

BUILDING PREDICTION SYSTEM FOR HEART DISEASES USING DEEP NEURAL NETWORK

NUR A'ISYAH BINTI AHMAD IRFAN¹, NUR ATIQA BINTI HASBULLAH², MOHAMAD ARIF DANIEL
BIN MUHAMADDUN³, ASSOC. PROF. TS. DR. AMELIA RITAHANI BINTI ISMAIL⁴

^{1,2,3,4} *Department of Computer Science, Kulliyah of Information and Communication Technology,
International Islamic University Malaysia*

1. Introduction

A variety of diseases that attack your heart is referred to as heart disease. Now, the heart is the most vital organ in our bodies. It circulates blood throughout our bodies, ensuring that oxygen reaches all parts of our bodies. A life-threatening condition may occur if something is wrong with the heart. That demonstrates how vital the organ heart is to us. Let us now discuss heart disease and why it is such a concern in our society. For your information, heart diseases were announced as the top killer in Malaysia for the year 2021 (Department of Statistics Malaysia, 2021). Malaysia is rich with varieties of foods from all regions around the world. Most foods in our country contain a sheer amount of oil and herbs used in it making them heavy in cholesterol and calories. It is undeniable that although our bodies require cholesterol to produce healthy cells, high cholesterol levels can increase the risk of heart disease. A high amount of cholesterol intake in the blood will lead to fatty blockage in the blood vessels, and those deposits eventually accumulate to the point that blood flow through your arteries becomes restricted. This condition will then cause a heart attack or heart disease. There are other indicators that can be used to determine whether someone is at risk of heart disease, which will be explored further in our research.

Patients' information and data are becoming more publicly available for use in

research on this topic every day. It makes it easier to get the dataset for us to do our research project. To be honest, applying machine learning to the medical field is no longer unusual. More researchers are collaborating with medical practitioners to employ current technologies to aid in identifying or predicting a patient's health status based on their diagnosis results. Thus, as a result of this matter, we decided to propose a deep neural network (DNN) based prediction system to predict heart diseases. The machine learning model will be used in our proposed system to predict the likelihood of patients developing heart disease. We will also compare the results after testing the system to see how accurate it is and why DNN would be the ideal machine learning model for this research topic.

2. Dataset Description

We used a heart dataset taken from Kaggle's website to make predictions of heart disease. This dataset contains 918 observations with 12 attributes. A brief description of the dataset is shown in Table 1 with the last attribute, 'HeartDisease' that displays whether the patient is in a normal condition, 0, or has heart disease, 1. Furthermore, in this dataset, 54.45% of patients in the age range between 37 and 76 years old have high chance of getting heart disease. The first attribute age contains the

age of the patient with 28 years old as the youngest and 77 years old as the oldest. Sex attribute classifies the gender of the patient based on male or female. ChestPain attribute is the chest pain type consisting of typical angina (TA), atypical angina (ATA), non-anginal pain (NPA), and asymptotic (ASY) chest pain. The restingBP is a blood pressure measurement taken while the patient is at rest with range values of 0 to 200 mm Hg. The cholesterol attribute is the amount of cholesterol taken from the patient using a BMI sensor. Next is fastingBS which is measured after the patient has done an overnight fast. RestingECG is the result of the test that measures the electrical activity of the patient's heart where the results consist of normal, ST-T wave abnormality, ST, and

showing probable or definite left ventricular hypertrophy by Estes' criteria, LVH. The maxHR attribute is the patient's maximum heart rate that was calculated by subtracting their age from 220 and is also the average of the maximum number of the heartbeat per minute during exercise. ExerciseAngina displays the information in which the patient got angina when doing exercises with the answer either yes or no. The oldpeak is the result of ST depression; a sign of myocardial ischemia, of which coronary insufficiency is a major cause that is measured by using an overlying electrode. The slope of the ST depression is shown in the ST_Slope attribute where the results are either up: upsloping, flat, or down: downsloping.

No.	Attributes	Description	Range Values
1.	Age	Age of the patient	28 - 77
2.	Sex	Gender of the patient	M of F
3.	ChestPain	Chest pain type	TA, ATA, NPA, ASY
4.	RestingBP	Resting blood pressure (mm Hg)	0 - 200
5.	Cholesterol	Cholesterol fetched via BMI sensor (mg/dl)	0 - 603
6.	FastingBS	Fasting Blood Sugar > 120 mg/dl	1 or 0
7.	RestingECG	Resting Electrocardiographic Results	Normal, ST, LVH
8.	MaxHR	Maximum heart rate achieved	60 - 202
9.	ExerciseAngina	Exercise induced angina	Yes or No
10.	Oldpeak	Previous peak	-2.6 - 6.2
11.	ST_Slope	Slope of the peak exercise ST segment	Up, Flat, Down
12.	HeartDisease	Presence or absence of heart disease	0 or 1

Table 1: Heart Dataset Description

2.1 Data Pre-processing

Data preprocessing is a step to make the dataset more presentable and easier to visualize that includes removing unwanted columns or null values. Since this dataset has no null values, we only apply another preprocessing step which is the standardization method using

StandardScaler(). Standardization is very helpful in creating consistent data as it involves the process of translating the data into a range between -1 and 1. As we can see in Table 1, some of the columns have categorical values. Hence, the conversion from categorical to numerical using scikit learn is done in order to standardize the

datatype of all attributes. The data standardization is using the following equation where X represents the column inputs and X_{ST} represents the transformed values:

$$X_{ST} = \frac{X - \text{Mean of } X}{\text{Standard Deviation of } X}$$

3. Deep Learning Model

Deep learning or deep structured learning is part of a broader family of machine learning methods based on artificial neural networks with representation learning (Wikipedia). There are many deep learning architectures and, in this research, we will try to classify heart disease using the Deep Neural Network (DNN). Deep Neural Network is an Artificial Neural Network (ANN) but with multiple layers between the input and output layers that allow models to become more efficient in learning complex features and make a better prediction.

3.1 Hyperparameter Tuning

A hyperparameter is a parameter of the model that is set before the start of the learning process. Different machine learning models have different hyperparameters. We will use the Grid Search method that takes in as many hyperparameters as we would like, and tries every single possible combination of the hyperparameters. This method will be used with Adam optimizer that will return us the best hyperparameters with the best accuracy. The first hyperparameter is the batch size and number of epochs. The batch size is the number of patterns shown to the network before the weights are updated whilst the number of epochs is the number of times that the entire training dataset is

shown to the network during training. We are evaluating the batch sizes of 10, 20 and 40, the number of epochs with the values of 10, 50 and 100 and the grid search returns 40 as the batch size and 10 as the epochs that have achieved the highest accuracy. Next, we will tune the numbers of neurons in each of the two hidden layers and the optimal number of neurons in each hidden layer is 16 and 8 respectively.

4. Experimental Setup

4.1 Flowchart of the steps in the experiment.

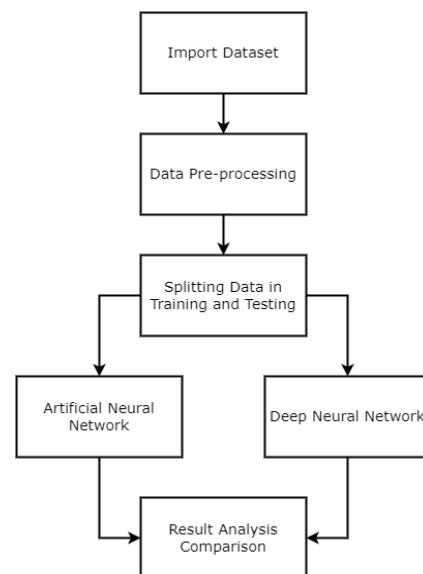


Figure 1: Experimental Flowchart

5. Analysis of Result

In figure 1, we splitted the preprocessed dataset into 80% training data and 20% testing data. We then apply the ANN and DNN model to make predictions and the analysis of the results are further discussed in Table 2.

		Predicted Values	
		0	1
Actual Values	0	True Negative (TN)	False Positive (FP)
	1	False Negative (FN)	True Positive (TP)

Figure 2: Confusion Matrix for Binary Classification

5.1 Confusion Matrix

In order to measure the performance of this supervised machine learning task with multiple classes as output, a performance metric tool called confusion matrix as can be seen in Figure 2 is used to measure the model's performance. As this dataset in our research employs the values 0 and 1 for classification, if the predicted value is 0 as the actual value should be then it will be classified as True Negative (TN). However, if the value displayed is 1 when the actual value should be 0, it will fall into the False Positive (FP) category. For the actual value of 1, if the predicted value turns out to be 0, it will become False Negative (FN), whereas if the value is 1 as it should be then the category is True Positive (TP). The formula for calculating these values is given below:

5.1.1 Accuracy

Accuracy (ACC) will reveal just how many of our predictions were completely right. For us to know that the model is operating successfully, accuracy should be at maximum.

$$ACC = \frac{(TP + TN)}{(TP + TN + FP + FN)} \times 100$$

5.1.2 Precision

Precision or Positive Predictive Value (PPV) is the accuracy of a model's positive prediction.

$$PPV = \frac{TP}{(TP + FP)} \times 100$$

5.1.3 Sensitivity

Sensitivity or Recall or True Positive Rate (TPR) is computed by dividing the number of positives by the number of positive predictions made. 1.0 is the best value for sensitivity while 0.0 is the worst of it.

$$TPR = \frac{TP}{(TP + FN)} \times 100$$

5.1.4 Specificity

Specificity or also known as True Negative Rate (TNR) is computed by dividing the total of negatives by the number of valid negative predictions. Specificity, like sensitivity, has the greatest value of 1.0 and the worst value of 0.0.

$$TNR = \frac{TN}{(FP + TN)} \times 100$$

5.1.5 F1 - Score

The arithmetic average of precision and recall is called F1-Score. It considers both false positives and false negatives. As a result, it functions well on a skewed dataset.

$$F1 - Score = \frac{(2 \times TP)}{(2 \times TP + FP + FN)} \times 100$$

5.2 Performance Comparison

To prove that DNN that we have applied has better performance, we compared it with another prediction method which is Artificial Neural Network (ANN) by using

evaluation metrics. In the following, we are discussing the performance comparison obtained by using ANN and DNN in a table below.

Evaluation Metrics	Artificial Neural Network (ANN)	Deep Neural Network (DNN)
Accuracy	82%	85%
Precision	86%	86%
Sensitivity	82%	82%
Specificity	82%	88%
F1-Score	84%	87%

Table 2: Comparison of the evaluation between Artificial Neural Network and Deep Neural Network

Table 2 displays the performance comparison and it shows that DNN has a higher accuracy of making correct predictions by 85%. Next, the precision or the positive cases that were correctly identified shows no difference between ANN and DNN with the percentage of 86%. Both models also have the same performance in correctly identifying the actual positive cases or also known as sensitivity but DNN is better in predicting the actual negative cases as the specificity is 88% higher than specificity of ANN. Due to the lower specificity in ANN, the F1-Score shows that DNN is a better classifier as it reaches 87% compared to ANN which is only 84%.

6. Conclusion

For this research project, we proposed a prediction system for heart disease using the deep neural network (DNN) algorithm. A DNN is a model that has a multilayer of

hidden layers compared to an artificial neural network (ANN) that only consists of one hidden layer (Johnson, 2020). This is why the DNN model outperforms the ANN model in terms of accuracy. Our comparison analysis using the evaluation metrics of comparing the accuracy, precision, sensitivity, specificity and the F1-Score, also reveals that the KNN model which belongs to DNN performs considerably better than the ANN model. We trust that by using the DNN technology, we will be able to anticipate cardiac disease at an earlier stage and save more lives in the near future. This will be a significant milestone in the medical field, particularly in the field of heart disease.

References

- [1] aniruddha. "Feature Scaling | Standardization Vs Normalization." 3 April 2020. Analytics Vidhya. <https://www.analyticsvidhya.com/b>

- log/2020/04/feature-scaling-machine-learning-normalization-standardization/
- [2] Bharti, Rohit, et al. "Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning." Computational Intelligence and Neuroscience 2021 (2021): 11.
https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=401&bul_id=R3VrRUhwSXZDN2k4SGN6akRhTStwQT09&menu_id=L0pheU43NWJwRWVSZklWdzQ4TlhUUT09
- [3] Johnson, Jonathan. "What's a Deep Neural Network? Deep Nets Explained – BMC Software | Blogs." 27 July 2020.
<https://www.bmc.com/blogs/deep-neural-network/>
- [4] Lakshmanamoorthy, Rajkumar. "Python Code for Evaluation Metrics in ML/AI for Classification Problems." 7 March 2021. Analytics India Magazine.
<https://analyticsindiamag.com/evaluation-metrics-in-ml-ai-for-classification-problems-wpython-code/>
- [5] Narkhede, Sarang. "Understanding Confusion Matrix | by Sarang Narkhede." 9 May 2018.
<https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62>
- [6] Brownlee, Jason. "How to Grid Search Hyperparameters for Deep Learning Models in Python with Keras." Machine Learning Mastery, machinelearningmastery.com/grid-search-hyperparameters-deep-learning-models-python-keras/.