

HEART DISEASE PREDICTION USING MACHINE LEARNING

ABSTRACT

The cardiovascular system is made up of the heart and blood vessels. Cardiovascular disease (CVD) is defined as any serious, abnormal condition of the heart or blood vessels(arteries, veins). Cardiovascular disease includes coronary heart disease (CHD), stroke, peripheral vascular disease, congenital heart disease, endocarditis, and many other conditions. Many cardiovascular diseases are preventable. Cardiovascular disease is one of the major causes of mortality in today's world. Cardiovascular disease is a type of disease that affects the heart or blood vessels. In cardiovascular disease, the heart is unable to push the required amount of blood to other parts of the body. Accurate and on time diagnosis of heart disease is important for heart failure prevention and treatment. The risk of certain heart diseases may be increased by smoking, high blood pressure, high cholesterol, unhealthy diet and obesity.

Cardiovascular disease has significantly increased over the last decade and has become the leading cause of death for people in most countries around the world. The symptoms of heart disease include shortness of breath, weakness of physical body, swollen feet, and fatigue. The American heart association reported that nearly half of American adults are affected by cardiovascular disease (CVD), equating to nearly 121.5 million adults. In Korea, heart disease is among the top three leading causes of death and contributed to nearly 45% of total deaths in 2018. The total number of deaths globally is projected to increase to around 22 million in 2030. As a result, early detection of cardiovascular disease can reduce the causes of mortality.

INTRODUCTION

GENERAL DESCRIPTION :

The term “heart disease” refers to several types of heart conditions. The most common type of heart disease in the United States is “coronary artery disease (CAD)”, which affects the blood flow to the heart. Decreased blood flow can cause a heart attack. Heart disease refers to any condition affecting the heart. There are several different types of heart disease, and they affect the heart and blood vessels in different ways. Heart diseases include: Blood vessel disease, such as coronary artery disease.

SYMPTOMS :

Sometimes heart disease may be “silent” and not diagnosed until a person experiences signs or symptoms of a heart attack, heart failure, or an arrhythmia. When these events happen, symptoms may include., Heart attack: Chest pain or discomfort, upper back or neck pain, indigestion, heartburn, nausea or vomiting, extreme fatigue, upper body discomfort, dizziness, and shortness of breath. Arrhythmia: Fluttering feelings in the chest (palpitations). Heart failure: Shortness of breath, fatigue, or swelling of the feet, ankles, legs, abdomen, or neck veins.

Risk factors for heart disease :

High blood pressure, high blood cholesterol, and smoking are key risk factors for heart disease. About half of people in the United States (47%) have at least one of these three risk factors.² Several other medical conditions and lifestyle choices can also put people at a higher risk for heart disease, including Diabetes, Overweight and obesity, unhealthy diet, Physical inactivity, Excessive alcohol use.

PROBLEM IDENTIFICATION

The main objective is to predict whether a person has heart disease or not .In our proposed model we are using which is used to diagnose the subject's heart disease status earlier. we are using Artificial Neural Network(ANN) outlier detection and removal and data balancing techniques which will remove outliers from the datasets and balance the data before applying the machine learning techniques. We are using Artificial Neural Network(ANN) technique which is used to predict the early stages of cardiovascular disease.

PROPOSED METHODOLOGY

In proposed system mainly focuses on drawbacks of the existing algorithms. Artificial Neural Networks (ANN) algorithms is used to predict the early stage of heart disease with high accuracy.

DEEP LEARNING :

Artificial Neural Network (ANN)

Recurrent Neural Network (RNN)

Convolutional Neural Network (CNN)

Artificial Neural Network

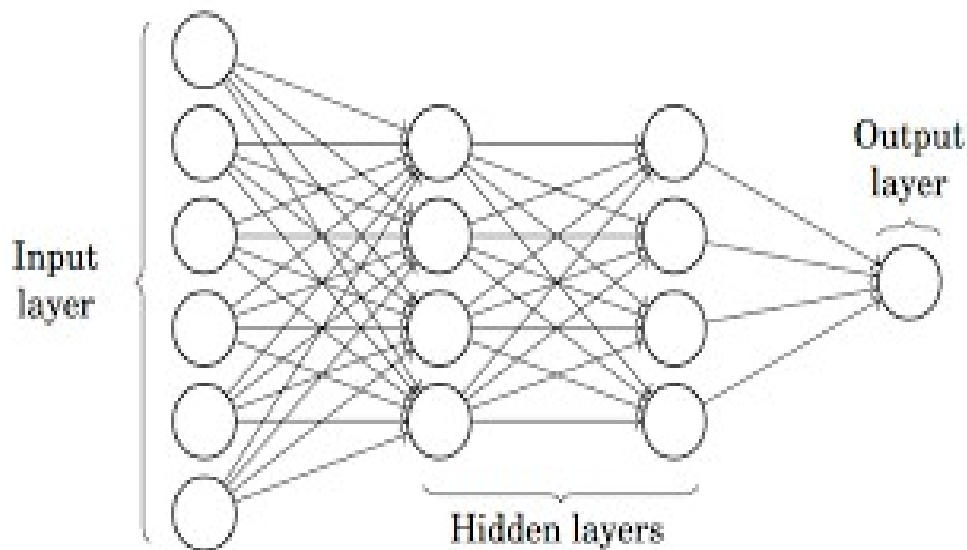
Neural Networks is a computational learning system that uses a network of functions to understand and translate a data input of one form into a desired output. The concept of the artificial neural network was inspired by human biology and the way neurons of the human brain function together to understand inputs from human senses. Neural Networks are a set of algorithms that tries to recognize the patterns, relationships, and information from the data through the process which is inspired by and works like the human brain/biology.

ANNs are composed of artificial neurons which are conceptually derived from biological neurons. Each artificial neuron has inputs and produces a single output which can be sent to multiple other neurons. The inputs can be the feature values of a sample of external data, such as images or documents, or they can be the outputs of other neurons. The outputs of the final output neurons of the neural net accomplish the task, such as recognizing an object in an image.

To find the output of the neuron, first we must take the weighted sum of all the inputs, weighted by the weights of the connections from the inputs to the neuron. We add a *bias* term to this sum. This weighted sum is sometimes called

the activation. This weighted sum is then passed through a (usually nonlinear) activation function to produce the output. The initial inputs are external data, such as images and documents.

Architecture of Neural Network



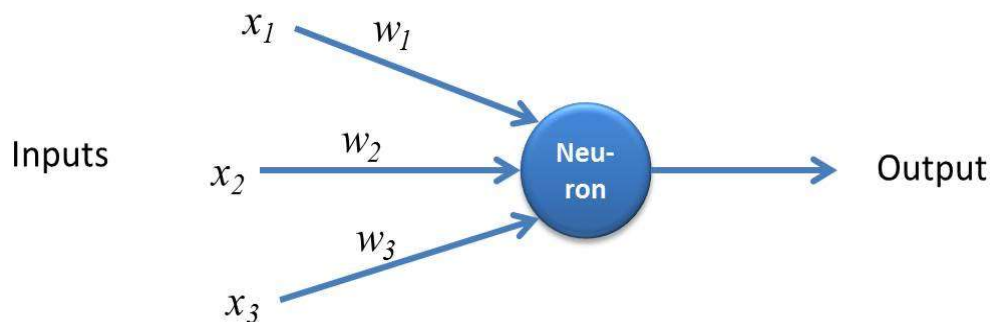
Input Layer: Input layer known as Input nodes are the inputs from the outside world is provided to the model to learn and derive conclusions.

Hidden Layer: Hidden layer is the set of neurons where all the computations are performed on the input data. There can be any number of hidden layers in a neural network.

Output layer: The output layer is the output of the model derived from all the computations performed. There can be single or multiple nodes in the output layer.

WORKING OF THE ANN

Perceptron is a simple form of Neural Network and consists of a single layer where all the mathematical computations are performed.



In first step, Input units are passed data is passed with some weights attached to it to the hidden layer. We can have any number of hidden layers. In the above image inputs $x_1, x_2, x_3, \dots, x_n$ is passed. Each hidden layer consists of neurons. All the inputs are connected to each neuron.

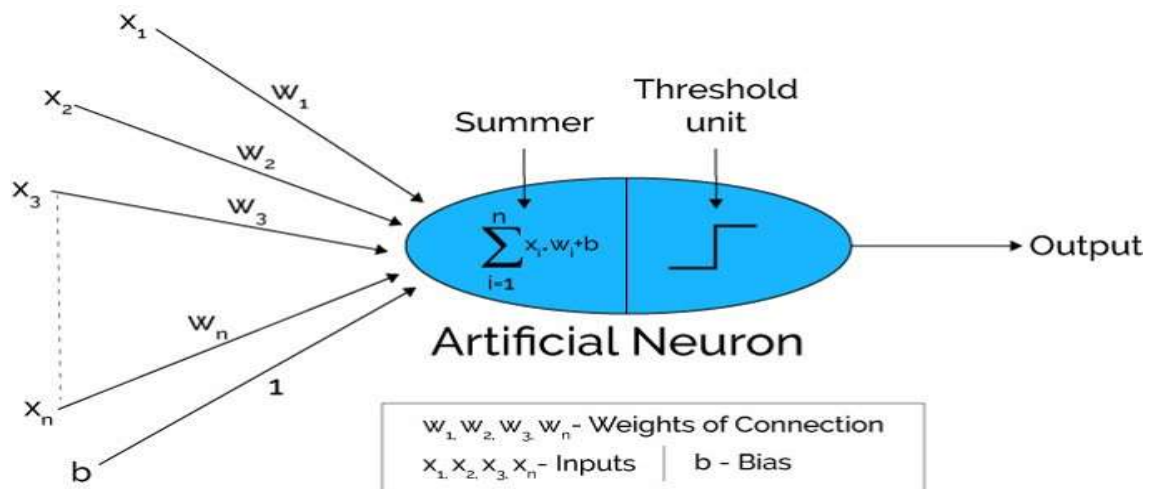
Computation performed in hidden layers are done in two steps which are as follows:

First of all, all the inputs are multiplied by their weights. Weight is the coefficient of each variable. It shows the strength of the particular input.

After assigning the weights, a bias variable is added. Bias is a constant that helps the model to fit in the best way possible.

$$Z_1 = W_1 * In_1 + W_2 * In_2 + W_3 * In_3 + W_4 * In_4 + W_5 * In_5 + b$$

W_1, W_2, W_3, W_4, W_5 are the weights assigned to the inputs $In_1, In_2, In_3, In_4, In_5$, and b is the bias. After getting the predictions from the output layer, the error is calculated i.e the difference between the actual and the predicted output.



The activation function is applied to the linear equation Z_1 . The activation function is a nonlinear transformation that is applied to the input before sending it to the next layer of neurons. The importance of the activation function is to inculcate nonlinearity in the model. If the error is large, then the steps are taken to minimize the error and for the same purpose, Back Propagation is performed.

