International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

ANALYSIS ON DEEP-LEARNING-FOR-HEART-DISEASE-PREDICTION

Saloni Fathima [1], Faculty in CSE, Kakatiya University College of Engineering,
salonifathimamd@gmail.com
Annadi Mukesh Reddy [2], Data Scientist, Amazon Development Center
mukeshreddy246@gmail.com

Abstract:
One of the dangerous diseases is heart disease. This condition affects a big population worldwide. Taking into consideration death rate and many people who have heart illness, the importance of early diagnoses of cardiac disease is emphasized. There are several classic prediction methods for such a disease, but not enough. A medical diagnostic system is urgently needed which can anticipate early diagnosis. Machine learning approaches are now becoming more popular. This paper is utilized to create a prediction and medical diagnostic system for Early Stadium Convolutionary neural nets (CNNs). During testing, more than 85% of the predictions of lack and presence of heart disease are shown to be exact.

1. Introduction

Heart disease represents a number of heart disease. Diseases under the umbrella of heart disease include diseased blood vessels, including heart disease, cardiac rheumatoid disorders (arrhythmia) and born cardiac defects, including congenital heart defects. In many respects, the word "cardiovascular disease" is interchangeable with that phrase. Cardiovascular disease is usually a problem involving blood arteries that are narrower or constricted and can lead to heart attack, chest pain (angina) or stroke. Other cardiac disorders are considered kinds of heart disease, including those affecting the muscles, valves or rhythm of the heart.

Heart disease is one of the world's largest causes of illness and death. Cardiovascular disease prediction is considered to be one of the most important topics in the analysis of clinical data. The data available in the medical sector is enormous. The enormous accumulation of raw healthcare data is transformed by data mines into information that can support educated judgments and forecasts.

In order to be able to detect concealed information for efficient decision-making, the health care business collects vast amounts of health care data.The researchers employ data mining techniques to aid health care practitioners in diagnosing heart illness, motivated by the increasing worldwide mortality of heart disease patients every year. They also have a big quantity of patient data to obtain helpful information. Data mining is the explorations of massive datasets, which extracts from classic statistical approaches hidden or unknown patterns, relations, and knowledge. Data mining therefore refers to the extraction or extraction of information from enormous amounts of data.Increasing health policy and prevention of hospital mistakes, early detection, disease prevention and preventable hospital mortality are to be used for data mining

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

applications. A heart prediction system can help health workers forecast heart disease based on the patient's clinical data. Thus, using data mining techniques and applying some kind of data mining on different heart disease variables, by adopting a prediction system for heart disease the potential of patients to be diagnosed with heart disease can be more likely to forecast.This paper proposes a new model which improves the precision of the Decision Tree in detecting people with heart disease. It uses the various Decision Trees algorithms.

This makes heart disease an important concern. However, a number of contributory risk factors, including diabetes, high blood pressure, excessive cholesterol, abnormal pulse rate and many others, make it difficult to detect the heart disease. Because of these limits, researchers have resorted to current methodologies such as Data Mining and Machine Learning for disease prediction.

Medical organizations worldwide collect information on many aspects connected to health. These data may also be used to obtain beneficial insights using different machine learning approaches. However, the data collected are quite vast and often highly noisy. These data sets, too big to understand by human minds, can be simply investigated utilizing various approaches of machine learning. These algorithms have therefore recently proved extremely useful to properly forecast the presence or absence of heart disease.

The biggest problem in the medical industries is the prediction and diagnosis of cardiac disease, depending on characteristics such as physical exam, symptoms and patient signals. Factors affecting heart disease include body cholesterol, smoking and obesity, family history of diseases, blood pressure and workplace. In forecasting heart disease, machine learning algorithms play a crucial and accurate part. Technological advances allow machine language to manage unstructured and exponentially expanding data, in conjunction with big data tools. [5].

In this paper presents Heart disease prediction using CNN. The paper is organized as follows. Section 1 consists of the general introduction of the heart disease prediction. Section 2 consists of literature review of heart disease prediction. Section 3 focuses on the proposed Deep-learning for heart disease prediction of the CNN. Section 4 has results for the simulations. Section 5 the important conclusions.

II. Literature Review

Singh et al., (2018) [1] An effective prediction system for heart disease (EHDPS) is established by means of the neural network to predict cardiac disease risk. The technique uses age, sex, blood pressure, cholesterol and obesity for prediction using 15 different medical indicators. The EHDPS forecasts the probability of cardiac illness for patients. It permits important knowledge to be established, e.g. connections between cardiovascular and pattern medical aspects. The multilayer neural perceptron network was used as the training algorithm with reverse reproduction. The results showed that the suggested diagnostic method can estimate the risk level of heart disease efficiently.

Indrakumari et al., (2020) [2] The author shows, with the help of the K-means

algorithm, considering and forecasting the risk variables causing the heart illness, and analyzing them using publicly made cardiac disease data. This dataset has a total of 209 records with 8 attributes: age, chest pain type, blood pressure, blood glucose level. K-means clustering algorithm, data analytics and visualization tool are utilized to prevent heart disease. The paper discusses pre-processing approaches, performance of classifiers and assessment measures.

Singh et al., (2020) [3],The prediction system for disease awareness is vital in order to tackle the problem. Machine learning is the branch of AI; it offers prestigious assist in the prediction of any type of occurrence that takes place through natural events. In this article we analyze the accuracy of machine learning algorithms to predict heart illness, for which we use a UCI repository data set for training and testing, k-nearing neighbour, decision-tree, linear regression and support vector machine.

Gao et al., (2021) [4] Ensemble learning approaches are employed in this author to improve the performance of cardiovascular prediction. In order to choose essential features from the dataset, two features of extraction methodologies: linear discriminant analysis (LDA) and main component analysis (PCA) are applied. For specified features a comparison is applied between machine learning algorithms and ensemble learning methods.

Bhatla et al., (2012) [5] The purpose of this research is to analyze the different data mining strategies for prediction of heart disease introduced in recent years. The observations have shown that 15-attribute neural networks have exceeded all other strategies for data mining. The investigation further concludes that with the help of genetic algorithm and the selection of sub-sets of features, decision trees have also showed good precision.

Sharma et al., (2016) [6] In light of the parameter offered on their health, the author planned a framework which could effectively identify the principles for forecasting patients' risk. Helping non-specialists make the right decision on the risk of heart disease is the key contribution of this work. In the form of original rules, pruned rules, rules without duplications, classified norms, sorting norms and Polish the rules generated by the proposed system will be given priority. The implementation of the framework is evaluated to show that the arrangement is accurate and the results show that the framework has amazing potential for a greater anticipation of the risk of coronary heart disease.

Sultana et al., (2016) [7] This author discusses the problem of heart disease prediction based on data mining approaches according to the input attributes. We examined the prediction of heart illness by employing Weka software, such as KStar, J 48, SMO and Bayes Net. The performance of various methods of data extraction will be tested by utilizing a standard data set and a data set to combine prediction accuracy findings, a ROC curve and an AUC value. The performances of KStar, Multilager Perceptron and J48 are optimal based on SMO and Bayes Net performance factors.

Almustafa et al. A comparative investigation of the different classifiers for the classification of the heart disease dataset has been done to correctly classify

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

and/or to forecast HD occurrences with minimum attributes. The set has 76 class attributes and only 1 025 patients from Cleveland, Hungs, Switzerland, and Long Beach use a sub-set of 14 attributes. The methods employed in this case are: K-Nearest Neighbor (K-NN), Naive Bayes, J48, JRip, SvM, Adaboost, Stochastic Gradient Decent (SGD) and Decision table to show the performance of the algorithms selected in classifications that best categorize or predict HD circumstances (DT).

III. ProposedMethod

Deep learning

In today's life, a system built on intelligence that takes its own decision is very necessary. The breakthrough is one of the largest failures in the artificial neural network that makes the new method termed deep learning one of the main reasons. The new method is the learning of machines that achieve Artificial Intelligence's underlying purposes. The world requires a deep system of learning that thinks like a neocortex. A number of models are learned from the training samples by the system in accordance with actual sense for forecasting and classification.Deep learning is the best way to train a big neural network. When considering the huge amount of data, a deep learning strategy delivers higher performance than any other methodology. The benefit of deep learning is that it enables the automatic extraction of features and makes learning more easy. With the help of a deep learning algorithm, this problem in uncontrolled and supervised learning is addressed.

Health researchers say that the danger of heart disease is significant and that it suddenly impacts patients' lives. The reasons of heart disease are diverse and some are caused by lifestyle, gene and smoking change. Many genetic variants raise the risk of heart disease. The treatment for cardiac disease such as medicines for decreasing cholesterol, insulin and blood pressure is initiated once the forecast of cardiac conditions is predicted accurately. It's not easier to predict an illness.The prediction of heart attack requires a continuous cholesterol and blood pressure monitoring during the course of life. The basis for heart attack plaques, which are significantly more sensitive to patients, should be identified. The health sector is investigating several studies on heart disease detection.

Table heart disease factors with symptoms

TABLE I. HEART DISEASE FACTORS WITH SYMPTOMS

Heart Disease -Risk Factors	Heart Disease- Symptoms
Age	Discomfort, pressure or heaviness,
Sex	Pain in the chest or arm, or below the breastbone.
Family History	Discomfort burning to the back, jaw, throat, or arm.
Smoking	Fullness, indigestion, or choking feeling.
Poor Diet	Sweating, nausea, vomiting, or dizziness.
High Blood Pressure	Extreme weakness, anxiety, or shortness of breath.
High blood cholesterol level	Rapid or irregular heartbeats
Diabetes	
Obesity	
Physical inactivity	
Stress	
Poor hygiene	

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Table 2. Heart Diseases types

TABLE II. HEART DISEASES TYPES

Heart Disease Types	Description
Coronary heart disease	Block in the coronary blood vessels, which leads to the reduction in supply of blood and oxygen to heart.
Angina pectoris	Chest pain will occur due to the insufficient supply of blood to heart.
Congestive heart failure	Heart is not able to pump enough blood.
Cardiomyopathy	Weakening or a change in the heart muscle.
Congenital heart disease	Defect in the structure of the heart or its functioning, leading to abnormal formation of heart.
Arrhythmias	Disorder in the rhythmic movement of the heartbeat.
Myocarditis	Inflammation of heart muscle by viral, fungal, and bacterial infections affecting the heart

Convolutional Neural Networks

A state-of-the-art version of deep neural networks with spectral levels is the CNN. CNNs are efficient models for statistical predictions, modeling, etc. Just three additional notions, i.e. local filters, max pooling and the sharing of weights make it stronger than DNNs. Figure 1 illustrates the architecture of CNN used to forecast cardiac illness. The CNN has a few convergence pairs and max bathing layers. The layer of pooling always follows the convolutionary layer.

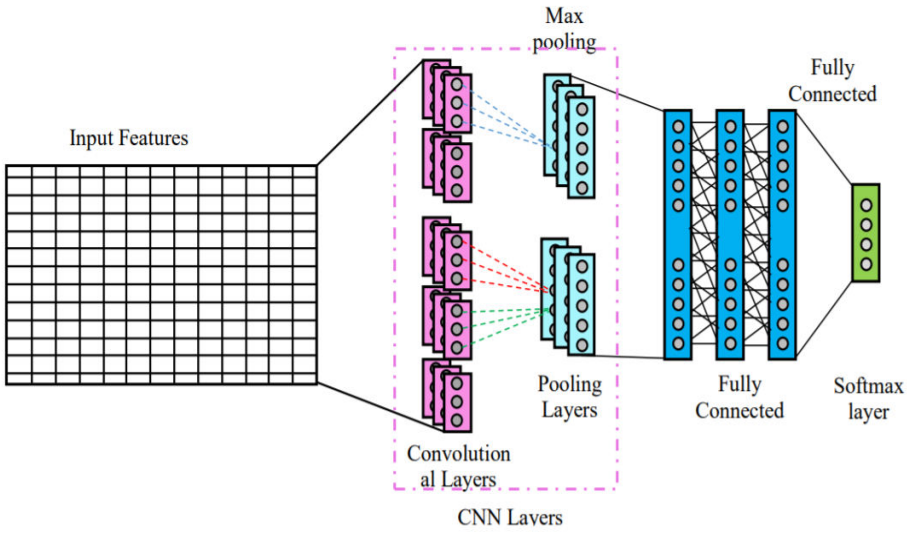

Figure 1: The architecture of CNN used for Prediction Model

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Convolution Layer

The first layer is the layer of convolution. The pixels of images are still linked to the layer. It discover its properties by doing this. This is done by dividing the image input into smaller pixel boxes. In principle, it is a mathematical process that takes two things into account: the image matrix and the filter/kernel.

Padding

The filter applied often doesn't adequately suit the input image. There are two possibilities available to us:

1. The addition of padding (zeroes for zero padding) for a better fit might be one remedy.
2. Using valid padding–weeding bits of an image that did not fit well with the filter and only using the rest.

ReLu

The rectified linear unit is the most often used deep learning activation function. If the input value is positive or not, ReLu functions by returning 0.

This can be shown as:

Two key purposes are the activation functions:

1. Help an interaction effects model account. This is when a variable A affects a forecast differently based on the value of B. For example, if my model had to know if there was an increased risk of diabetes for a given body weight, it would have to know a person's height. ReLu(a) = Max (0, a) shows that there are high risks for small people, whereas huge people have high health levels. Thus, the influence of body weight on diabetes risk is height-dependent, and we would say weight and height interact.
2. Help a nonlinear effects model account. That simply implies that I do not graph a variable on the horizontal axis and my predictions on the vertical axis. In other words it is different in different values from this predictor to increase the predictor by one.

Pooling

The bundling area is utilized to delete extraneous parameters if necessary. Spatial pooling decreases input data dimension, but keeps its significant characteristics.

This can be of various kinds:

• Max Pooling: It takes the most important thing from the map.

• Average Pooling: This might also be called the largest element.

• Sum Pooling: The sum of each element is calculated and utilized in the feature map.

Fully Connected Layer

Next, the fully connected layer or FC layer takes the matrix once it becomes vectors. The map matrix is changed to the form of a vector (V1, V2, V3 and ...). The FC layer uses and incorporates these vector-like properties into a model. Finally, another categorization activation function is performed. The CNN activation function is commonly Softmax.

Softmax Function

Function Softmax estimates the distribution of probability for the event over 'n' various events. Below are the few softmax function properties.

• Calculated probabilities will range from 0 to 1.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

• The sum of all these computed probabilities will be equalto 1.

….

Inputs of the Fusion Process

Because of the crucial color correction in underwater, we apply to the original image our white balance procedure. This process tries to improve the appearance of the image by removing undesirable color casts created by different illuminants.

White balance has noticeable impact with water deeper over 30 ft, as the colors absorbed are difficult to restore. Consequently, we do Histogram Equalization to get our initial input.Equalization of histogram aims at rectifying the global contrast and is relevant, because the histogram is generally overly bright. This adjustment enhances the difference between darker and lighter parts at the cost of features being lost in regions under or above exposed.

To offset this loss, the second input is the sharp form of an equalized histogram image. We therefore follow the unsharp, filtered form of the image to sharpen the image.

IV. Results

Figure 3 shows that the majority of those affected are 58 years of age and 57. The condition is mostly experienced by adults in the age group 50+.

Figure 3 Variation of age for each target class

Figure 4 shows the age and gender distribution for each target class. Figure 4. This is older than males for females who have the condition.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

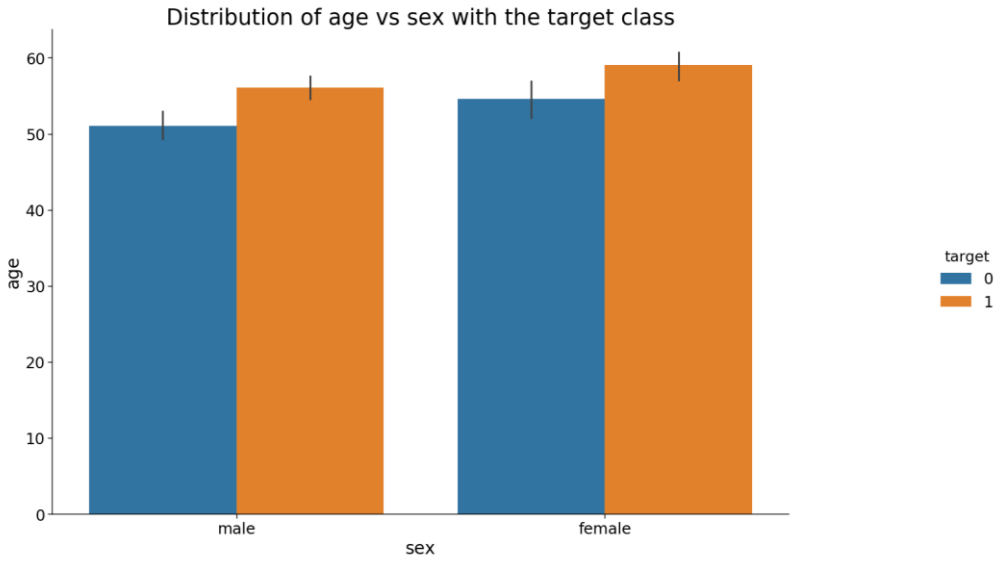

Figure 4 the distribution of age and gender for each target class

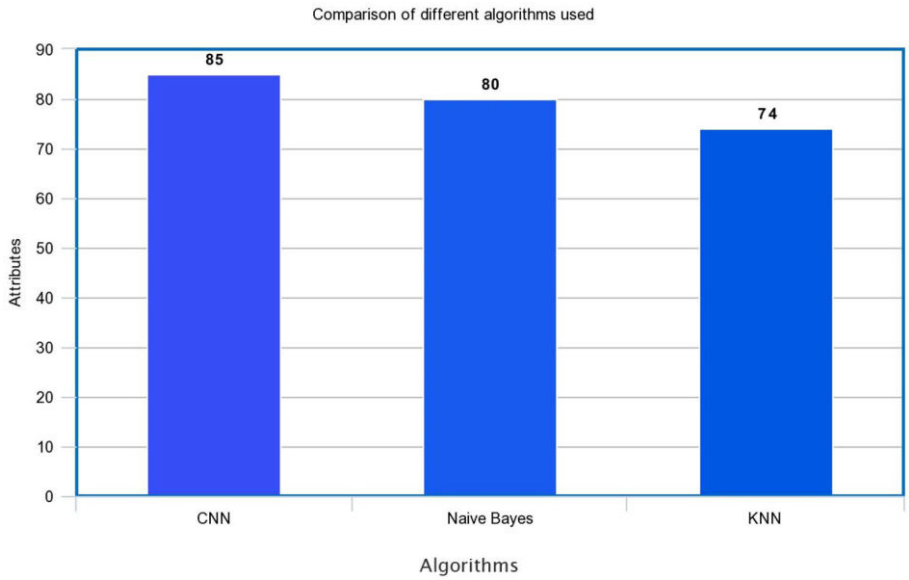

Fig. 9 Graphs of different algorithms

Figure 9 displays a comparison for different heart prediction methods, including CNN, Naïve Bayes, KNN, etc. The model accuracy is highest for the model developed using CNN and then comes with an algorithm from Naive Bayes with little more precision than CNN and then with a low accuracy of the KNN method.

Conclusion

Heart disease is today one of society's most important issues. The chances of becoming heart disease based on risk factors may hardly be manually determined. Machine learning approaches are, nonetheless, useful for the prediction of current data output. CNN algorithm is a way of determining risk of early heart

disease using structured data. The accuracy of our models can be as high as 85-88%. We propose that our technique be extended to include unstructured data in future endeavour. From now, the medical physicians have accepted all attributes and laboratory tests considered.

Reference

[1]. Singh, Poornima, Sanjay Singh, and Gayatri S. Pandi-Jain. "Effective heart disease prediction system using data mining techniques." *International journal of nanomedicine* 13, no. T-NANO 2014 Abstracts (2018): 121.

[2]. Indrakumari, R., T. Poongodi, and Soumya Ranjan Jena. "Heart Disease Prediction using Exploratory Data Analysis." *Procedia Computer Science* 173 (2020): 130-139.

[3]. Singh, Archana, and Rakesh Kumar. "Heart disease prediction using machine learning algorithms." In *2020 international conference on electrical and electronics engineering (ICE3)*, pp. 452-457. IEEE, 2020.

[4]. Gao, Xiao-Yan, Abdelmegeid Amin Ali, Hassan Shaban Hassan, and Eman M. Anwar. "Improving the Accuracy for Analyzing Heart Diseases Prediction Based on the Ensemble Method." *Complexity* 2021 (2021).

[5]. Bhatla, Nidhi, and Kiran Jyoti. "An analysis of heart disease prediction using different data mining techniques." *International Journal of Engineering* 1, no. 8 (2012): 1-4.

[6]. Sharma, Purushottam, Kanak Saxena, and Richa Sharma. "Heart disease prediction system evaluation using C4. 5 rules and partial tree." In *Computational intelligence in data mining—volume 2*, pp. 285-294. Springer, New Delhi, 2016.

[7]. Sultana, Marjia, Afrin Haider, and Mohammad Shorif Uddin. "Analysis of data mining techniques for heart disease prediction." In *2016 3rd international conference on electrical engineering and information communication technology (ICEEICT)*, pp. 1-5. IEEE, 2016.

[8]. Almustafa, Khaled Mohamad. "Prediction of heart disease and classifiers' sensitivity analysis." *BMC bioinformatics* 21, no. 1 (2020): 1-18.