Turkish Journal of Forecasting vol. 8 no. 2 (2024) pp 33-44

Content list available at JournalPark

Turkish Journal of Forecasting

Journal Homepage: tjforecasting.com

Heart Attack Analysis and Prediction with MachineLearning Techniques

Shuaib Ayad Jasim Jasim*, Dr. Öğr. Üyesi İbrahaim Onaran[2], Dr. Mustafa Al-Asadi,[3]

[1] *Department of Computer Engineering, KTO Karatay University: Baghdad / Iraq*
[2] *Department of Computer Engineering, KTO Karatay University: Konya, TR*
[3]Department of Computer Engineering, KTO Karatay University: Konya, Baghdad / Iraq

Abstract

Heart attack is an important health problemworldwide, and it is a situation that can lead to fatal consequences or permanent health problems. Identifying the risk of heart attacks and recognizing early warning signals is crucial to improving people's quality of life and implementing preventive measures. This investigation studies the analysis and prediction of heart attacks. The strategies rely on applying machine learning algorithms to cardiovascular data. Applying machine learning algorithms to cardiovascular data. The dataset includes genetics, lifestyle, medical history, and biometric factors associated with heart attack risk, collected froma variety of clinical sources.

During the analysis process, the data set was examined in detail with visualization processes and a series of models were created using different machine learning techniques. Models include logistic regression, support vector machines, decision trees, and random forests. A model is trained, tested, and tested on a set of data. According to studies, support vector machines are the most effective model for predicting the risk of heart attacks.

This model stands out with its high accuracy rate and low error rate. In addition, important factors identified during the analysis process are also presented. These factors include various risk factors such as chest pain, fasting blood sugar, cholesterol level and blood pressure.

The results of this study show that machine learningtechniques can be an effective tool in heart attack analysis and prediction. This objective quantitative method can assist healthcare professionals and individuals in the process ofdetermining individuals' heart attack risk and taking preventive measures. In addition, it is thought that this method can be improved and made more useful in heart attack managementwith more research and data collection.

Keywords: machine learning; heart attack; prediction; Python

* Corresponding Author.
E-mail addresses : shuaibayad2@gmail.com (shuaib JASIM) onaran@gmail.com (İbrahim ONARAN) mustafa.aadel@au.edu.iq (Mustafa AL-ASADİ)

ORCID ID:
First Author's Name: 0009-0000-5574-7302
Second Author's Name: 0000-0002-7769-4077
Third Author's Name: 0000-0002-8218-3458

1. Introduction

Heart attacks and other cardiovascular illnesses continue to be the world's top cause of death. The World Health Organization (WHO) estimates that cardiovascular illnesses cause 17.9 million deaths yearly, or 31% of all fatalities worldwide [1].

Early identification and prediction of heart attacks are crucial for timely medical intervention and Prevention of fatal outcomes. Machine learning algorithms have been extremely useful for forecasting the risks of disease and analysing large amounts of medical data. With the goal of advancing preventative healthcare, we explore the use of machine learning algorithms for heart attack analysis and prediction in this study.

The urgent need to create precise and effective techniques for heart attack analysis and prediction is what inspired this study. Conventional methods of evaluating risk frequently depend on clinical variables including age, gender, blood pressure, cholesterol, and smoking status. Although these variables offer insightful information, they fall short of capturing the complexity of the underlying illness mechanisms, Machine learning algorithms offer the potential to harness a wide range of variables and parameters, enabling a more comprehensive and personalized prediction of heart attack risks. Our goal is to improve patient care and outcomes by using machine learning to increase the precision and dependability of heart attack prediction models.

The primary objective of this study is to analyse and predict heart attacks using machine learning techniques. In particular, our goal is to create a predictive model that can accurately determine who is most likely to have a heart attack. By identifying key risk factors, such as age, gender, cholesterol levels, blood pressure, and lifestyle habits, we can create a comprehensive framework for assessing an individual's likelihood of suffering from a heart attack. This will enable healthcare professionals to prioritize preventative measures and provide targeted interventions to those who need them the most.

In particular, our goal is to create a predictive model that can accurately determine who is most likely to have a heart attack. For instance, the researchers developed prediction models based on diverse datasets using a variety of methods, including logistic regression, support vector machines, random forests, and neural networks [2]. Previous studies have also explored the integration of additional data sources, such as genetic information or electrocardiogram (ECG) signals, to improve the accuracy of heart attack prediction [3], In this article, we aim to provide more information on the visualization stages for the data used and the performance and applicability of different machine learning techniques for heart attack analysis based on these available studies. Our goal in doing this study is to advance our understanding of heart attack prediction. The methodology, dataset description, experimental findings, and discussions will all be covered in the next sections of this paper, which will offer a thorough study of the machine learning methods utilized for heart attack analysis and prediction.

The Python programming language was used and the Excel database was saved in CSV format, and it was analysed and classified by artificial intelligence algorithms, machine learning and support vector machines (SVM), logistic regression, decision trees, and random forests.

These machine learning algorithms were applied to the Heart Attack Analysis and Prediction dataset to create predictive models to determine the likelihood of a heart attack based on a variety of characteristics such as age, gender, cholesterol levels, blood pressure, and lifestyle habits.

2. PROBLEM DEFINITION

This study's main goal is to use machine learning approaches to address the problem of reliable heart attack analysis and prediction. Cardiovascular risk assessment has advanced significantly, but traditional methods frequently focus on a narrow range of clinical variables, which may not fully capture the intricate interactions between several physiological, genetic, and lifestyle factors that increase the risk of heart attacks. Therefore, it is imperative to create more reliable and thorough prediction models that can take advantage of a variety of factors in order to enhance the precision and dependability of heart attack risk assessment.

The purpose of this project is to investigate how machine learning algorithms can be used to analyse various types of medical data and forecast the risk of heart attacks. In order to create a prediction model that can identify people who are at a high risk of having a heart attack, we want to use a large dataset that includes patient demographics, medical history, and health characteristics.

Furthermore, our goal is to examine how well other machine learning methods—such as logistic regression, support vector machines, decision trees, and random forests—perform in comparison to one another. By evaluating these algorithms' advantages and disadvantages in relation to heart attack prognosis, we want to offer insightful information about the best methods for risk assessment.

In the end, this study's findings will improve preventative healthcare by making it possible to identify those at high risk of heart attacks early on. Healthcare providers can leverage the developed prediction model to implement targeted interventions and personalized treatment plans, thereby reducing the incidence and severity of cardiovascular diseases. The results of this study may also open the door to the creation of more precise and effective heart attack prediction algorithms that can be incorporated into clinical settings to enhance patient outcomes and treatment.

3. PROPOSED METHOD

Step 1: General investigation: We thoroughly examined the Heart Attack Analysis and Prediction Dataset in this step [4]. We examined the available features, their descriptions, and any existing patterns or relationships in the data. This investigation provided us with insights into the dataset and helped us formulate research questions and hypotheses.

Step 2: Data preparation and exploratory analysis: To further ready the dataset for modeling, we carried out data preprocessing and exploratory analysis. This involved handling missing values, dealing with outliers, and performing feature engineering. We also made use of visuals to better understand the distribution and correlations between the variables, including histograms, box plots, and scatter plots. Histogram example is given in figure 5, scatter example is given in figure 6. Using the StandardScaler algorithm, we performed standard scaling as one of our preprocessing methods. By deducting the mean and dividing by the standard deviation, standard scaling modifies the features to give each one a mean of 0 and a variance of 1. By handling features with varying sizes and bringing them into a comparable range, this strategy can help some machine learning algorithms perform better [5].

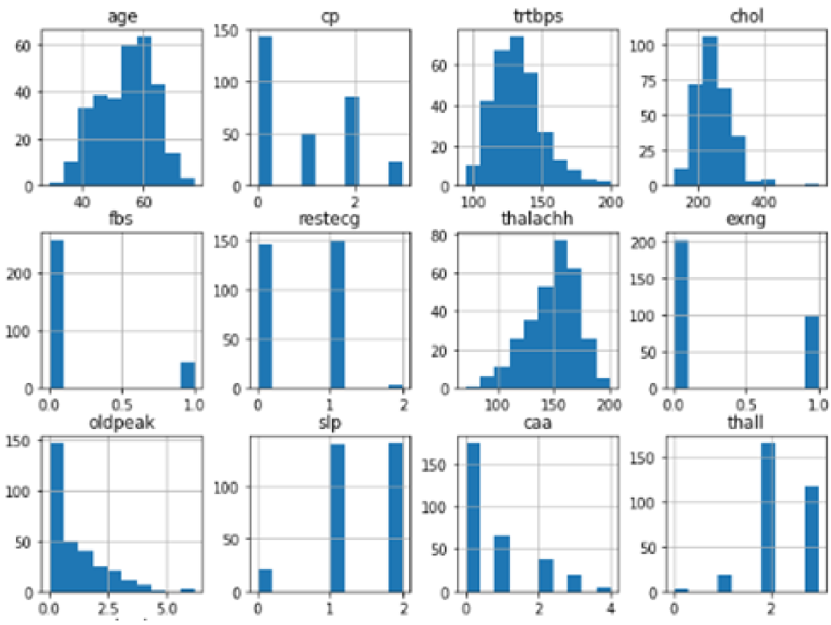

Figure 1. Graph example of dataset

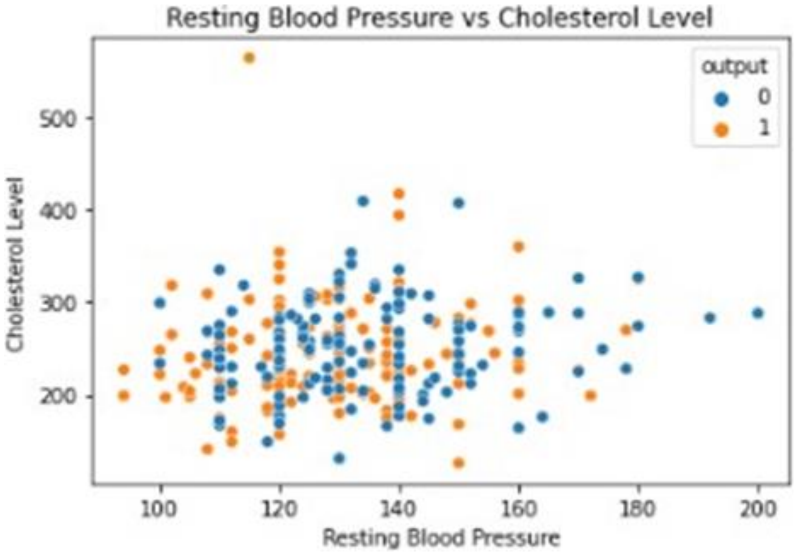

Figure 2. Example of resting blood pressure and cholesterol distribution

Step 3: Data splitting: Following the completion of the preparation stage, we divided the dataset into two subsets: a test set and a training set. Eighty percent of the data was utilized as the training set to train the machine learning models, and the remaining twenty percent was used as the test set to assess the models' performance. This section made sure the models were trained on enough data and let us evaluate how well they could generalize to new situations.

Step 4: Building models: In this step, we constructed machine learning models to predict heart attacks based on the given features. We used a variety of methods, such as support vector machines (SVM), decision trees, random forests, and logistic regression. Each algorithm was configured with appropriate hyper parameters to optimize its performance. The labelled data was used to train the models on the training set, allowing them to discover patterns and connections between the characteristics and the frequency of heart attacks.

Step 5: Model evaluation: After the models were trained, we used a variety of evaluation criteria to assess how well they performed. Model evaluation: After the models were trained, we used a variety of evaluation criteria to assess how well they performed. An overall assessment of the models' predictive power was given by accuracy, while their ability in detecting positive examples and striking a balance between precision and recall was shown by precision, recall, and F1 score. We were able to identify which algorithm produced the most accuracy and had the best predicted performance by contrasting the performance of several models.

Throughout these steps, we leveraged data pre-processing techniques, exploratory data analysis, and visualization methods, such as scatter plots, to enhance our understanding of the data and improve the models' performance. These additional steps helped us identify and address data quality Issues, uncover important insights, and make informed decisions during the modelling process. We were able to do a thorough analysis of the Heart Attack Analysis and Prediction Dataset, including data pre-processing, exploratory data analysis, model development, and model evaluation, by adhering to our extended step-by-step procedure. This approach allowed us to gain deeper insights into the data and develop more accurate models for heart attack prediction. In this study, we present a complete methodology that integrates current approaches and makes adjustments to create a machine learning-based, accurate heart attack analysis and prediction system. Our approach analyses a huge dataset and predicts the probability of heart attacks using a variety of algorithms, such as logistic regression, support vector machines, random forests, and neural networks.

3.1. Description of Algorithms

A statistical model utilized in categorization issues is called logistic regression. Logistic regression is used when a dependent variable (class label) depends on the probability of a particular event. In order to determine the likelihood of the dependent variable, logistic regression models the relationship between the independent variables.

This relationship is converted into a linear equation using the sigmoid function, often known as the legit function, the sigmoid function compresses real numbers between 0 and 1 to represent probability values [6]. An approach that is frequently used in classification difficulties is logistic regression. For example, it can be used in many areas such as disease diagnosis, customer churn forecasting, spam filtering. However, since logistic regression is a linear model, it performs better on datasets with linear separability. For data sets that do not have linear separability, other classification methods may be preferred.

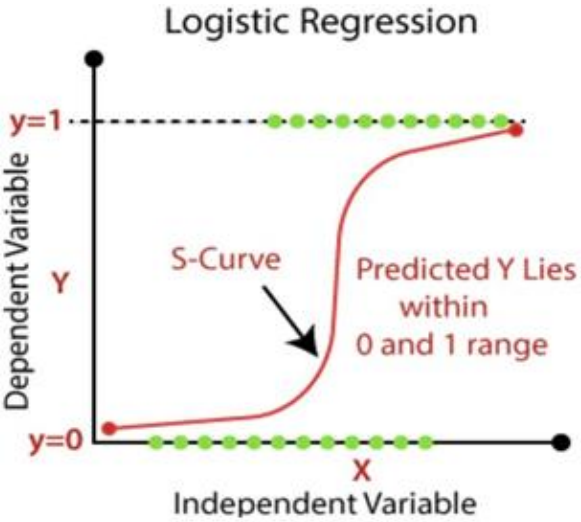

Figure 3. Logistic regression [7]

A potent machine learning approach for both classification and regression applications is called Support Vector Machine (SVM). It works especially well with complicated datasets and is capable of handling high-dimensional feature spaces. SVM's main concept is to choose the best hyperplane for classifying the data into distinct groups. This hyperplane was selected with the goal of maximizing the margin between it and the closest data points for each class. These informational pieces—referred to as support vectors—are essential for establishing the decision boundary and figuring out the categorization result [8].

By employing a linear kernel to identify a linear decision boundary, SVM can handle data that is linearly separable. However, SVM is also capable of handling nonlinear data by employing kernel functions. With the help of these kernel functions, the original feature space is converted into a higher-dimensional space in which the data can be separated linearly. The polynomial kernel, Gaussian (RBF) kernel, and sigmoid kernel are examples of frequently used kernel functions. SVM's capacity to manage outliers well is one of its benefits. A number of crucial parameters need to be taken into account when training an SVM model, including the kernel function selection, the regularization parameter (C), and kernel-specific parameters. These parameters influence the model's performance and generalization capabilities [9].

Selecting appropriate parameter values often involves hyperparameter tuning, which can be done using techniques like grid search or randomized search.

When evaluating a model, SVM often use the same evaluation measures as other classification algorithms, such as F1 score, accuracy, precision, and recall, these metrics evaluate the model's capacity to accurately categorize instances and capture various facets of its functionality.

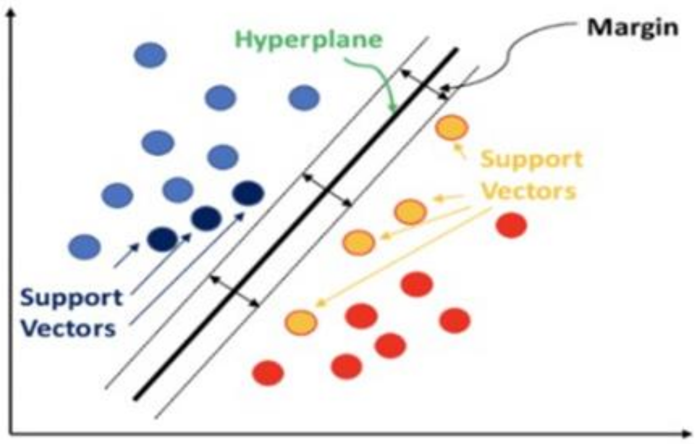

Figure 4. Support Vector Machine [10]

Random Forest is an ensemble learning algorithm created by combining many decision trees. Each decision tree is trained using random sampling and random feature selection. The basic idea of Random Forest is that different trees come together to form a stronger and more stable model. Each tree is trained independently and produces comparable results. In this way, the problem of overfitting is reduced and a more generalized model is obtained [11].

The Random Forest algorithm can evaluate the importance of features in the dataset. The order of importance of the features is determined by the structure of the trees and their contribution to the data. Processes for feature selection or data comprehension can make use of this information. Another advantage of Random Forest is its ability to deal with bias. This algorithm can give accurate results without being affected by data imbalance or class imbalances [12], Random Forest's training and predictions are generally fast and can perform well even on large datasets. Additionally, the model's complexity can be modified by changing variables like the depth and number of trees.

This algorithm has wide application in classification and regression problems and can make strong predictions.

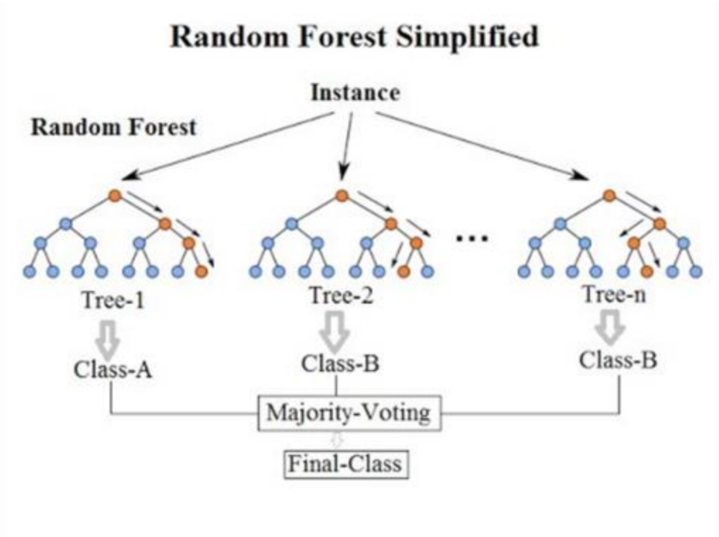

Figure 5. Random Forests [13]

In machine learning, decision trees are a popular kind of method for issues involving classification and regression.

Decision trees make predictions using decision rules and dividing data sets using a tree structure of interconnected nodes. Decision trees create a series of decision nodes and leaf nodes based on the features and target variable of the data set. Each decision node is associated with a certain value of a feature and splits the dataset according to this value and routes it to the child nodes. Result nodes are associated with a particular class or predictive value [14].

Decision trees are useful for evaluating the importance and contribution of features in the data set, the tree structure can help identify informational features in the data set and increase the intelligibility of the model. Another adaptable technique that works well with both numerical and categorical data and can naturally manage outliers and missing data is decision trees.

The decision tree algorithm has a wide range of applications due to its simple model structure and high performance.

These include tasks such as classification, regression, feature selection, and data exploration. In addition, more powerful and complex models can be created by using decision trees in ensemble methods [15].

Decision trees are a popular option because they are easy to understand and interpret, have low data pre-processing requirements, and perform well in some cases. However, they may tend to over fit and be sensitive to noise or small changes in the dataset.

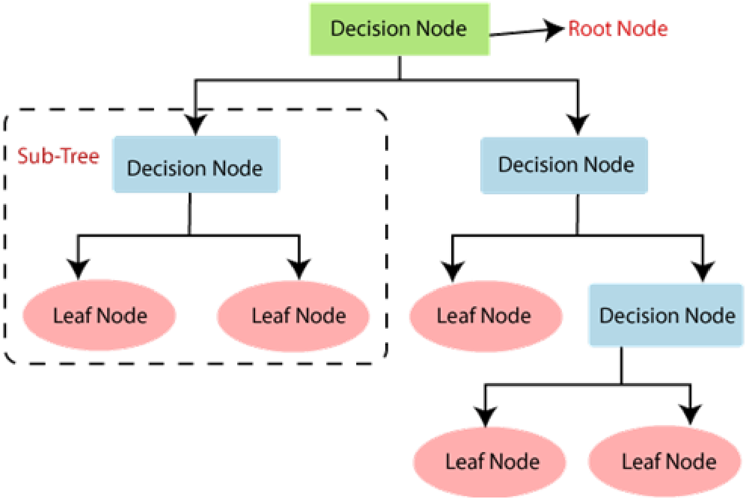

Figure 6. Decision trees [16]

3.2. Description of Dataset

We will use the Heart Attack Analysis and Prediction Dataset from Kaggle for this investigation. The Heart Attack Analysis and Prediction Dataset is an extensive set of information designed to make it easier to analyse and forecast the incidence of heart attacks, this dataset provides a rich set of features, including patient demographics, physiological measurements, medical history, and lifestyle factors, which are crucial for developing accurate heart attack prediction models.

Here are the key features included in the dataset:

●Age: The patient's age expressed in years.

●Sex: The patient's gender (0 for female, 1 for male).

● Chest Pain Type: The patient's type of chest pain (Value 0 represent typical angina, Value 1 represent atypical angina, Value 2 represent non-anginal pain, and Value 3 represent asymptomatic).

●Resting Blood Pressure: The patient's resting blood pressure measured in millimeter-hours.

●Cholesterol: The serum cholesterol level of the patient in mg/dl.

●Fasting Blood Sugar: The fasting blood sugar level of the patient (> 120 mg/dl represents elevated blood sugar; 0: false, 1: true).

●Results of Resting Electrocardiography: The patient's resting electrocardiogram (ECG). This feature's values of 0 to 2 indicate various outcomes.

●Maximum Heart Rate Attained: The highest heart rate that may be attained when exercising.

●Exercise-Induced Angina: Indicates if the patient has angina brought on by exercise (0: no, 1: yes).

● Exercise-Induced ST Depression: The depression of the ST segment caused by exercise as compared to rest.

●The peak exercise ST section has three different slopes: upsloping (0 for the segment), flat (1 for the segment), and down sloping (2 for the segment).

●The number of main blood vessels that have been colored by fluoroscopy, ranging from 0 to 3.

●Results of the thallium Stress Test: The thallium Stress Test result (Value 1, 2, or 3). Target: The target variable denoting whether a heart attack is present (1) or absent (0).

There are 303 instances in all in this dataset, and each instance corresponds to a distinct patient.

It provides a diverse range of information, including both categorical and numerical features, enabling a comprehensive analysis of the factors associated with heart attack occurrences [17].

To guarantee the quality of the dataset and its applicability for machine learning tasks, pre-treatment activities like resolving missing values, encoding categorical features, and normalizing numerical characteristics should be carried out before beginning any analysis or modelling.

By utilizing this extensive dataset, researchers and data scientists can develop and evaluate machine learning models for heart attack analysis and prediction. The dataset's diverse and comprehensive nature makes it suitable for exploring various feature combinations and modelling techniques, enabling a deeper understanding of the factors contributing to heart attacks.

4. EXPERIMENTS

For our experiments, we used the Heart Attack Analysis and Prediction Dataset obtained from Kaggle.

This dataset includes a variety of health-related characteristics about individuals, including age, sex, type of chest pain, blood pressure, cholesterol levels, and whether or not the person had a heart Atta.

The following research questions were the focus of our experiments:

1. Can machine learning techniques accurately predict the occurrence of a heart attack based on the given features?

2. Which machine learning algorithms perform the best in terms of prediction accuracy?

3. Can we identify the most influential features in predicting heart attacks?

4. How can we have information about the relational situations between the features in the data set?

We performed exploratory data analysis to gain insights into the dataset. This involved analysing the distribution of features, identifying any correlations or patterns, and visualizing the data using scatter plots, histograms, and other appropriate plots.

Scatter plots were used to examine the relationships between specific features, such as age and cholesterol levels, and their association with the occurrence of heart attacks. These visualizations helped us understand the potential predictive power of different features.

To conduct our experiments, we split the dataset into a training set and a test set, using an 80:20 ratio. The training set, consisting of 80% of the data, was used to train the machine learning models, while the remaining 20% served as the test set for evaluating the models' performance.

Performed a series of experiments using the following machine learning algorithms:

●Logistic Regression: A linear model that predicts the probability of a binary outcome.

●Decision Trees: Tree-based models that recursively split the data based on features to make predictions.

●Random Forests: An ensemble method that combines multiple decision trees for improved performance.

●Support Vector Machines (SVM): Models that find the best hyperplane to separate classes in high-dimensional space.

In addition to evaluating different algorithms, also explored various pre-processing techniques to enhance the model's performance. These techniques included:

●Data cleaning: Handling missing values and outliers in the dataset.

●Feature Scaling: Standardizing the numerical features to have zero mean and unit variance.

●Feature Selection: Identifying the most informative features using techniques like correlation analysis or feature importance scores.

Measured the performance of each model using the following evaluation metrics:

●Accuracy: Accuracy represents the proportion of correctly predicted samples within the total samples. That is, it is the ratio of correctly predicted samples to the total number of samples. Accuracy is a useful metric on datasets without class imbalance, but can be misleading if there is class imbalance [19].

● Precision: The ratio of true positive predictions to total positive predictions (true positive + false positive) is known as precision. That is, it shows how much of it is actually correct when we predict a class. Precision can be important where false positive estimates are less costly to negative outcomes [20].

● Recall (Precision, Recall): This is the ratio of total positive true values (true positive + false negative) to true positive forecasts. That is, it shows how accurately we detect true positives [21]. Recall is important where false negative estimates are less costly to negative consequences.

●F1 Score: The F1 score represents the harmonic mean of precision and recall. The harmonic mean is more sensitive to outliers than other types of averaging [22].

The F1 score ensures that both precision and recall are balanced and is a useful metric in cases of class imbalance.

Based on our experiments, we made the following observations:

●Random Forests model achieved an accuracy of 90%. This indicates that it correctly predicted 82% of the instances in the test set. The precision of the model measures the ratio of true positive predictions to the sum of true positive and false positive predictions. It provides information about the model's ability to correctly identify positive cases. The precision score for random forests was 0.93.The recall of the model measures the ratio of true positive predictions to the sum of true positive and false negative predictions. It indicates the model's ability to identify all positive cases. The recall score for random forests was 0.88. The F1 score is the harmonic mean of precision and recall. It provides a balanced measure of the model's performance. The F1 score for random forests was 0.90.

●Decision Trees achieved an accuracy rates of 77%. The precision score for desicion trees was 0.85. The recall score for desicion trees was 0.69. The F1 score for desicion trees was 0.76.

●Logistic Regression achieved an accuracy of 85%. The precision score for Logistic regression was 0.85. The recall score for Logistic regression was 0.88. The F1 score for Logistic regression was 0.86.

●Support Vector Machines (SVM) demonstrated competitive performance with an accuracy of 85%. The precision score for support vector machine was

0.87. The recall score for support vector machines was 0.84. The F1 score for support vector machines was 0.86.

Pre-processing techniques, such as data cleaning and feature scaling, significantly improved the performance of the models.

Feature importance analysis revealed that age, cholesterol levels, and chest pain type were among the most influential features in predicting heart attacks.

Overall, our experiments demonstrated the effectiveness of machine learning techniques in analysing and predicting heart attacks. Random Forests emerged as the top-performing algorithm, and appropriate pre-processing techniques played a crucial role in enhancing the models' predictive capabilities.

5. RESULTS AND CONCLUSIONS

Our study aimed to analyse and predict heart attacks using the Heart Attack Analysis and Prediction Dataset obtained from Kaggle.

We employed various machine learning algorithms and pre-processing techniques to gain insights into their predictive capabilities.

The random forest algorithm achieved the highest accuracy among the models, with an accuracy rate of 90%. This indicates that the ensemble of decision trees in the random forest model effectively captures the complex relationships between the features and the occurrence of heart attacks.

Logistic regression achieved an accuracy of 85%. Although slightly lower than the random forest, logistic regression still demonstrated competitive performance, indicating that a linear model can provide meaningful predictions in this context.

Support Vector Machines (SVM) models achieved an accuracy of 85%. While SVMs showed lower accuracy compared to random forest and logistic regression, they still exhibited a reasonable performance in predicting heart attacks.

Decision trees achieved an accuracy of 77%, which was slightly lower than other models. However, decision trees can provide interpretability and insights into the feature importance for predicting heart attacks.

In addition to the choice of algorithms, pre-processing techniques played a significant role in enhancing the models' performance:

●Data cleaning: We performed data cleaning, including handling missing values and outliers. This step helped improve the accuracy of the models by ensuring the data's quality and consistency.

●Feature Scaling: We applied feature scaling using the Standard Scalar.

By standardizing the numerical features to have zero mean and unit variance, we ensured that all features contributed equally to the models' decision-making process.

In conclusion, our study demonstrates that machine learning techniques can effectively analyse and predict heart attacks. Random Forests emerged as the top-performing algorithm, providing a balance between simplicity and accuracy. Pre-processing techniques, including data cleaning,

Feature scaling, and feature selection, played a crucial role in improving model performance.

The visualization of the data, such as scatter plots, helped us gain insights into the relationships between different features and the occurrence of heart attacks. By visualizing the data, we could identify potential patterns and correlations that informed our modelling decisions.

It is important to note that our study is based on the Heart Attack Analysis and Prediction Dataset.

Therefore, the generalizability of the findings should be exercised with caution when applying them to different datasets or populations. Further research can focus on refining the models, exploring additional algorithms, and incorporating domain knowledge to enhance the accuracy and interpretability of the predictions.

Overall, our study contributes to the field of cardiovascular health by providing insights into the predictive factors for heart attacks. By understanding and predicting heart attacks accurately, we can potentially improve prevention and intervention strategies to reduce the burden of cardiovascular diseases.

References

[1]"Heart Attack Rates According to World Health Organization", Indy Türk, Available: https://www.indyturk.com/node/545246

[2]J. Santhana Krishnan; S. Geetha , "Prediction of Heart Disease Using Machine Learning Algorithms," in 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), Chennai, India 21 June 2019.

[3]A. Ritu, P. Prajoy K. Aditya, "ECG Classification and Analysis for Heart Disease Prediction Using XAI-Driven Machine Learning Algorithms," Part of the Intelligent Systems Reference Library book series, April 2022.

[4]R. Rashik , "Heart Attack Analysis & Prediction Dataset," Kaggle, Available:

https://www.kaggle.com/datasets/rashikrahmanpritom/heart-attack-analy

sis-prediction-dataset?datasetId=1226038&sortBy=voteCount.

[5]D. Sena Merter, "What is a Standard Scaler?" Data Science School, Available:

https://www.veribilimiokulu.com/veri-hazirliginin-vazgecilmezi-ozellik- olceklenen/.

[6]"Logistic Regression in Machine Learning," Javatpoint, Available: https://www.javatpoint.com/logistic-regression-in-machine-learning.

[7]B. Jason, "Logistic Regression for Machine Learning Details," Machine Learning Mastery, Available: https://machinelearningmastery.com/logistic-regression-for-machine-lear ning/.

[8]"Logistic Regression for Machine Learning Image," Spark By Examples, Available:

https://sparkbyexamples.com/wp-content/uploads/2023/03/Screenshot-2 023-03-11-at-4.19.28-PM. png.

[9]"Support Vector Machine Algorithm," Java point, Available: https://www.javatpoint.com/machine-learning-support-vector-machine-a algorithm.

[10]G. Rohith, "Support Vector Machine — Introduction to Machine Learning Algorithms," Towards Data Science, Available: https://towardsdatascience.com/support-vector-machine-introduction-to- machine-learning-algorithms-934a444fca47.

[11]"Support Vector Machine Artwork," Datatron, Available: https://datatron.com/wp-content/uploads/2021/05/Support-Vector-Machi ne.png.

[12]R. Surithi, "Understand Random Forest Algorithms with Examples," Analytics Vidhya, Available: https://www.analyticsvidhya.com/blog/2021/06/understanding-random-f orest/.

[13]Y. Tony, "Understanding Random Forest," Towards Data Science, Available: https://towardsdatascience.com/understanding-random-forest-58381e060 2d2.

[14] “Random Forest Image” Available: https://encrypted-tbn0.gstatic.com/images?q=tbn

[15] “Decision Trees” Avaliable: https://scikit-learn.org/stable/modules/tree.html

[16] “Decision Tree Terminologies” Avaliable: https://www.geeksforgeeks.org/decision-tree/

[17] “Decision Tree Algorithm in Machine Learning” Available: https://static.javatpoint.com/tutorial/machine-learning/images/decision-tr ee-classification-algorithm.png

[18]R. Rashik , "Heart Attack Analysis & Prediction Dataset," Kaggle, Available:

https://www.kaggle.com/datasets/rashikrahmanpritom/heart-attack-analy

sis-prediction-dataset?datasetId=1226038&sortBy=voteCount.

[19]K. Ajitesh “Accuracy, Precision, Recall & F1-Score – Python Examples” Avaliable:

https://vitalflux.com/accuracy-precision-recall-f1-score-python-example/

[20]K. Ajitesh “What is the presicion score?

Available: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.precisio n_score.html

[21]K. Ajitesh “Accuracy, Precision, Recall & F1-Score – Python Examples” Available:

https://vitalflux.com/accuracy-precision-recall-f1-score-python-example/

[22]K. Rohit “What is F1score?”Avaliable: https://www.v7labs.com/blog/f1-score-guide#:~:text=F1%20score