Risk of Heart Attack Analysis

Yifeng Tang(3030065676), Lingfei Zhong(3030089347)
Fall 2022, COMP 9501
The University of Hong Kong

1 introduction

Heart diseases claimed about 71300 inpatient discharges and inpatient deaths in all hospitals, and 6561 registered deaths in 2020. It was the third commonest cause of deaths in Hong Kong and accounting for 13% of all deaths in 2020 (1). According to the Department of Health, people with multiple risk factors, such as smoking, raised blood pressure, raised cholesterol might have a higher risk of heart attack (2). Besides the risk factors mentioned, there are more relevant risk factors of heart attack that people might not be aware of. Oftentimes, people have a hard time identifying the most important risk factors and taking actions to prevent heart attack. Hence, though it is important for government to implement population-based interventions to reduce the risk of heart attack, it is also important for individuals, especially the elders, to gain more awareness of heart attack and some factors that might lead to high risk of heart attack. This study will focus on identifying and explaining some of the most important causes for heart attack by analyzing relevant dataset using various machine learning models such as support vector machine, random forest classifier and multi-layer neural network. We will also recommend some of the actions that people can take to prevent heart attack based on our analysis.

2 Dataset

We obtained the Heart Attack Analysis & Prediction Dataset from Kaggle (3). The data contains 300 data points. The predictive variable of this project is the "output" variable, which classify the sample into high risk/low risk of heart attack categories. In the dataset, we have five numeric variables and eight categorical variables. Some categorical variables might have numerical meaning. For example, data point with higher caa (number of major vessels) value indicates that this patient has more major vessels colored by fluoroscopy though caa is a categorical variables. Each sample data point is associated with the following variables displayed in Table 1.

3 Exploratory Data Analysis

We can group people from high risk/low risk catetory and see the average values of relevant variables, which are summarized in Table 2. From Table 2, we can see that out of those people who have lower risk of heart attack, 82.6 % of them are female. Out of those people who have higher risk of heart attack, 56.4 % of them are female. From Table 3, we can see that male usually have higher average value in thalachh(maximum heart rate), chol(cholestoral level), and trtbps(resting blood pressure). The exploratory data analysis also gives us some counter intuitive results. For example, Table 2 shows that people that are older will have a lower risk of heart attack. Table 2 also shows that people with higher resting blood pressure and cholestoral values are more likely to high risk of heart attack. From EDA, we can see that is is less likely for a person with thalassemia (inherited condition that may affect heart health) to have heart attack. In fact, we might expect someone with inherited heart condition to have higher risk of heart attack.

To further understand multicategorical variables, we calculated the total number of high risk patients in each category for different variables. From 5, we can see that most people with high risk of heart

Table 1: Variables and Descriptions

Variables	Explanations	Data Type
Age	Age of the patient	numeric
Sex	Sex of the patient (0 = male, 1 = female)	binary categorical
exang	exercise induced angina (1 = yes; 0 = no)	binary categorical
caa	number of major vessels (0-4) colored by fluoroscopy	categorical
cp	Chest Pain type chest pain type 1: typical angina 2: atypical angina 3: non-anginal pain 4: asymptomatic	categorical
trtbps	resting blood pressure (in mm Hg)	numeric
chol	cholestoral in mg/dl fetched via BMI sensor	numeric
fbs	(fasting blood sugar > 120 mg/dl) (1 = true; 0 = false)	binary categorical
rest_ecg	resting electrocardiographic results 0: normal 1: having ST-T wave abnormality 2: showing probable or definite left ventricular hypertrophy by Estes' criteria	categorical
thalach	maximum heart rate achieved	numeric
oldpeak	ST depression induced by exercise relative to rest	numeric
thall	thalassemia (1 = true, 0 = false)	binary categorical
slp	slope of the peak exercise ST segment	categorical
output	predicted variable: whether this patient has heart attack 0= less chance of heart attack 1= more chance of heart attack	binary categorical

attack fall into the category of having non-anginal pain. This result is also counter intuitive because angina is chest pain caused by reduced blood flow to the heart muscles, which might be a warning sign of high risk of a heart attack (4). From 5, we can see that people with fewer major vessels colored by fluoroscopy will have a higher risk of heart attack. The fluoroscopy will highlight the blood flow in the vessels, which is an indicator of smooth blood flow and healthy heart condition. Having fewer major vessels colored by fluoroscopy might be an indicator of clogged blood flow towards heart and high risk of heart attack. From 6, we can see that people with probably or definite left ventricular hypertrophy by Estes' criteria in their electrocardiographic results will have higher risk of heart attack.

However, different variables could be related to each other and contribute together to one's heart condition and EDA usually can't reflect these intertwined effects. For numerical variables, differences in average between high risk population and low risk population might be resulted from noises in the dataset. We can not simply conclude that a person with higher/lower value for a numeric variable will have a higher risk of heart attack. Similarily, we should refrain from making the conclusion that people who fall into one category of a categorical variable will have a higher risk of heart attack. As a result, more advanced machine learning analysis is needed to evaluate the importance of different variables for the risk of heart attack prediction and to help us explain some of the counter intuitive results from EDA earlier.

4 Methodology

This section presents the four popular classification methods we used in this project. We start with the most basic algorithm *perceptron*, then *support vector machine (SVM)*, *random forest classifier*, and finally *multi-layer neural network.*

4.1 Perceptron

A perceptron (5) can be considered a basic unit of a neural network, which provides the basic function of linear classification. The basic idea of the perceptron is quite simple, doing a linear computation for the inputs with the stored weights and then calling an activation function for outputs. The math formula of the perceptron is shown in Equation 1. Generally, given inputs $\mathbf{x}$, the perceptron first times

Table 2: The Average Value for Each Variables for High/Low Risk

Variable Name	Average for Low Risk	Average for High Risk
caa	1.166667	0.363636
cp	0.478261	1.375758
thalachh	139.101449	158.466667
oldpeak	1.585507	0.583030
thall	2.543478	2.121212
age	56.601449	52.496970
chol	251.086957	242.230303
trtbps	134.398551	129.303030
exng	0.550725	0.139394
sex	0.826087	0.563636
slp	1.166667	1.593939
restecg	0.449275	0.593939
fbs	0.159420	0.139394

Table 3: The Average Value for Each Variables for Male/Female

Variable Name	Average for Female	Average for Male
caa	0.811594	0.552083
cp	0.932367	1.041667
thalachh	148.961353	151.125000
oldpeak	1.115459	0.876042
thall	2.400966	2.125000
age	53.758454	55.677083
chol	239.289855	261.302083
trtbps	130.946860	133.083333
exng	0.371981	0.229167
slp	1.386473	1.427083
restecg	0.507246	0.572917
fbs	0.15942	0.12500
output	0.449275	0.750000

Table 4: Chest Pain Type and Risk of Heart Attack

Chest Pain Type	Total Number of High Risk
1: typical angina	39
2: atypical angina	41
3: non-anginal pain	69
4: asymptomatic	16

Table 5: Number of Major Vessels Colored and Risk of Heart Attack

Number of Major Vessels Colored	Total Number of High Risk
0	130
1	21
2	7
3	3
4	4

Table 6: Resting Electrocardiographic Results and Risk of Heart Attack

Resting Electrocardiographic Results	**Total Number of High Risk**
0	68
1	96
2	1

them with the corresponding weights to get $\mathbf{w}^T\mathbf{x}$. Then with a sign function, perceptron outputs $+1$ or -1 as the final output.

$$\begin{aligned} &t = f(\sum_{i=1}^{n} w_i x_i + b) = f(\mathbf{w}^T\mathbf{x}) \\ &where\ f(n) = \begin{cases} +1 \ \ if\ n \geq 0 \\ -1 \ \ if\ n < 0 \end{cases} \end{aligned} \tag{1}$$

While the perceptron is a fundamental structure, training of the perceptron is also uncomplicated. The math formula of an epoch of the training is shown below:

$$\begin{aligned} &\omega_i \leftarrow \omega_i + \Delta\omega_i \\ &where\ \Delta\omega_i = \eta\,[t - o]\,x_i \end{aligned} \tag{2}$$

where t is the target value, o is the output of the perceptron, and η is the learning rate.

4.2 SVM

SVM (6) targets to find the best hyperplane to classify the data, which has the largest separation or margin between the two categories of the data. The key is to choose the hyperplane with the largest distance to the nearest training data point of any class or, formally, the maximum-margin hyperplane. Figure 1 gives an example of three hyperplanes that classify the data, where it is not difficult to find $H3$ as the SVM's target hyperplane. The formal optimization problem of SVM is shown below:

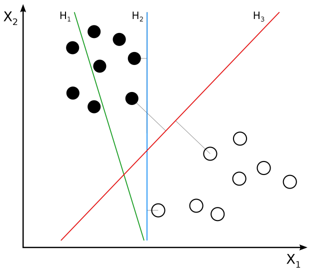

Figure 1: An example of the data with three different hyperplanes (7)

$$\begin{aligned} &min\ \ \frac{1}{2}||w||^2 \\ &s.t.\ \ y_i(w^T x_i + b) \geq 1 \end{aligned} \tag{3}$$

where x, y are the training data points, w, b is the parameter of the hyperplane. The problem is further transferred to its dual problem:

$$
\begin{aligned}
&max_{\lambda} \quad \left[\sum_{j=1}^{n} \lambda_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j y_i y_j (x_i \cdot x_j) \right] \\
&s.t. \quad \sum_{i=1}^{n} \lambda_i y_i = 0 \ \ \lambda_i \geq 0
\end{aligned}
\tag{4}
$$

which is a quadratic programming problem and can be solved efficiently with mature solutions.

4.3 Random Forest Classifier

Random forest (8) classifier is a group of multiple decision trees, and the classifier's output is the class chosen by the most trees. It is considered an extension of the bagging or bootstrap aggression method (9). Compared with the traditional strategy tree, which considers all candidate variables for each data, the random forest chooses only a subset of variables (e.g., $K = log_2 d$) each time. Therefore, because of the different variables chosen, the generalization of the random forest is better than the normal strategy tree. Figure 2 presents the structure of a random forest with simple strategy trees. With different randomly chosen subsets of the dataset and the variables, different strategy trees output different results. The random forest method chooses the majority among all outputs.

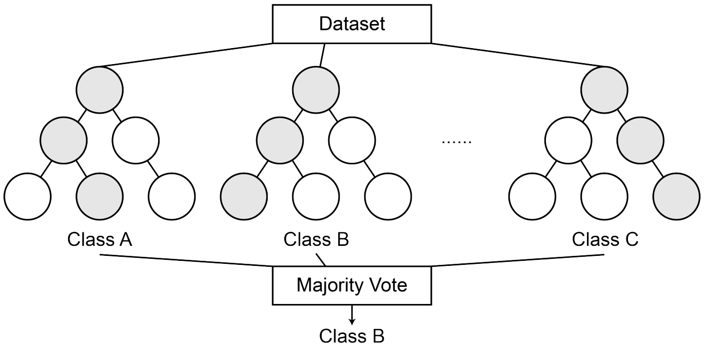

Figure 2: An example of the data with three different hyperplanes (7)

In addition to classification, the random forest method can also evaluate the importance of the variables. The idea is to quantify how much each variable contributes to the strategy trees of the random forest and then compare the average values. In this project, we introduce the Gini index for evaluation. Supposing we have J variables $X_1, ...X_J$, I strategy trees, C categories, then the Gini index of node q of the strategy tree i is computed as:

$$GI_q^{(i)} = 1 - \sum_{c=1}^{|C|} (p_{qc}^{(i)})^2 \tag{5}$$

where GI is the Gini index, p_{qc} is the percentage of the class c in node q. Then the importance of X_j at node q is computed as:

$$VIM_{jq}^{(Gini)(i)} = GI_q^{(i)} - GI_l^{(i)} - GI_r^{(i)} \tag{6}$$

where VIM reprents Variable Importance Measures or the importance we talked, $GI_l^{(i)}, GI_r^{(i)}$ are the new nodes after branching. Let us consider set Q, which contains all nodes where X_j exists, then the importance of X_j of tree i is:

$$VIM_j^{(Gini)(i)} = \sum_{q \in Q} VIM_{jq}^{(Gini)(i)} \tag{7}$$

then we have:

$$VIM_j^{(Gini)} = \sum_{i=i}^{I} VIM_j^{(Gini)(i)} \to \frac{VIM_j^{(Gini)}}{\sum_{j^{`}=1}^{J} VIM_{j^{`}}^{(Gini)}} \tag{8}$$

which is the final importance we need.

4.4 Multi-layer Neural Network

Multi-layer neural network is one of the most popular machine learning algorithms. The most straightforward multi-layer neural network, Feedforward Neural Network or Multi-Layer Perceptron (MLP), has three essential components, input layer, hidden layer, and output layer. The input layers only input the data to the network without any modifications. Each node in the hidden layer is a perceptron that receives the outputs from the former layer as the inputs. The output layer is an activation function in most situations. Compared with the single perceptron, the multi-layer perceptron has the ability to fit the non-linear function. Figure 3 shows the structure of a 4-layer multi-layer perceptron, which has two hidden layers. The method to train an MLP is backpropagation (10), which computes the gradient of the loss function with respect to each weight by the chain rule.

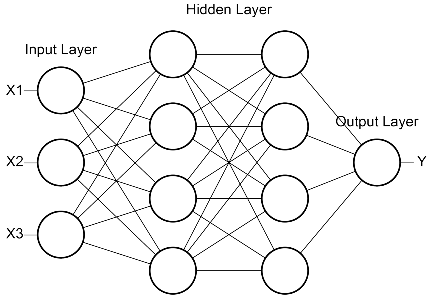

Figure 3: The structure of a 4-layer multi-layer perceptron

In this project, we implement a 4-layer multi-layer perceptron. Either of the hidden layers has eight perceptrons, and the output node is a Sigmoid activation function.

5 Evaluations

In this section, we evaluate the accuracy of the four methods mentioned in the last section with the heart attack datasets. We finish all the experiments on an Apple M1 Max chip with *scikit-learn* (11) and *Keras* (12) packages. For each experiment, we randomly choose 20% of the dataset as the test set for evaluation after training with the other 80% data. All the data are $l2$ normalized.

In addition to the accuracy results, we also collect each method's training time and the variables' importance with the random forest classifier.

5.1 Accuracy & Training Time Results

Table 7: Accracy Results on the Heart Attack Dataset

Method Name	Mean Abs. Err	Accuracy	Traning Time
Perceptron	0.2131	78.68%	0.5569 ms
SVM Classifier	0.1639	83.60%	0.9610 ms
Random Forest Classifier	0.1147	88.52%	101.66 ms (100 trees, 8 threads)
4-Layer MLP	0.2482	85.24%	1900.32 ms (150 epochs)

Table 7 lists the accuracy and the training time results we collected. In general, for accuracy results, the random forest classifier reports the best accuracy with the result of 88.52%, while the perceptron reports the worst result of 78.68%. Especially perceptron and SVM show worse accuracy results than the other two methods. Since the former two methods are linear classifiers while the latter two are non-linear classifiers, it represents that the heart attack data we used cannot be linearly classifiable.

For the training time results, the perceptron and the SVM classifier are trained extremely fast with the results of < 1 ms. In contrast, the random forest classifier needs 101.66 ms to finish training, even with eight threads executing concurrently. The 4-layer MLP converges at about 150 epochs and needs 1900.32 ms, much longer than other methods.

Combining the accuracy and the training time results, we conclude that the random forest classifier method should be the best method for this dataset's prediction work.

5.2 Variable Importance

Table 8: Variable Importance on the Heart Attack Dataset

Variable Name	Importance Value	Description
caa	0.133137	Number of major vessels colored by fluoroscopy
cp	0.123768	Chest pain type
thalachh	0.115053	Maximum heart rate achieved
oldpeak	0.103460	ST depression induced by exercise relative to rest
thall	0.094571	Thalassemia
age	0.091523	Age
chol	0.080982	Cholesterol in mg/dl fetched via BMI sensor
trtbps	0.072010	Resting blood pressure (in mm Hg)
exng	0.060947	Exercise induced angina
sex	0.054320	Sex
slp	0.040743	Slope of the peak exercise ST segment
restecg	0.019957	Resting electrocardiographic results
fbs	0.009528	Fasting blood sugar > 120 mg/dl

Table 8 lists the dataset's variables with the corresponding importance values. From the results, we can easily observe that the number of major vessels colored by fluoroscopy, the chest pain type, and the maximum heart rate are the top 3 important variables to predict a heart attack.

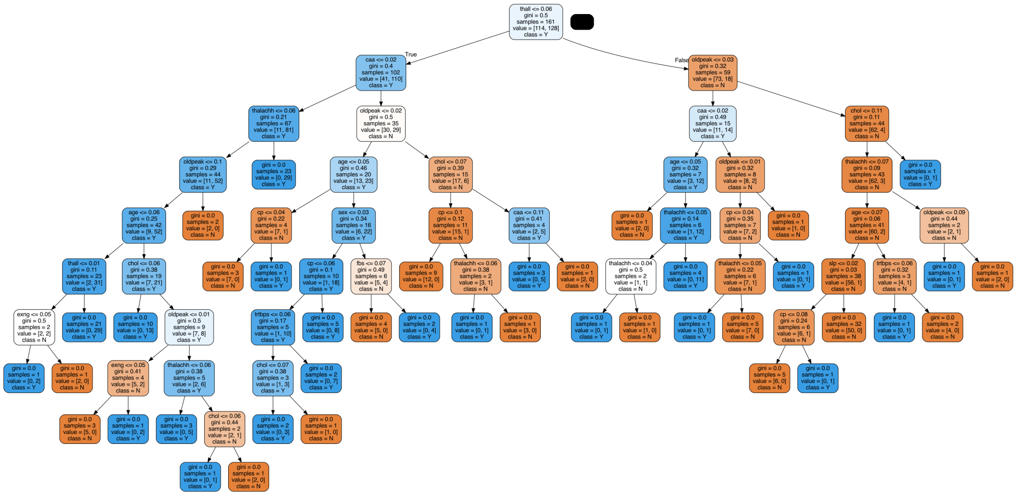
thall <= 0.06 gini = 0.5 samples = 161 value = [114, 128] class = Y
[redacted]
True
caa <= 0.02 gini = 0.4 samples = 102 value = [41, 110] class = Y
False
oldpeak <= 0.03 gini = 0.32 samples = 59 value = [73, 18] class = N
thalachh <= 0.06 gini = 0.21 samples = 67 value = [11, 81] class = Y
oldpeak <= 0.02 gini = 0.5 samples = 35 value = [30, 29] class = N
caa <= 0.02 gini = 0.49 samples = 15 value = [11, 14] class = Y
chol <= 0.11 gini = 0.11 samples = 44 value = [62, 4] class = N
oldpeak <= 0.1 gini = 0.29 samples = 44 value = [11, 52] class = Y
gini = 0.0 samples = 23 value = [0, 29] class = Y
age <= 0.05 gini = 0.46 samples = 20 value = [13, 23] class = Y
chol <= 0.07 gini = 0.39 samples = 15 value = [17, 6] class = N
age <= 0.05 gini = 0.32 samples = 7 value = [3, 12] class = Y
oldpeak <= 0.01 gini = 0.32 samples = 8 value = [8, 2] class = N
thalachh <= 0.07 gini = 0.09 samples = 43 value = [62, 3] class = N
gini = 0.0 samples = 1 value = [0, 1] class = Y
age <= 0.06 gini = 0.25 samples = 42 value = [9, 52] class = Y
gini = 0.0 samples = 2 value = [2, 0] class = N
cp <= 0.04 gini = 0.22 samples = 4 value = [7, 1] class = N
sex <= 0.03 gini = 0.34 samples = 16 value = [6, 22] class = Y
cp <= 0.1 gini = 0.12 samples = 11 value = [15, 1] class = N
caa <= 0.11 gini = 0.41 samples = 4 value = [2, 5] class = Y
gini = 0.0 samples = 1 value = [2, 0] class = N
thalachh <= 0.05 gini = 0.14 samples = 6 value = [1, 12] class = Y
cp <= 0.04 gini = 0.35 samples = 7 value = [7, 2] class = N
gini = 0.0 samples = 1 value = [1, 0] class = N
age <= 0.07 gini = 0.06 samples = 41 value = [60, 2] class = N
oldpeak <= 0.09 gini = 0.44 samples = 2 value = [2, 1] class = N
thall <= 0.01 gini = 0.11 samples = 23 value = [2, 31] class = Y
chol <= 0.06 gini = 0.38 samples = 19 value = [7, 21] class = Y
gini = 0.0 samples = 3 value = [7, 0] class = N
gini = 0.0 samples = 1 value = [0, 1] class = Y
cp <= 0.06 gini = 0.1 samples = 10 value = [1, 18] class = Y
fbs <= 0.07 gini = 0.49 samples = 6 value = [5, 4] class = N
gini = 0.0 samples = 9 value = [12, 0] class = N
thalachh <= 0.06 gini = 0.38 samples = 2 value = [3, 1] class = N
gini = 0.0 samples = 3 value = [0, 5] class = Y
gini = 0.0 samples = 1 value = [2, 0] class = N
thalachh <= 0.04 gini = 0.5 samples = 2 value = [1, 1] class = N
gini = 0.0 samples = 4 value = [0, 11] class = Y
thalachh <= 0.05 gini = 0.22 samples = 6 value = [7, 1] class = N
gini = 0.0 samples = 1 value = [0, 1] class = Y
slp <= 0.02 gini = 0.03 samples = 38 value = [56, 1] class = N
trtbps <= 0.06 gini = 0.32 samples = 3 value = [4, 1] class = N
gini = 0.0 samples = 1 value = [0, 1] class = Y
gini = 0.0 samples = 1 value = [2, 0] class = N
exng <= 0.05 gini = 0.5 samples = 2 value = [2, 2] class = N
gini = 0.0 samples = 21 value = [0, 29] class = Y
gini = 0.0 samples = 10 value = [0, 13] class = Y
oldpeak <= 0.01 gini = 0.5 samples = 9 value = [7, 8] class = Y
trtbps <= 0.06 gini = 0.17 samples = 5 value = [1, 10] class = Y
gini = 0.0 samples = 5 value = [0, 8] class = Y
gini = 0.0 samples = 4 value = [5, 0] class = N
gini = 0.0 samples = 2 value = [0, 4] class = Y
gini = 0.0 samples = 1 value = [0, 1] class = Y
gini = 0.0 samples = 1 value = [3, 0] class = N
gini = 0.0 samples = 1 value = [0, 1] class = Y
gini = 0.0 samples = 1 value = [1, 0] class = N
gini = 0.0 samples = 1 value = [0, 1] class = Y
gini = 0.0 samples = 5 value = [7, 0] class = N
cp <= 0.08 gini = 0.24 samples = 6 value = [6, 1] class = N
gini = 0.0 samples = 32 value = [50, 0] class = N
gini = 0.0 samples = 1 value = [0, 1] class = Y
gini = 0.0 samples = 2 value = [4, 0] class = N
gini = 0.0 samples = 1 value = [0, 2] class = Y
gini = 0.0 samples = 1 value = [2, 0] class = N
exng <= 0.05 gini = 0.41 samples = 4 value = [5, 2] class = N
thalachh <= 0.06 gini = 0.38 samples = 5 value = [2, 6] class = Y
chol <= 0.07 gini = 0.38 samples = 3 value = [1, 3] class = Y
gini = 0.0 samples = 2 value = [0, 7] class = Y
gini = 0.0 samples = 5 value = [6, 0] class = N
gini = 0.0 samples = 1 value = [0, 1] class = Y
gini = 0.0 samples = 3 value = [5, 0] class = N
gini = 0.0 samples = 1 value = [0, 2] class = Y
gini = 0.0 samples = 3 value = [0, 5] class = Y
chol <= 0.06 gini = 0.44 samples = 2 value = [2, 1] class = N
gini = 0.0 samples = 2 value = [0, 3] class = Y
gini = 0.0 samples = 1 value = [1, 0] class = N
gini = 0.0 samples = 1 value = [0, 1] class = Y
gini = 0.0 samples = 1 value = [2, 0] class = N

Figure 4: Tree 34 from Random Forest

From Figure 4, we can see that people with lower value in the chest pain type variable is more likely to have higher risk of heart attack. In the tree, smaller chest pain value is more often to be classified as the high risk category. This result contradicts what what we observe from EDA as people with high risk of heart attack are more likely to have higher value in the chest pain type. It is not too surprising that people with higher maximum heart rate value will have a higher risk of heart attack, which is consistent with the result from EDA. In EDA (2), people who have lower risk of heart attack have lower average value of maximum heart rate than people with high risk of heart attack. From Figure 4, we can also see that people with fewer major vessels colored by fluoroscopy can have higher risk of heart attack. The reason is that the color doppler would color the running blood in the vessels blue and red to mark the directions. When the blood volume is insufficient, the fluoroscopy could fail to color, representing a high possibility of a heart attack.

We can gain some other insights by ananlyzing variables listed in Table 8. People might often think that inherited heart condition is detrimental to heart health, but at least from this study, inherited heart condition is not the most important factor. Inherited heart condition might affect the three most important risk factors identified above, but undesired values in these variables might not be directly resulted from inherited heart condition. Though Table 8 also points out that inherited heart condition might be a more important risk factor than age and cholesterol level, we should keep in mind that heart attack is very closely related to lifestyle and acquired eating/exercising habits.

The results displayed in Table 8 also highlights the difficulties for individuals to identify risk factors heart attack on their own. People might be able to noticed chest pain, but many people might not be able to identify the reason of the chest pain. It is also difficulty for people to get accurate estimation of their value of number of major vessels colored by fluoroscopy and maximum heart rate without going to the hospital. Hence, it is very important for people to closely monitor their health and do body check regularly to identify potential problems.

6 Discussion

In this study, we use a data with 300 datapoints. In Section 5.1, we can observe that though random forest classifier gives the highest accuracy, it takes a significantly longer time than perceptron and SVM classifier. If the dataset is larger, it will be more and more expensive to train random forest classifier. Though it takes longer to train random forest trees, random forest is better at selecting the most relevant features to make accurate predictions. Feature selection is particularly important in this study. Many factors might be related to high risk of heart attack, but they might not be significant. Identifying the most relevant risk factors will allow more people to have better understanding of heart attack and act early to prevent heart attack from happening.

7 Conclusion

From machine learning algorithms, we have identified number of major vessels colored by fluoroscopy, chest pain type and maximum heart rate achieved to be the most important three factors. Though inherited heart conditions, sex, and age are also potential risk factors of heart attack, they are not as important as the three risk factors mentioned above. Hence, people should pay more attention to maintain healthy eating and exercising patterns if they would like to lower the risk of heart attack. This study also implies that it might be difficult for one to monitor the values of the three important variables mentioned above, and that people should do body check at the hospitals regularly if they would like to prevent potential heart attack.

For the public, the results from this study highlights the importance of maintaining a healthy lifestyle, as inherited heart condition and age are not the most important risk factor of heart attack. People should try to avoid behaviors that might increase the risk of having clogged vessels such as eating high-fat or high-sugar food and smoking. As mentioned previously, it is important for people to keep the risk factors of heart attack in mind and pay attention to the connections between their daily behaviors and risk factors. A small unhealthy chronical action could build up its affect and eventually reflect negative influence on the body, whereas a small healthy pattern could help people to maintain desired heart health throughout their lives.

References

[1] T. G. o. t. H. K. S. A. R. Department of Health, "Heart diseases." https://www.chp.gov.hk/en/healthtopics/content/25/57.html. Accessed:2022-12-01.

[2] T. G. o. t. H. K. S. A. R. Department of Health, "Prevent heart attacks and strokes through drug therapy and counselling." https://www.change4health.gov.hk/en/saptowards2025/target8.html. Accessed:2022-12-02.

[3] R. RAHMAN, "Heart attack analysis prediction dataset." https://www.kaggle.com/datasets/rashikrahmanpritom/heart-attack-analysis-prediction-dataset. Accessed:2022-12-02.

[4] NHS, "Angina." https://www.nhs.uk/conditions/angina/. Accessed:2022-12-02.

[5] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115–133, 1943.

[6] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[7] Wikipedia, "Support vector machine." https://en.wikipedia.org/wiki/Support_vector_machine. Accessed: 2022-11-29.

[8] T. K. Ho, "Random decision forests," in *Proceedings of 3rd international conference on document analysis and recognition*, vol. 1, pp. 278–282, IEEE, 1995.

[9] L. Breiman, "Bagging predictors," *Machine learning*, vol. 24, no. 2, pp. 123–140, 1996.

[10] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[11] S. Community, "scikit-learn." https://scikit-learn.org/stable/. Accessed: 2022-11-29.

[12] K. Community, "Keras." https://keras.io/. Accessed: 2022-11-29.