model is that it takes the clinical real data and diagnose those to classify the heart disease.

In [16] the author had developed a model based on the EDCNN and implemented on IoMT for evaluation purpose. The proposed model is being compared with ANN, RNN and LSTM models and shown that the accuracy level of the developed model is 96% and precision is 99% while dealing with the real clinical data.

The author in [17] had developed one deep learning-based model for heart disease prediction by using enhanced deep learning-based CNN of EDCNN. The accuracy level shown for the proposed model is shown as 97.6% that is high as compared to the ML classification algorithms.

In [18] an autoencoder based ANN model had been proposed to predict the heart disease. The Sparse AE is used in order to obtain better representation at the training level and then the ANN is being applied to predict the disease. The accuracy level of the proposed model is shown as 95% and the comparison had been done with different ML and DL algorithms.

In [19] Ashraf et al had used the DNN for developing the deep learning-based heart disease prediction system with some benchmark data and shown that the accuracy level for the developed model is 87.64%.

Baviskar et al in [20] had used Genetic algorithm (GA) and Particle Swarm Optimization (PSO) as the feature selection algorithm and LSTM as the classification model for predicting the heart disease. It is shown that the GA-LSTM approach had an accuracy level of 93.5% whereas the PSO-LSTM approach had an accuracy level of 93.33%. But the precision for GA-LSTM and PSO-LSTM is 97.2% and 95% respectively. The author had suggested that the GA-LSTM performs well for heart disease prediction.

4. Conclusion

In this research article, the main focus is on CVD diagnosis and prognosis. The basic concepts of ML and DL are being elaborated in providing an automated model for disease prediction. This article focuses on the deep learning approach because the ML approach is unable to handle the image data which is the basis for CVD prediction. The next few well-known ML and DL algorithms have been described. Finally, survey work has been done based on different DL approaches towards CVD prediction and classification. The future scope of this research article to provide a deep learning model in addition to feature selection to deal with the image dataset and to provide a more accurate model based on some influencing parameter such as the accuracy, F1 score, specificity and sensitivity.

Reference

1. L. Ali *et al.*, “An Opt imized St acked Support Vect or Machines Based Expert System for the Effect ive Predict ion of Heart Failure,” *IEEE Access*, vol. 7, pp. 54007–54014, 2019, doi: 10.1109/ACCESS.2019.2909969.
2. A. Javeed, S. Zhou, L. Yongjian, I. Qasim, A. Noor, and R. Nour, “An Intelligent Learning System Based on Random Search Algorithm and Opt imized Random Forest Model for Improved Heart Disease Det ect ion,” *IEEE Access*, vol. 7, pp. 180235– 180243, 2019, doi: 10.1109/ACCESS.2019.2952107.
3. M. Gjoreski, A. Gradisek, B. Budna, M. Gams, and G. Poglajen, “Machine Learning and End-to-End Deep Learning for the Det ect ion of Chronic Heart Failure from Heart Sounds,” *IEEE Access*, vol. 8, pp. 20313–20324, 2020, doi: 10.1109/ACCESS.2020.2968900.
4. L. Ali, A. Rahman, A. Khan, M. Zhou, A. Javeed, and J. A. Khan, “An Automat ed Diagnost ic Syst em for Heart Disease Predict ion Based on $\chi 2$ St at ist ical Model and Opt imally Configured Deep Neural Network,” *IEEE Access*, vol. 7, pp. 34938–34945, 2019, doi: 10.1109/ACCESS.2019.2904800.
5. M. R. Ahmed, S. M. Hasan Mahmud, M. A. Hossin, H. Jahan, and S. R. Haider Noori, “ A cloud based four-t ier architecture for early detect ion of heart disease with machine learning algorithms,” *2018 IEEE 4th Int. Conf. Comput. Commun. ICCC 2018*, pp. 1951–1955, 2018, doi: 10.1109/CompComm.2018.8781022.
6. J. Schmidhuber, “Deep Learning in neural networks: An overview,” *Neural Networks*, vol. 61, pp. 85–117, 2015, doi: 10.1016/j.neunet.2014.09.003.
7. Manogaran, G., Varatharajan, R., & Priyan, M. K. (2018). Hybrid recommendation system for heart disease diagnosis based on multiple kernel learning with adaptive neuro-fuzzy inference system. Multimedia tools and applications, 77(4), 4379-4399.

8. Ali, L., Rahman, A., Khan, A., Zhou, M., Javeed, A., & Khan, J. A. (2019). An automated diagnostic system for heart disease prediction based on $\chi 2$ statistical model and optimally configured deep neural network. IEEE Access, 7, 34938-34945.
9. Poplin, R., Varadarajan, A. V., Blumer, K., Liu, Y., McConnell, M. V., Corrado, G. S., ... & Webster, D. R. (2018). Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. Nature Biomedical Engineering, 2(3), 158-164.
10. Liu, J., Zhang, Z., & Razavian, N. (2018, November). Deep ehr: Chronic disease prediction using medical notes. In Machine Learning for Healthcare Conference (pp. 440-464). PMLR.
11. Mehmood, A., Iqbal, M., Mehmood, Z., Irtaza, A., Nawaz, M., Nazir, T., & Masood, M. (2021). Prediction of Heart Disease Using Deep Convolutional Neural Networks. Arabian Journal for Science and Engineering, 46(4), 3409-3422.
12. Sharma, S., & Parmar, M. (2020). Heart Diseases Prediction using Deep Learning Neural Network Model. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 9(3).
13. Khan, M. A. (2020). An IoT framework for heart disease prediction based on MDCNN classifier. IEEE Access, 8, 34717-34727.
14. Ali, F., El-Sappagh, S., Islam, S. R., Kwak, D., Ali, A., Imran, M., & Kwak, K. S. (2020). A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. Information Fusion, 63, 208-222.
15. Baccouche, A., Garcia-Zapirain, B., Castillo Olea, C., & Elmaghraby, A. (2020). Ensemble Deep Learning Models for Heart Disease Classification: A Case Study from Mexico. Information, 11(4), 207.
16. Pan, Y., Fu, M., Cheng, B., Tao, X., & Guo, J. (2020). Enhanced Deep Learning Assisted Convolutional Neural Network for Heart Disease Prediction on the Internet of Medical Things Platform. IEEE Access, 8, 189503-189512.
17. Vennila, V., Akillesh, T., Bharathi, S., Karthik, C., & Sharmila, V. (2021). Enhanced Deep Learning Assisted Convolutional Neural Network for Heart Disease Prediction. Annals of the Romanian Society for Cell Biology, 8467-8474.
18. Mienye, I. D., Sun, Y., & Wang, Z. (2020). Improved sparse autoencoder based artificial neural network approach for prediction of heart disease. Informatics in Medicine Unlocked, 18, 100307.
19. Ashraf, M., Rizvi, M. A., & Sharma, H. (2019). Improved heart disease prediction using deep neural network. Asian Journal of Computer Science and Technology, 8(2), 49-54.
20. Baviskar, V., & Verma, M. (2021). A Model for Heart Disease Prediction Using Feature Selection with Deep Learning. In Advanced Computing: 10th International Conference, IACC 2020, Panaji, Goa, India, December 5-6, 2020, Revised Selected Papers, Part I (Vol. 1367, p. 151). Springer Nature.