The Predictability of the Likeliness of Having a Heart Attack

Prisha Mathur

Inspirit, IA, USA

ABSTRACT

Heart Disease is the leading cause of deaths in the entire world. They are happening to more than 17.9 million people each year, and each year is a higher increase for heart diseases. With using an accurate model to predict the likelihood of having a heart attack, you can get the help needed, and prevent further healthcare problems. This paper uses the logistic regression machine learning method to show the Predictability of the Likeliness of a Heart Attack. With AI models increasing in their accuracy with predictions, models are becoming safer. The methodology used for this dataset is to utilize the user's information, and uses specific equations of linear models, neural networks, and decision tree to find the weight of the attributes, and evaluate the performance on the tested. For example, the neural network accuracy is 0.81 There are 14 attributes tested in this dataset. The model will use the user's information to predict the likeliness of the heart attack. The methodology employing this is using logistic regression. First the model will predict the likeness of a heart attack based on the columns provided. It will use the neural network, decision tree, and linear model to predict, find out the weight of the attributes, and tell the model threshold predictability from 0 or 1.The significance in accurate prediction is so there isn't a lack of trust with the patient and this dataset. If the data is not accurate, it can lead to misjudgement of the patient's information.

Introduction

Cardiovascular diseases are happening to more than 17.9 million people each year. According to About 695,000 people in the U.S. died from cardiovascular disease, which is 1 in every 5 people. They are not only happening to elderly people, but a 30% increase in people from the ages 25-44, during the years of the Global Pandemic. The pandemic has had a setback on a lot of people with different factors, especially healthwise. Many individuals were busy on their phones, staying at home, and not getting enough exercise that is necessary for us humans. We have been accustomed to staying at home, and not really getting enough sunshine and exercise in our everyday life. Ever since Covid-19, there has been a significant increase in cardiovascular disease from 2020. It is very important to make sure your body is healthy, and be aware of what you can do to help. With the role of AI in predictive healthcare, you can find out the predictability of your likeliness of having a heart attack. Some of the main risk factors include poor diet, tobacco smoking, physical inactivity, overweight or obesity, alcohol intake, raised blood pressure, raised blood glucose, and raised blood lipids.

In addition, Lipoprotein A multiplies your chances of getting a heart attack if it is high. It is one of the main factors of causing a heart attack, as well as having a sedentary lifestyle. Heart Attacks and strokes are caused by a blockage that prevents your blood from passing to the heart or brain. Symptoms of having a heart attack are chest pain or discomfort; Shortness of breath; Pain or discomfort in the jaw, neck, back, arm, or shoulder; Feeling nauseous, light-headed, or tired. However, age is one of the main factors towards cardiovascular disease. It promotes oxidative stress, and atherosclerosis, which is a disease from the arteries. Some of the most common types of cardiovascular diseases include coronary heart disease, high blood pressure, cardiac arrest, congestive heart disease, arrhythmia, peripheral artery disease, stroke, and congenital heart disease. Over

three quarters of CVD deaths take place in low- and middle-income countries. They don't have the doctors and necessities they need to care for themselves. That is what is making the cardiovascular diseases increase significantly as poverty is the main reason. The inflation and influx of money being the main problems, lots of people can't care for themselves. In poor countries, there aren't the necessary medical supplies and workers to take care of each individual person.

Dataset

With AI models increasing the accuracy with their predictions, models are becoming safer, and decreasing the number of mortality rates. The dataset I used was from Kaggle. This dataset provided the dataset file, with a graph of all the information. There are 13 features included in this dataset including age, sex, chest pain type (4 values), resting blood pressure, serum cholesterol in mg/dl, fasting blood sugar > 120 mg/dl, resting electrocardiographic results (values 0,1,2), maximum heart rate achieved,exercise induced angina. This dataset is based on a 80-20 split between 80% for the training, and 20% for the testing. There are 303 people total in this dataset. We split the dataset between 80-20, because we don't want all the data to be training, and we want some people to be tested so we can see how accurate the dataset is with using people that haven't been trained.

Link of the dataset: https://www.kaggle.com/datasets/nareshbhat/health-care-data-set-on-heart-attack-possibility?resource=download
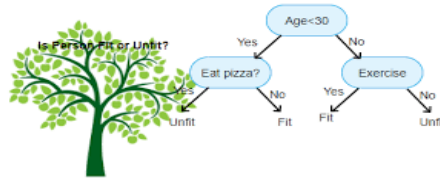
Model

Models are a very important role in AI healthcare. They require a set of algorithms and the data used to train those algorithms to make accurate and beneficial predictions. The purpose of these machine-learning models is to recognize patterns in data. Models have various advantages including it is a safer and cheaper option, as well as beneficial to use for testing before you create a product or system. This model uses 3 models : decision tree, neural network, and the linear model. These models use 13 features : age, sex, chest pain type (4 values: Angina, Heart Attack, Aortic dissection, pericarditis), resting blood pressure, serum cholesterol in mg/dl, fasting blood sugar > 120 mg/dl, resting electrocardiographic results, maximum heart rate achieved,exercise induced angina. AI models are becoming more accurate in their predictions, as we get more data.

Linear model: Every single number is being multiplied by the weights, and then it will predict. (W1 x p1) + (W2 x p2)... if the summation is higher than the threshold, predict 1 otherwise predict 0.
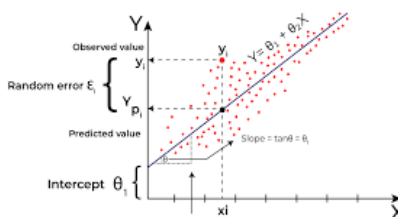
Decision Tree: This model classifies or regresses the data using true or false questions to certain questions. They use a tree-like structure and each brach represents the outcome of a certain value. They can handle both numerical and categorical information.

Neural Network: They are models that are made to recognize non-linear patterns in data. They produce outputs for things including classification, as well as regression. Neural Networks have more parameters which requires more data. With using large datasets, they can improve accuracy, and will continue to be stronger.
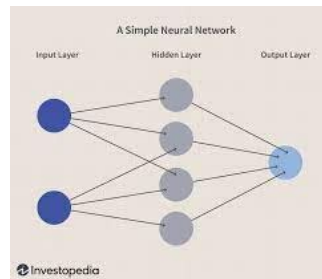
The majority vote in the code ensures that at least 2 of the models have agree with each other. Instead of having these models work independently, it shows how they will work combined. This is why the majority vote is vital in the models because at every image it will predict what the majority says. Each model will display each of the accuracy percentages.With 3 models, we can aggregate the predictions in order to get the final prediction, and the percentage accuracies. With this strategy the accurate prediction is 93%.

Training Accuracies	Testing Accuracies
Decision Tree : 1.0	Decision Tree : 0.72
Neural Network : 0.87	Neural Network : 0.81
Linear Model : 0.88	Linear Model : 0.77
Majority Strategy : 91%	Majority Strategy : 93%

Discussion

If someone were to continue this work, they should focus on finding more people to do this research on. With more people they can broaden the evidence with the numbers. Furthermore, an improved dataset should add lipoprotein A, because that is one of the main risk factors in having a heart attack. In addition, my long term goal for this project is to open up new options in healthcare. Instead of going to a hospital or a doctor to check if you are all right, AI is the future in healthcare, and can predict if you are all right or not. AI healthcare, especially heart attacks, inspired me because of my dad. He was a 45 year old male with a healthy lifestyle. However, this heart attack came so suddenly and was almost fatal. If we were more aware of AI models in healthcare, we would have been more prepared for this to prevent it from happening. This model's purpose is

so no one else has to feel as unprepared as we did. This can be to prevent anyone's family from having to go through this horrible experience.

Conclusion

In conclusion, this paper focuses on the predictability of heart attacks using 3 models. The neural network accuracy and linear model accuracy were the closest, with them being 85% and 83%. All the models had an accuracy of 75% and higher, but can continue to be improved with more user's information.

Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

References

1. https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

2. https://www.bhf.org.uk/informationsupport/conditions/cardiovascular-heart-disease

3. https://www.nm.org/healthbeat/healthy-tips/heart-facts-infographic

4. https://aphasia.org/aphasia-resources/stroke/?gad_source=1&gclid=Cj0KCQjw-_mvBhDwARIsAA-Q0Q6aujYQZ4CEVbssA0pBF-C_63MEb42qK_rKnVCyOJ6JkABPpzOvXHMaAucBEALw_wcB

5. https://www.cdc.gov/heartdisease/heart_attack.htm#:~:text=Heart%20attack%20signs%20and%20symptoms,%2Dheaded%2C%20or%20unusually%20tired