

# Heart disease prediction using neural network

Maisha Jannat (2018-1-60-026)

Rubaiat Hasnain Khan (2017-2-60-048)

Syed Sharful Islam Sifat (2017-2-60-019)

Moonwar Al Wardiful (2017-2-60-107)

## Contents

<b>1</b>	<b>Abstract</b>	<b>2</b>
<b>2</b>	<b>Introduction</b>	<b>2</b>
2.1	objectives . . . . .	3
2.2	motivation . . . . .	3
2.3	existing . . . . .	3
2.4	necessity . . . . .	4
<b>3</b>	<b>Methodology</b>	<b>4</b>
<b>4</b>	<b>Implementation</b>	<b>5</b>
4.0.1	Data collection . . . . .	5
4.1	Data processing . . . . .	6
4.2	Model development . . . . .	6
4.2.1	Results . . . . .	8
<b>5</b>	<b>Conclusions</b>	<b>8</b>
5.1	Challenges . . . . .	8
5.2	Limitations . . . . .	9
5.3	Future work . . . . .	9
<b>6</b>	<b>References</b>	<b>10</b>

# 1 Abstract

This paper presents an effective neural network with convolutions layers for classifying clinical data that is highly class-imbalanced. Machine learning models were employed to produce effective conclusions in the heart disease prediction because the healthcare industry contains a significant amount of psychiatric data. Using machine learning approaches, it is possible to reliably classify people as healthy or unhealthy. The goal of this article is to use neural networks to forecast cardiac disease. The model will be investigated and improved in the future in order to become a robust and end-to-end dependable research instrument. In the case of ECG report data, the data format was changed to improve the action of the convolutions neural network utilized in our study, and in the risk prediction module, we chose attributes for training and deploying the multi-layered neural network we constructed. We discovered that group algorithms for heart disease prediction are more successful than baseline algorithms.

# 2 Introduction

Any ailment that affects the heart is referred to as heart disease. Heart disease is one of the most frequent diseases today, and early detection of the condition is critical for many health care providers in order to protect their patients and save lives. Predicting cardiac disease may be done using a variety of methods. Each property in this document has a predetermined value. The construction, assessment, and optimization of increasing-depth NN architectures for heart disease detection is our key contribution. Machine learning is a technique for teaching a machine to learn and predict when it is presented with fresh data. Machine learning has a plethora of applications, including recommended systems, medical diagnosis, and so on. Predictive analytic is used in customer relationship management, healthcare, and a variety of other industries. There are enough models in predictive modeling, such as Naive bayes, logistic regression, neural networks, and support vector machines, to name a few. The input layer of a neural network contains a large number of nodes. These input layer nodes are connected to the hidden layer nodes. The weights are assigned to each input. The data is passed from the input nodes in the network to the hidden layer nodes, which execute specific tasks or computations before sending the processed data to the output node.

The node that delivers the final result is in the output layer.

## **2.1 objectives**

The goal of this study is to use neural networks to predict heart disease. Patients will be categorized into various degrees of coronary artery disease based on factors such as blood pressure, cholesterol levels, heart rate, and other characteristics. A total of 1026 patients will be used in this experiment.

## **2.2 motivation**

The goal of this study is to use neural networks to predict heart disease. Patients will be categorised into various degrees of heart disease based on factors such as blood pressure, cholesterol levels, heart rate, and other characteristics. The SCI Machine Learning Repository will provide a data set of patients for this study. The risk variables acquired via logistic regression are fed into a neural network, which is then utilized to determine whether or not the person has heart disease. In order to predict cardiac disease, the merger of logistic regression and neural networks is used.

## **2.3 existing**

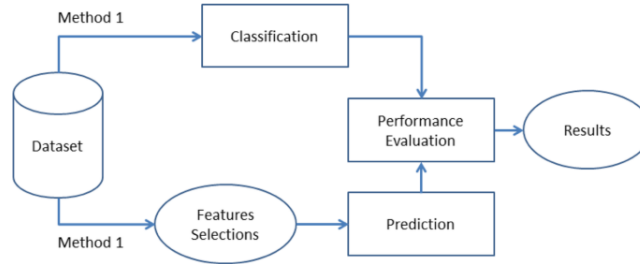
V. Sree Hari Rao et al. [1] used the in-built imputation technique and particle swarm optimization to identify risk variables for Coronary heart disease (CHD) or atherosclerosis. Physical inactivity has been discovered to be a risk factor for CHD. The decision rules are used to forecast heart disease risk variables. For the prediction of cardiac disease, Carlos Ordonez et al. [2] employed association rules. When these association rules are applied to a medical dataset, a large number of rules are generated that are irrelevant. In order to determine which regulations are in effect. Using search constraints, which examines the association rules in the training dataset and then validates on the test set, actually crucial for predicting heart disease is determined. Sikander Singh Khurl et al., [3] used decision trees and an apriori algorithm to identify risk variables for heart attack, also known as myocardial infarction. Chest discomfort, diabetes, smoking, gender and physical inactivity, age, lipids, cholesterol, triglyceride, and blood pressure are the risk variables found as effective in the identification of heart attack using these approaches.

## 2.4 necessity

Heart disease is one of the most debilitating and dangerous chronic diseases that is on the rise in both industrialized and developing nations, and it is the leading cause of mortality. This harm can be significantly mitigated if the patient is recognized early and receives appropriate therapy. The system would primarily be designed to reduce the death rate associated with cardiac disorders while also increasing the rate of effective diagnosis. Building an intelligent system that can forecast sickness based on risk factors, saving money and time on medical tests and check-ups while also assuring that the patient can monitor his own health and plan preventative measures and therapy at the earliest stages of disease. As a result of the neural network system, medical professionals' errors, such as misdiagnosing, would be removed, and a prompt preventative plan may destroy the disease at its core.

## 3 Methodology

In practice, neural networks are used to deliver very accurate outcomes. The proposed technique improves the accuracy of classification. There are two parts to the data set: testing data and training



data. The neural network was given the training data set. Neural networks are a collection of algorithms for recognizing patterns. Activation functions make up the layers of the neural network. The input layer provides the network with the training features. The features are assigned to a hidden layer, which does the actual processing using a weighted connection. The network's output layer is linked to the hidden layer. The goal of the predictive model was to generate hypotheses using deep learning models. The relationship between data that may be tested by gathering data and making observations is referred to as a hypothesis. By minimizing the error

in the training cases, we may create the hypothesis. There are two classes in the binary classification of heart disease. The positive class is one, and the negative class is the other. There are also  $t$  occurrences in it.  $E_a$  represents the expected values. The expected value of two independent value based on null hypothesis can be calculated as

$$E_a = (a+b) \frac{a+b}{t}$$

$$\chi^2 = \frac{1}{d} \sum_{k=1}^n \frac{(O_K - E_K)^2}{E_K}$$

After feature ranking, which is denoted by  $n$ , a threshold for the number of features must be determined. Exhaustive search is used to find the optimum number of features from a subset of features with  $n=1$ . NN receives a subset of the features. The NN performance is measured using grid search. After saving the result of the first subset, a new subset with  $n=2$  is used to find the best feature, which is then applied to NN and the result is saved. These steps are repeated until all of the characteristics have been added to a subgroup of features. The optimal outcome is the subgroup of features that produces the best performance result.

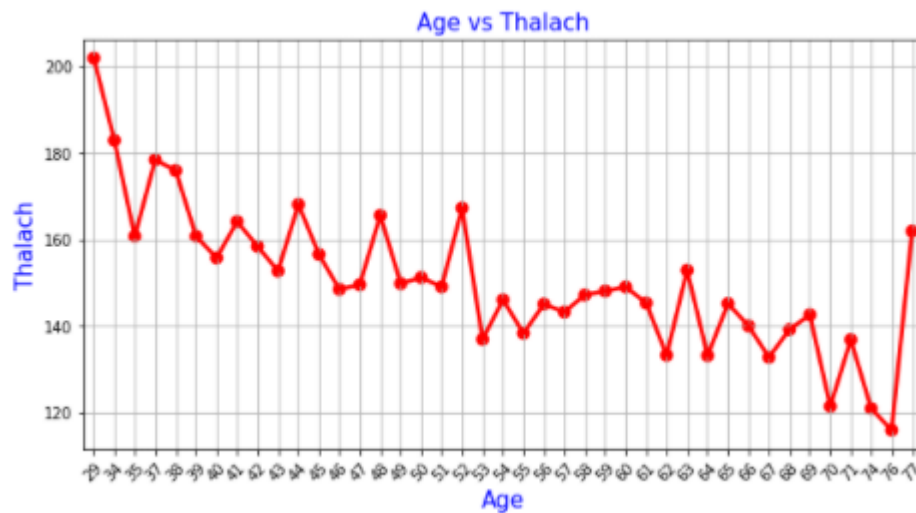
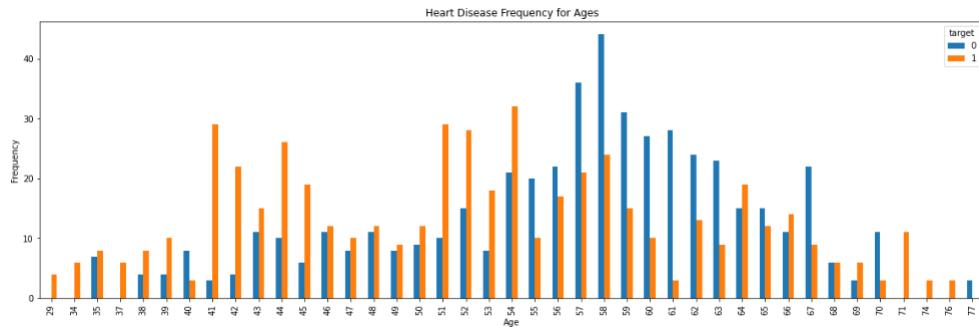
## 4 Implementation

In this section, we used a binary classification model based on visualizing the data. And there are some visual graphs and based on this graph we made our model and method

Then we also visualize

### 4.0.1 Data collection

This data set provides patient data from several locations throughout the world on heart disease diagnosis. There are 76 attributes, including age, sex, resting blood pressure, cholesterol levels, echo cardiogram data, exercise habits, and many others. To data, all published studies using this data focus on a subset of 14 attributes - so we will do the same.



## 4.1 Data processing

Here there were many null fields in the data set, in those cases, we used some mean or average values. Since we used a binary classification problem, here we prepare data as when the data is greater than 0 then it keeps the value 0 and when the data is less than zero then it keeps the values zero.

## 4.2 Model development

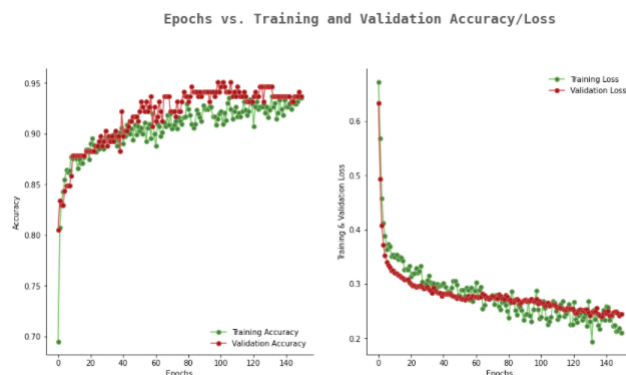
Here we used anaconda and as a language we used python. As an environment/IDE, we used jupyter-notebook, Which is a very good platform for python development and Data science. Otherwise, we used some libraries like NumPy, Pandas, Sklearn, Matplotlib, TensorFlow, Keras, and seaborn.

This is a model based on a neural network. In this model, there are multiple hidden layers, input layers, and output layers. Here at the first step, there are 16 hidden layers and used activation function 'relu' and dropout 25 percent. In the second step, there are hidden layers and used activation function 'relu' and dropout 25 percent. At the last stage it gives output and here dense layer 1 and activation sigmoid. We used a learning rate of .0001.

Layer (type)	Output Shape	Param #
dense_6 (Dense)	(None, 16)	224
dropout_4 (Dropout)	(None, 16)	0
dense_7 (Dense)	(None, 8)	136
dropout_5 (Dropout)	(None, 8)	0
dense_8 (Dense)	(None, 1)	9
Total params: 369		
Trainable params: 369		
Non-trainable params: 0		
None		

This figure describes our full model with different functions and models.

While we fit our model then we used batch size as 10 and epoch 150 which gives us the more accurate result. After building this model we found some graphs with accuracy and loss. Here is the graph and we can analyze our model by this graph



### 4.2.1 Results

Here we used binary classification that's why it gives only 0 and 1 values. That's mean based on the model it gives when Yes or No. It will not give other values. After training the model this data set test with testing data set and it gives about 93 percent accuracy and here is the detailed result of our model

---

Results for Binary Model					
0.9365853658536586					
	precision	recall	f1-score	support	
0	0.94	0.93	0.93	100	
1	0.93	0.94	0.94	105	
accuracy			0.94	205	
macro avg	0.94	0.94	0.94	205	
weighted avg	0.94	0.94	0.94	205	

## 5 Conclusions

With the rising number of deaths due to heart disease, it is becoming increasingly important to build a system that can effectively and accurately forecast heart disease. The goal of the research was to discover the most effective machine learning system for detecting cardiac problems. The accuracy of Neural Network binary classification methods for predicting heart disease is compared in this study. deliver improved outcomes and aid health professionals in accurately and efficiently identifying heart disease

### 5.1 Challenges

Heart Disease refers to the realm of health and medical science, which is distinct from the IT industry, whereas ML refers to an emerging field of AI that is a new technology that is opening up new possibilities. Because ML efforts are always technology-driven, getting to the real product is difficult. These are some of the high-level problems that will be confronted as the project develops. The followings are addressing the challenges while developing this project as listed below: 1. Finding details about the Heart Disease's characteristics is challenging. As a result, determining the Heart Disease's proper



property becomes difficult. 2. The key goal of this project will be to build a classifier model that can predict whether or not a patient has Heart Disease. This implies that the main challenge will be to train the classifier model such that the system can accurately determine whether or not the patient has Heart Disease. 3. Machine learning is a new technology that offers a solution to a current issue. Every significant corporation, such as Google, Facebook, and Microsoft, is working on machine learning projects in their own environments, such as Tensorflow (launched in 2017), PyTorch (published in 2017), and Microsoft Cognitive Toolkits, respectively. The study is still in its early phases. It may not be ready for production or may be close to being ready. As a result, the issue will arise throughout the implementation of this project.

## **5.2 Limitations**

Due to domain and technical knowledge, this model is allocated for the unique purpose of Heart Disease. Due to the limited time available to complete this project, it will only perform the following tasks:

1. Heart Disease Prediction System is only capable of detecting the presence of HD in patients. We will improve the model in the future to forecast specific types of HD.
2. Due to the difficulty of gathering cardiac patient data, the HD prediction model can only be trained with 1026 data of HD patients; however, in the future, we will collect bigger data sets and train models with high accuracy.
3. Heart Disease Prediction System can only predict heart disease in the long run. We will integrate a healthcare system in the future.
4. Due to a time constraint, we were unable to investigate additional aspects of the system, such as the system's ability to forecast different types of heart disease.

## **5.3 Future work**

More research will be done to include new analytical models of other diagnostic tests for the identification of different forms of heart disease. The tools were intended exclusively for the aim of study, not to make diagnostic judgements. In the future, the Heart Disease Prediction System will be upgraded to anticipate a specific form of Heart Disease, such as heart attacks, CVD, CAD, and so on. Hospitals, clinics, smart phones, smart wear,

hospital/police emergency systems, and interaction with fitness mobile applications are all possibilities for the Heart Disease Prediction System. This algorithm will be integrated into hospital and clinic systems to predict heart illness. We'll also incorporate this model into a smart phone app that will allow us to quickly screen for heart disease. In the event of an emergency, we will connect smart wear to hospital and police emergency systems to save the patient's life.

## 6 References

1. <https://www.hindawi.com/journals/cin/2021/8387680/>
2. <https://medium.com/ai-techsystems/heart-disease-prediction-using-neural-networks-a9486ee3cb81>
3. <https://turcomat.org/index.php/turkbilmat/article/download/8438/6618/15156>
4. <https://www.ijert.org/neural-network-based-heart-disease-prediction>
5. Neural Network based Heart Disease Prediction – IJERT
6. Prediction of heart disease and classifiers' sensitivity analysis — BMC Bioinformatics — Full Text (biomedcentral.com)