Journal of Pharmaceutical Research International

33(58A): 206-212, 2021; Article no.JPRI.76780
ISSN: 2456-9119
(Past name: British Journal of Pharmaceutical Research, Past ISSN: 2231-2919, NLM ID: 101631759)

Prediction of Accuracy in Emergency Health Records using Hybrid Machine Learning Model

G. S. Raghavendra [a,b*], Shanthi Mahesh [a,b] and M. V. P. Chandra Sekhara Rao [a,b]

[a] *Atria Institute of Technology, Bangalore, Karnataka, India.*
[b] *RVR & JC College of Engineering, Guntur, Andhra Pradesh, India.*

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JPRI/2021/v33i58A34107

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/76780

Original Research Article

Received 20 September 2021
Accepted 30 November 2021
Published 15 December 2021

ABSTRACT

The quantity of digital information contained in electronic health records(EHR) has increased dramatically during the last ten years. Numerous researchers have discovered that these records may be used for a variety of other purposes as well, including applications in clinical informatics. Additionally, within the same time period, significant advancements in the area of deep learning have been made by the machine learning community. Using EHR data, we examine the existing research on applying deep learning to clinical activities. In this article we will discuss various deep learning techniques used for the classification of electronic health records along with proposing of Hybrid model for finding classification accuracy of various models.

Keywords: MHR; CNN; botlzmann machine; hybrid model; naïve bayes.

1. INTRODUCTION

Recently, most approaches for assessing large amounts of EHR data have relied on traditional machine learning and statistical techniques like as logistic regression, support vector machines, and support vector machines with reinforcement learning. Machine learning methods such as support vector machines (SVMs) and random forests are examples. Deep learning techniques have recently seen great success in a wide range of areas. Effectively capturing long-range connections in data is a challenging task. Domains are created via the use of deep hierarchical feature generation and the capture of long-range relationships in data. Because of the

**Corresponding author: E-mail: raghavendragunturi@gmail.com;*

growing popularity of deep learning techniques, as well Deep learning is becoming more essential as a result of the massive quantity of patient data being collected. Furthermore, there has been a rise in the number of articles that have been published in recent years. The data from electronic health record (EHR) systems is being processed using deep learning techniques in order to fulfil healthcare information tasks. The characteristics of new techniques, in addition to being more efficient and less time-consuming than conventional methods, may also provide higher performance while needing less time-consuming pretreatment and optimizing engineering than old ways Recent research trends in machine learning techniques in healthcare have put a strong focus on the applicability of the models, which is a good development in this field. But, when it comes to certain smart healthcare, performance of the model is critical; however, in others, interpretability is frequently chosen above accuracy. Starting with a simple model, such as a generalized additive model (GA2M), and then making it more complex (and therefore more accurate), or starting with a complex model, such as Adaboost, and then making it that much more complex, is the optimal solution in which the model has high accuracy while also being easy to understand and interpret (high interpretability) [1] Worldwide, about 15 million infants are delivered prematurely each year. This is out of a total of 1 billion neonates born each year, according to the World Health Organization (WHO). Unplanned pregnancies are those that occur before the mother's 37th week during pregnancy has been finished [1]. The consequences of preterm delivery claim the lives of about one million infants each year in the United States alone[1]. A significant number of survivors will be profoundly handicapped for the rest of their lives, including those who have visual and hearing issues, and those who have learning disabilities, as a result of the disaster. Among these diseases, pregnancy-related retinopathy is the most common and severe, causing substantial vision impairment in young children as a consequence [2]

2. LITERATURE SURVEY

i) Medical Health Records*:* An electronic health record or electronic medical record is a collection of patient health information that is stored in a digital format, such as a computer file or an electronic medical record, and can be accessed electronically [2]. They may be categorized according to the functions that they perform, which are as follows: EHRs are classified into three types: simple electronic health records (EHRs) without clinical notes, simple electronic health records (EHRs with clinical notes), and comprehensive systems[2]. Even in its most basic form, electronic health records provide researchers with a huge amount of data that is useful for their work[2]. It has already been stated that data may be transferred via a network and that it can consist of a number of various kinds of information. EHRs were initially meant to be used for internal hospital administrative tasks, and there are a variety of different schemata that may be found in different organizational structures, all of which are described here[2]. In order to address this, one of the most important difficulties is the integration and harmonization of data from a variety of different sources. Aside from that, the varied nature of the many kinds of data, which includes numerical data, date and time objects, free text, and so on, causes significant problems when dealing with electronic health records (EHRs), as previously stated [3].

In addition to deep learning, there are other types of Neural Network (NN), including convolutional neural networks and convolutional neural networks. Deep learning can be represented as a directed acyclic graph, in which the input layer takes in signal vectors and one or more hidden layers process the outputs of the preceding layer. That approach, known as deep learning, is now being developed with the goal of creating a complete system that can learn from raw data and perform specific tasks without the need for human supervision. Deep Neural Networks (DNNs) are considerably more complex than neural networks in that they include many more layers and nodes in each layer than neural networks. When compared to neural networks, the number of parameters that must be changed increases, and they cannot be learned without a significant quantity of data and strong processors [3].

iii) Clinical risk prediction*:* Medical professionals are increasingly relying on clinical risk prediction models for a number of reasons, and these models are growing more popular. Primary care physicians may use them to target treatments by identifying people who are at higher risk of acquiring diseases such as coronary heart disease (CAD) [4].

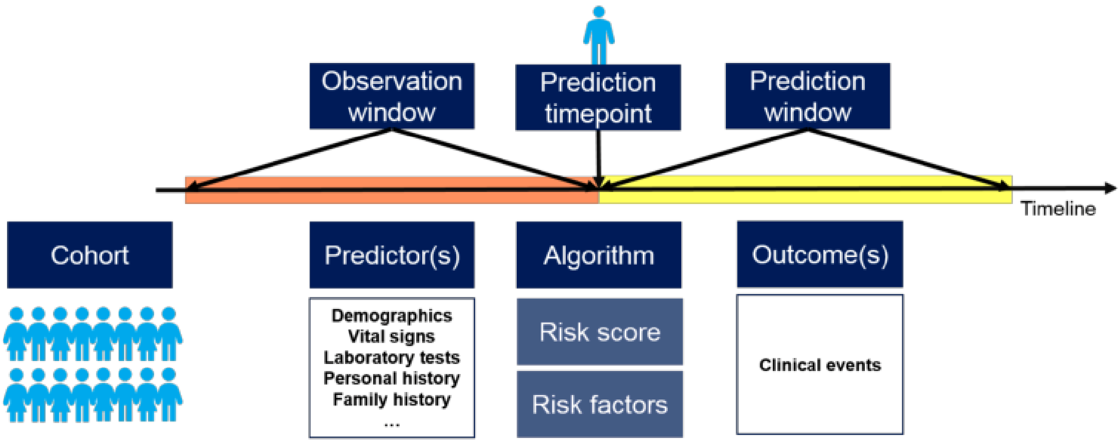


Fig. 1. Clinical risk assessment/prediction

In order to reduce the amount of time spent in the Emergency Department, the Deep Learning Emergency Department Admission Risk Score (DREADS) was evaluated in order to reduce the amount of time spent there. There were a total of 1000 people' medical records collected from a range of healthcare institutions. A prospective cohort study was conducted in order to collect these data, which were obtained during patient visits, which occur on a yearly basis. Detailed information on each patient was collected by a single researcher who was not involved in the actual treatment of the patients. The number of patient visits is on an annual basis, and in order to get the whole information, a prospective cohort study was carried out [5]. A single researcher who was not involved in patient treatment collected all of the material necessary for the calculation of the DREADS score.[5]

3. METHODOLOGY

3.1 Machine Models for Brain Tumor Segmentation and Detection

i) Processing: Images from magnetic resonance (MR) are subjected to pre-processing. Non-brain parts of the picture may be removed to fix non-brain defects. When MR images have consistent characteristics and are merged into a single spatial and intensity space, they may be utilized for further processing The development of new design looks similar neural networks (NN) architecture, according to (Mohsen, H. (2017), and requires fewer hardware requirements and processing time suited for large data sets. photos & illustrations (256 x 256). Working with three dimensional data and dense inference One MR brain picture was processed in a matter of seconds, according to CNNs, Brosch et al. (2015), and Urban et al. (2014), who claimed a total processing time of a few seconds and a maximum of one minute. Noisy image removal, converting a colour picture to grayscale, salt and peppering, and using other methods such as increasing contrast to improve image clarity are all possibilities. Process and Pre-process stages are included [5]. As a result of this, Preparation is critical for subsequent phases like as Sizing, categorization, and grading the categorization procedure becomes much easier if the pre-processing is done correctly [6].

3.1.1 Multilayer perceptrons

In this case, the input is provided by multilayer perceptrons, which is a kind of artificial neural network. When the term "perceptron" is used to refer to a single neuron model that serves as the foundation for a larger neural network, it is one of the most often seen applications. An input layer, a hidden layer, and an output layer are the three main layers of nodes that make up a network topology [6]. In an MLP, there are three main layers of nodes that make up the node structure. It is handled as if it were a neuron in both the hidden and output layers, and each node in both the hidden and output layers has a nonlinear activation function. The back propagation technique, which is supervised learning in nature, is used by MLP as part of its training process [6]. Weights are given to each neuron separately when a neural network is first created during its setup phase. As a consequence, back propagation aids in the modification of the weights of neurons in order to produce output that is more comparable to that expected. MLPs are best suited for projects including machine learning tasks, classification prediction problems, and regression prediction challenges [6].

3.1.2 Recurrent neural networks

In Recurrent Neural Networks (RNNs), the output from a previous phase is transmitted back into the current stage as input, resulting in a loop. The RNN's hidden layer, which can be located inside the RNN, is responsible for enabling this feedback mechanism. During a sequence of stages, this hidden state may be utilized to store information about the stages that came before it[5]. Because of the RNN's'memory,' it is much simpler for the model to remember all of the information that has been calculated in the past. In order to produce the output, it then uses these same parameters for each of the inputs, reducing both the number of parameters and the complexity of the process. It is one of the most widely used types of neural networks, primarily because it has a higher learning capacity and the ability to perform more complex tasks, such as learning handwriting or language recognition, than other types of neural networks. It is also one of the most widely used types of neural networks. Prediction problems, machine translation, video tagging, text summarization, and even music composition are some of the additional domains in which RNN may be used to improve performance [5].

3.1.3 Restricted boltzmann machine

Boltzmann Machine (RBM) is a kind of neural network that may be used to learn the probability distribution across an array of inputs in a generative and non-deterministic (stochastic) neural network. RBMs are shallow, two-layer neural networks that are utilized as building blocks for deep-belief networks, such as those used in machine learning. RBMs are also known as shallow, two-layer neural networks. With respect to RBM, the first layer is referred to as the visible layer, or the input layer, while the second layer is referred to as the hidden layer or the output layer. Each layer is made up of nodes, which are neuron-like units that are linked to one another across levels but are not connected to one another inside the same layer. In the development of applications like as dimensionality reduction, recommender systems, and topic modeling, RBMs are often used. Radial basis models (RBMs), on the other hand, have been progressively phased out in recent years due to the development of generative adversarial networks.

4. PROPOSED HYBRID MODEL FOR CLASSIFICATION OF MHR (RESTRICTED BOLTZMANN WITH CONVOLUTION NEURAL NETWORK FOR BACK PROPAGATION)

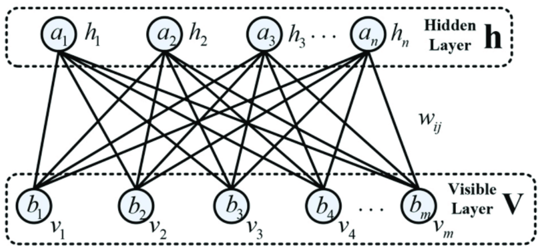


Fig. 2. Restricted Boltzmann machine

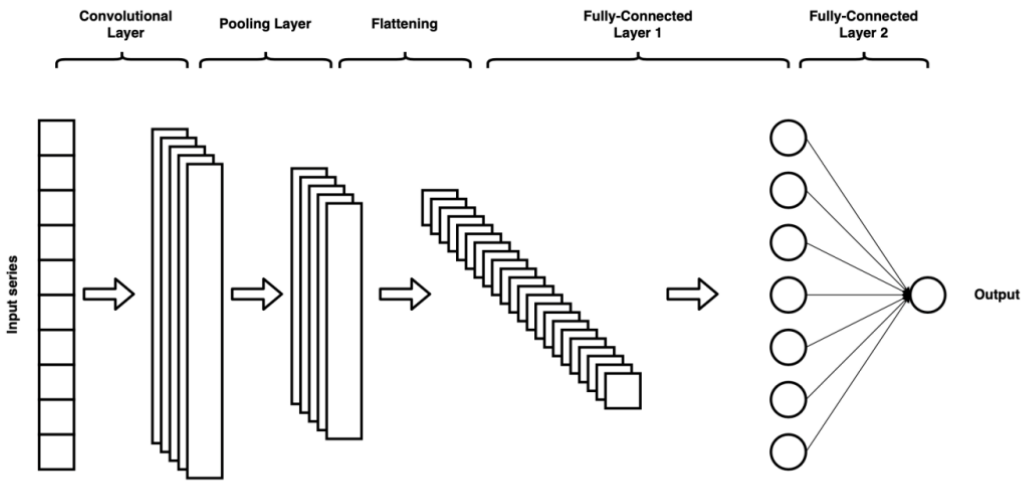


Fig. 3. CNN networks

4.1 Cons of Using a Boltzmann Machine

Boltzmann machines are algorithms that have been used for many years. However, they have caused many difficulties.

1. In order to compute probabilities, time must be spent collecting statistical data.
2. At what rate do the weights change?
3. How to make temperature adjustments during simulated annealing.
4. What to look for to determine whether the network has achieved equilibrium temperature
5. In comparison to back propagation, the primary drawback of Boltzmann learning is that it is considerably slower.

5. RESULTS AND DISCUSSION

5.1 Accuracy of SVM and CNN

When the pixel-based reflectance samples were employed, without the segmentation size, CNN was shown to have a static-significant advantage over SVM in terms of overall correctness. Fallopia japonica was the most misclassified class for both classifiers, with class correctacies all over 70%. The Fallopia japonica class was misclassified as numerous species had the fewest training samples, and continuously poor class correctacies. Most of the time, Acer platanoides were correctly identified**.** The influence of pixel-based training samples was considerable in SVM and CNN classifications. The difference between object-based and pixel-based training samples may be observed as an increase in sample size. Although the total number of samples is the same for all sample types, pixel-based reflectance data give extra spectral reflectance values to classifiers trained. Support vector machines successfully employ just a fraction of a dataset as training material. This is due to the fact that they can dependably identify the decision boundary based just on support vectors. As a result, for well-separated classes, the number of observations necessary to train an SVM isn't large. Instead, neural networks are trained using batches of data that are fed into them. Instead, neural networks are trained using batches of data that are fed into them. This means that the neural network's particular decision boundary is very reliant on the sequence in which the data batches are supplied to it. This, in turn, necessitates the processing of the entire training dataset; otherwise, the network may perform horribly.

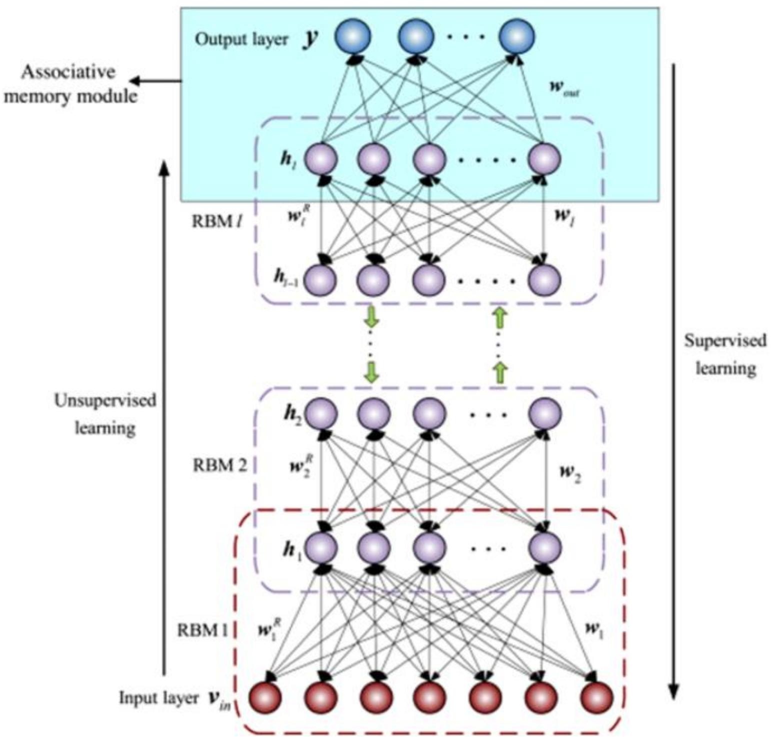
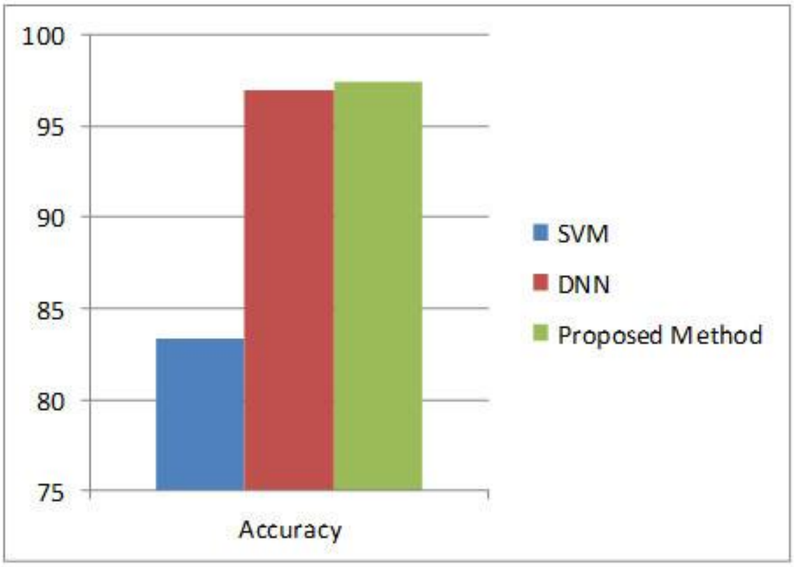


Fig. 4. Hybrid model



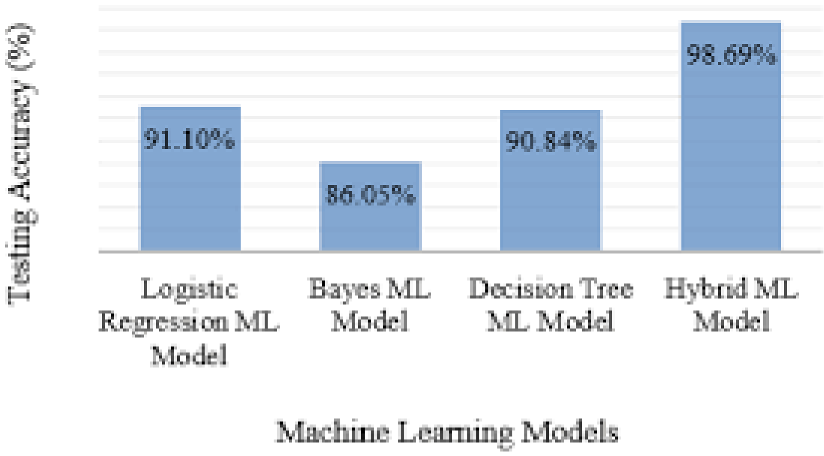


Fig. 5. Accuracy of models

5.2 Accuracy of Hybrid Model (Proposed Method)

The goal of this work was to improve the prediction accuracy of a hybrid machine learning method. As a consequence, this part presents the analytical findings and a discussion of the prediction accuracy of individual and proposed hybrid machine learning classifiers on a selected dataset. We have chosen two supervised machine learning algorithms: the Jrip method and the Nave Bayes algorithm. By merging two carefully chosen algorithm. The graphs below indicate the outcomes of both methods of implementing our planned hybrid, namely R programming language and Weka software tool. The average probability combination rule of ensembling is used while utilizing the Weka software tool because it provides superior prediction performance on the given dataset.

6. CONCLUSION

The goal of this work was to improve the prediction accuracy of a hybrid machine learning method. As a consequence, this part presents the analytical findings and a discussion of the prediction accuracy of individual and proposed hybrid machine learning classifiers on a selected dataset With the use of hybrid model which may be combination of any two or three classification models are giving an accuracy of over 98 percent and it can be further modified for clustering process also.

CONSENT

It is not applicable.

ETHICAL APPROVAL

It is not applicable.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Benjamin Shickel, Patrick James Tighe, Azra Bihorac, Parisa Rashidi. Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR). IEEE Journal of Biomedical and Health Informatics. Sept. 2018;22(5) .
2. Tamer Karatekin, Selim Sancak, Gokhan Celik, Sevilay Topcuoglu, Guner Karatekin, Pinar Kirci, Ali Okatan. Interpretable machine learning in healthcare through generalized additive model with pairwise interactions (GA2M): Predicting Severe Retinopathy of Prematurity 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML).
3. Luciano Caroprese, Pierangelo Veltri, Eugenio Vocaturo, Ester Zumpano. Deep learning techniques for electronic health record analysis. 2018 9th International Conference on Information, Intelligence, Systems and Applications (IISA).
4. Zhengxing Huang, Wei Dong, Huilong Duan, Jiquan Liu. A regularized deep learning approach for clinical risk prediction of acute coronary syndrome using electronic health Records IEEE Transactions on Biomedical Engineering. May 2018;65(5).
5. Subiksha KP. Accelerating health care inference using DL Boost 2019 9th International Conference on Emerging Trends in Engineering and Technology - Signal and Information Processing (ICETET-SIP-19)
6. Somasundaram S, Gobinath R. Current trends on deep learning models for brain tumor segmentation and detection – A review. 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)

© 2021 Raghavendra et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/76780