THE UNIVERSITY OF SYDNEY

Project Report for ENGG2112

CardioCare: An Intelligent Heart Attack Predictor

Sithma Gunawardena, 510510006, Biomedical Engineering

Faculty of Engineering

May 26, 2023

Executive Summary

Project overview and Objectives

The project aims to use a cardiovascular dataset to create a machine learning system that predicts the likelihood of a patient experiencing a heart attack based on their personal information for both current status and after 20 years. Early detection and prevention of cardiovascular diseases have the potential to save lives and reduce the healthcare system's burden. By implementing this model can have significant advantages for the medical industry and society, offering physicians a valuable tool to identify high-risk patients, prioritize interventions, and personalize treatment strategies. This has the potential to reduce heart attacks and improve patient outcomes.

Methods

Data normalization was conducted for all classification and regression models. Various evaluation measures, including accuracy score, area under curve, and precision score for classification models. In addition, a feature simulator function was developed to estimate the likelihood of a heart attack occurrence in the next 20 years and the current risk. Furthermore, cross-validation techniques were applied to evaluate the models' ability to generalize and learn from new data.

Findings

Among the classification models evaluated, KNN demonstrated the best fitness, achieving an average accuracy score of 0.8955 and a precision score of 0.8837. On the other hand, Naive Bayes performed the worst. While the linear regression model showed a higher Area Under Curve (AUC) score, KNN maintained consistent performance across multiple evaluation metrics. Thus, making KNN the preferred choice for predicting the likelihood of patients experiencing a heart attack.

Potential for Wider Adoption

Despite initial barriers, plans are in place to develop sophisticated models, improve feature engineering, and enhance the user interface. The technology, currently used in healthcare, has potential across other industries requiring predictive analytics. While AI's growth in healthcare is slower than expected, it is anticipated to bring about significant benefits in diagnostics and treatment. We foresee our technology, post improvements, being widely adopted and potentially commercialized, benefiting the medical sector and society at large.

Conclusion

In conclusion, KNN demonstrates the best performance in accurately predicting the likelihood of patients experiencing a heart attack in the next 20 years and their current risk.

Contents

1 Background

The Heart Attack Analysis & Prediction Dataset provides values of multiple health factors and personal details to reflect an individual's health status within the targeted population.

- Personal details include:
  - Age
  - Sex
- Health factors include:
  - exercise-induced angina
  - number of major vessels
  - Chest pan type
  - Resting blood pressure
  - Cholesterol fetched via BMI sensor
  - Fasting blood sugar
  - Resting electrocardiographic results
  - Maximum heart rate achieved
  - Have or have not experienced a heart attack

This dataset is closely relevant and useful for our project, it provides most of the information required for our machine learning system to be trained. Our project aims to train a cardiovascular dataset into a machine learning system to predict a patient's heart attack percentage based on their input including both their details and personal health factors.*[1]*

A credible and reliable source should consist of a well-respected publisher or an expert, unbiased analysis of the topic, the currency and the relativity to your project.*[2]* The Heart Attack Analysis & Prediction Dataset was published By Rashik Rahman. Rahman achieved a bachelor's degree in computer science and an engineering degree from the University of Asia Pacific as a lecturer. He also achieved many professional certifications including IBM data science, advanced machine learning on Google Cloud etc. Who also published journals and contributed to many academic papers. This indicates the dataset has a credible author.*[3]* This quantitative dataset is number-based, which is countable and measurable for many uses. The nature of the quantitative dataset has discriminated bias as it is produced from first-hand experiments. This indicates the dataset is unbiased. This dataset was updated in 2021, thus indicating a relatively up-to-date dataset containing recent information that is much more meaningful to develop further research compared to old datasets. This dataset is highly relevant to our designed project as it provides all the information required for us to train our cardiovascular dataset into a machine learning system to predict a patient's heart attack percentage based on their personal details input. Thus, all of the above elements contribute to the credibility of the source.

The Heart Attack Analysis & Prediction Dataset is highly reliable not only because it is published by a well-known publisher, but has been downloaded more than a hundred thousand times for further research purposes. It also has been labelled by the Kaggle data website with a score of 10/10 for its completeness, credibility and compatibility. Thus, all of the above elements contribute to the reliability of the source. Overall the Heart Attack Analysis & Prediction Dataset is highly credible and reliable as a source, it is also highly relevant to our group project. *[1]*

2 Problem Statement

According to research, Cardiovascular disease (CVDs) is the leading cause of global mortality, resulting in 17.8 million deaths annually. Due to the lack of early-stage symptoms, detecting CVDs is extremely challenging, making early detection crucial in reducing fatalities.

3 Objectives, Motivation

Cardiovascular disease (CVD) contributes to the loss of 17.9 million lives annually. CVD encompasses a range of heart and blood vessel disorders, including heart attacks and strokes. The primary cause of heart attacks and strokes is the obstruction of veins around the heart, preventing proper blood flow to the heart and brain. Unhealthy lifestyles, tobacco use, and physical inactivity are major factors leading to the accumulation of fatty deposits in blood vessel walls, resulting in blockages. Risk factors for CVD include elevated blood pressure, blood lipids, blood glucose levels, and obesity [6]. Detecting underlying heart attacks is challenging due to the lack of noticeable symptoms, making timely prevention difficult. Immediate medical attention is necessary when patients experience chest pain, breathing difficulties, lightheadedness, or faintness. Cardiovascular disease is the leading cause of global mortality, particularly in low and middle-income countries. These countries account for at least three-quarters of cardiovascular disease-related deaths due to inadequate primary healthcare programs for high-risk individuals [5]. This evidence emphasizes the significance of early detection and prevention to save lives and reduce fatalities.

Our project aims to develop a machine learning system trained on a cardiovascular dataset to predict the probability of a patient experiencing a heart attack based on their personal information both in the present and after a 20-year timeframe. Referring to the aforementioned research, early detection and prevention of cardiovascular diseases can have profound impacts, saving millions of lives and alleviating the burden on healthcare systems. By utilizing publicly available datasets, we aim to build a model that can estimate an individual's heart attack risk. Implementing this model would offer significant benefits to the medical industry and society at large. It would provide healthcare professionals with a powerful tool to identify high-risk patients, enabling timely interventions and personalized treatment strategies. This could result in a reduction in the number of heart attacks and improved patient outcomes [7]. Moreover, it would enhance individuals' awareness of their cardiovascular health and motivate the adoption of healthier lifestyles to mitigate the risk of heart attacks. This increased awareness would have a ripple effect on public health, leading to a decrease in heart attack incidence and associated medical costs [4]. Ultimately, this project will contribute to the community by saving lives in the medical field and fostering awareness of the importance of a healthy lifestyle to prevent cardiovascular disease at its roots.

4 Methodology

In this project, two classification models and a regression algorithm are utilized to predict the likelihood of patients experiencing a heart attack. By using a simulator function, the accuracy and precision of the algorithms were forecasted to predict the probability of an occurrence of a heart attack within the next 20 years and its present risk.

4.1 Data Pre-Processing

Normalisation

In the dataset, an observation was made that the features have different scales, which can lead to biased results in algorithms like KNN and WNN. In order to eliminate bias, it was necessary to scale all features on a similar scale. This step helped reduce bias and improve the performance of distance-based algorithms.

Firstly, import the pandas library to read the dataset. Then, using the MinMaxScaler() and fit_transform() function from scikit-learn to transform the features of the dataset to a common range, between 0 and 1. In mathematical terms, it is defined by the formula: normalize_data = (current_data - min)/(max - min). Lastly, convert resulting normalized data back to a DataFrame by using pd.DataFrame().

4.2 Feature Extraction

Feature extraction is also crucial. Firstly, divide the dataset into target (y) and features (X). In this example, the remaining columns serve as the characteristics, and the variable "output" serves as the target variable, denoting the likelihood of a heart attack.

4.2.1 KNN & WNN Feature Extraction

In the KNN and WNN algorithms, the target variable (y) represents the outcome of the test and can take two values: 0 and 1. A value of 0 indicates a lower chance of a heart attack, while a value of 1 indicates a higher chance of a heart attack.

The features (x) include age, sex, exercise-induced angina, number of major vessels, chest pain type, resting blood pressure, cholesterol level, fasting blood sugar, resting electrocardiographic results, and maximum heart rate, which are potential factors related to the occurrence of a heart attack.

4.2.2 Linear Regression Extraction

In the Linear Regression algorithm, "features" is the features used for machine learning as referenced, "target" is the representation of the output as aforementioned in the KNN and WNN feature extraction taking two values of 0 and 1. Identically, to the KNN and WNN feature extraction, the same features were utilized.

A test size of 0.2 was utilized, indicating that 20% of the dataset was allocated for testing, while the remaining 80% was allocated for training purposes. This was to ensure an appropriate evaluation of the model's performance and generalizability.

The random_state parameter in the train_test_split function was used to set the random seed for reproducibility. When a specific random_state value is provided, the data splitting process will be deterministic, meaning that it will produce the same results each time the code is run with the same random_state value.

4.2.3 Naive Bayes Extraction

Similarly to Logistic Regression, the Naive Bayes algorithm was also divided into training and testing sets, with training sets using 80% of the data and testing sets using 20%.

4.3 Simulation

4.3.1 KNN & WNN

In KNN and WNN, cross-validation was used to determine the best training-testing split value. By setting a range of split ratios, such as 0.78, 0.77, 0.76, and 0.75, Then compute the average scores for each split ratio. The average scores obtained from cross-validation are as figure 4.4.1-a.

```
Split Ratio: 0.78, Average Score: 0.8390070921985815
Split Ratio: 0.77, Average Score: 0.8282146160962072
Split Ratio: 0.76, Average Score: 0.7913043478260869
Split Ratio: 0.75, Average Score: 0.8014492753623188
```

Figure 4.4.1-a

By comparing the average scores for different split ratios, The optimal training-testing split value that yields the highest average score is determined, which is 0.78. Thus, the training-testing split value in both two algorithms is 78-22.

After determining the training-testing split value using in the algorithm, the best value for k in classification is determined using a validation curve. It shows the accuracy of the model changes in a range of k values. It reveals that the optimal k value for both algorithms is found to be 7, when a 78-22 training-testing split is selected. The validation curve for KNN and WNN are figure 4.4.1-b and 4.4.1-c.

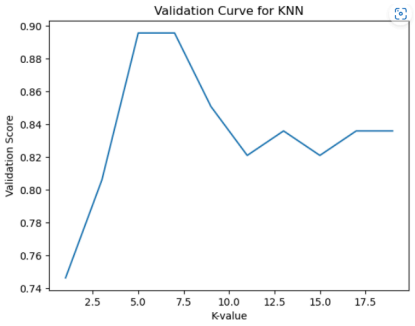


Figure 4.4.1-b

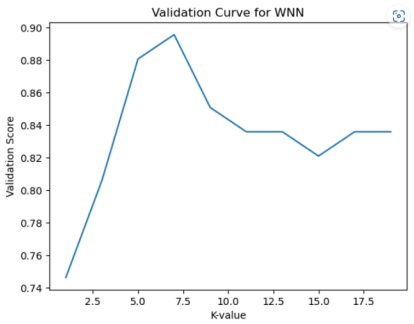


Figure 4.4.1-c

In the KNN simulations, the results are depicted in Figure 4.4.1-d, while in the WNN simulations, the outcomes are illustrated in Figure 4.4.1-e.

```
Accuracy Score:
0.8955223880597015

Precision Score:
0.8837209302325582

KNN AUC:
0.9101851851851852
```

Figure 4.4.1-d

```
Accuracy Score:
0.8955223880597015

Precision Score:
0.8837209302325582

WNN AUC:
0.9092592592592592
```

Figure 4.4.1-e

By comparing the performance metrics of accuracy and precision, both KNN and WNN exhibit similar scores. However, when considering the AUC metric, KNN outperforms WNN significantly, indicating that KNN is more effective in classifying the data. Consequently, based on the higher AUC score, KNN is considered the preferred choice for classification in our project.

Based on the information above, we have developed a high accuracy and precision KNN model. But, it only can predict the chance of having a heart attack right now. Based on this foundation, some innovations were made. As develop a simulate function that predicts the changes in patient features over the next 20 years in the absence of any health improvements. It will simulate the changes in each feature of the patients as shown in the figure 4.4.1-f.

```python
# Create a functions that can simulate the feachers change in the further 20 years.
def feature_simulater(s_data):
    simulated_data = s_data.copy()
    simulated_data.loc[0, 'age'] = simulated_data.loc[0, 'age'] + 1 # age
    if simulated_data.loc[0, 'cp'] >= 0.1:
        simulated_data.loc[0, 'cp'] = simulated_data.loc[0, 'cp'] - random.choice([0, 0.1, 0.2]) # Chest Pain type chest pain type
    simulated_data.loc[0, 'trtbps'] = simulated_data.loc[0, 'trtbps'] + 2 # resting blood pressure
    simulated_data.loc[0, 'chol'] = simulated_data.loc[0, 'chol'] + 3 # cholestoral in mg/dl fetched via BMI sensor
    simulated_data.loc[0, 'fbs'] = simulated_data.loc[0, 'fbs'] + random.choice([0, 0.1]) # fasting blood sugar
    simulated_data.loc[0, 'thalachh'] = simulated_data.loc[0, 'thalachh'] - 1 # maximum heart rate achieved
    if simulated_data.loc[0, 'exng'] < 1:
        simulated_data.loc[0, 'exng'] = simulated_data.loc[0, 'exng'] + random.choice([0, 0.05]) # exercise induced angina
    return simulated_data
```

figure 4.4.1-f

After determining that the patient currently has a lower chance of experiencing a heart attack, the function generates a simulated dataset representing the patient's features for the next year. It will be used to predict whether the patient has a high chance of developing heart attack in the year after, and this process will be repeated 20 times until a prediction of a high chance of heart attack is made for patients.

In the first simulation, a simulation dataset was created using features corresponding to a result indicating a high probability of having a heart attack, and the KNN prediction module produced results that matched the correct outcomes, as shown in the figure 4.4.1-g.

```
There's a high chance you'll have a heart attack now.
```

figure 4.4.1-f

In the second simulation, a simulation dataset was generated using features associated with a result indicating a lower probability of experiencing a heart attack. The KNN prediction module accurately predicted the outcome and indicated that the patient might have a heart attack after 5 years, as depicted in figure 4.4.1-h.

```
There's a small chance you'll have a heart attack now.
But you might have a heart attack after 5 years.
```

figure 4.4.1-f

In the third simulation. Firstly, a simulation dataset was generated using features associated with a result indicating a higher probability of experiencing a heart attack. Then, to align the features with a result indicating a low probability of having a heart attack, it is necessary to modify a subset of the features accordingly.The KNN prediction module accurately predicted the outcome and indicated that the patient might have a heart attack after 4 years, as depicted in figure 4.4.1-i.

```
There's a small chance you'll have a heart attack now.
But you might have a heart attack after 4 years.
```

figure 4.4.1-f

According to the predicted results above, KNN algorithm can accurately predict the correct outcomes. When the features of patients, who have a high chance of having a heart attack, are modified, KNN model still predicts that the patients may have a chance of having a heart attack after a few years.

4.3.2 Naive Bayes

A simulation using the Heart Attack Analysis & Prediction Dataset to assess the effectiveness of the Naive Bayes method for heart attack prediction. The dataset was divided into features and the target variable for preprocessing, with 80% of the data being used for training and 20% being utilised for testing.
By using the Python scikit-learn module to the training data to build a Gaussian Naive Bayes classifier. The Naive Bayes algorithm makes the assumption that every feature is distinct from every other feature and has a Gaussian distribution. The model learned the probabilities and correlations among the characteristics for predicting the likelihood of a heart attack by fitting the classifier to the training data.

Following the classifier's training, the use of testing data to make predictions and evaluated the classifier's performance using a number of measures. Out of all the test samples, an accuracy score is generated, which expresses the percentage of heart attack cases that were accurately predicted.
Additionally, in order to evaluate how well the classifier performed in differentiating between positive and negative instances, the plotted the Receiver Operating Characteristic (ROC) curve. The curve shows the trade-off for various classification thresholds between the true positive rate (sensitivity) and the false positive rate (1-specificity). The performance of the classifier is summarised by the Area Under the ROC Curve (AUC), which was also calculated. Better discriminating between positive and negative situations is indicated by a higher AUC score.

By evaluating the performance of the Naive Bayes algorithm for heart attack prediction by examining the simulation's outcomes, the accuracy score and the AUC value provide information about the classifier's discriminatory and predictive abilities. These results advance knowledge of how the algorithm might aid in the early detection and prevention of heart attacks.

4.3.3 Linear Regression

In the Logistic Regression model, the same training dataset was utilized, where independent variables (X) were employed to predict the target variable (y). The model generated probabilities for each sample, which were then compared to the actual target values to calculate the accuracy score.

This involved rounding the probabilities to obtain predicted classes (0 or 1) and determining the percentage of correct predictions. The precision score was also calculated, representing the accuracy of positive predictions, which was computed as the ratio of true positives to the sum of true positives and false positives.

Additionally, a Precision-Recall curve was obtained to assess the model's performance. The Precision-Recall curve depicts the trade-off between precision and recall (also known as sensitivity) for different classification thresholds. It provides insights into how well the model can identify positive instances while minimizing false positives. The curve is generated by calculating precision and recall values at various thresholds, allowing for a comprehensive analysis of the model's performance in differentiating between positive and negative instances.

The model's performance was then assessed using the Receiver Operating Characteristic (ROC) curve. By adjusting the classification threshold, the true positive rate (TPR) and false positive rate (FPR) were calculated for each threshold. This analysis provided insights into the model's ability to balance between true positive identifications and false positive errors.

Using the combined aforementioned data obtained, the overall performance of the Logistic Regression model was evaluated by calculating the Area Under the Curve (AUC) of the ROC curve.

The AUC value provides a measure of the model's predictive power, with a higher AUC indicating better performance in distinguishing between positive and negative instances. These steps allowed for a comprehensive assessment of the Logistic Regression model's effectiveness in predicting the likelihood of heart attacks based on the given input features.

Using the feature simulator function as before mentioned, the corresponding prediction of a patient suffering a heart attack in the following 20 years was printed.

5 Simulation Results

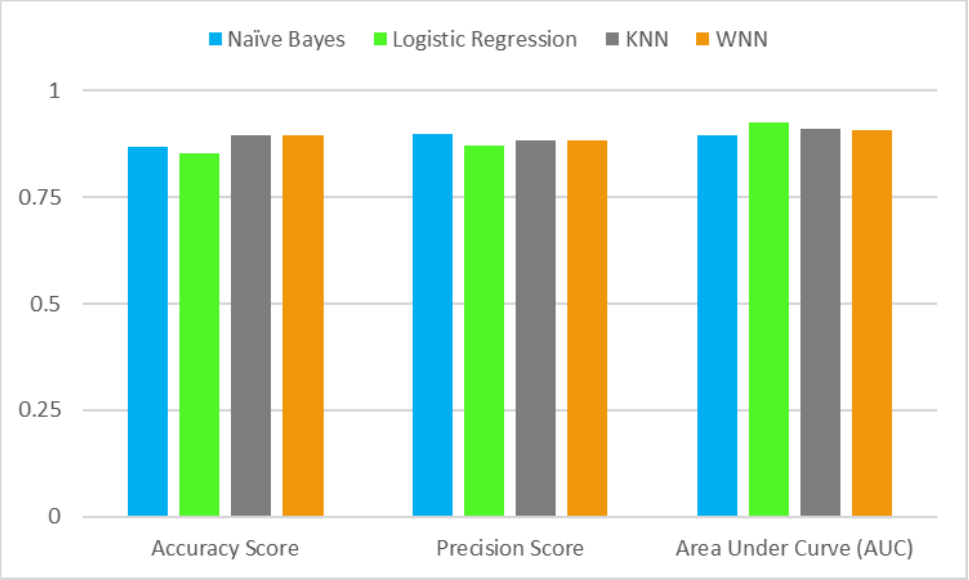
5.1 Key Findings and Significance

Comparing classification models

The table below shows the accuracy score, precision score and area under curve for each model tried. And the bar chart shows a comparison between each model.

Classification Models	Accuracy Score	Precision Score	Area Under Curve (AUC)
Naïve Bayes	0.8688	0.9	0.8943
Logistic Regression	0.8524	0.8709	0.9256
KNN	0.8955	0.8837	0.9101
WNN	0.8955	0.8837	0.9092

(Fig 2: The table of the classification models' performance measures)



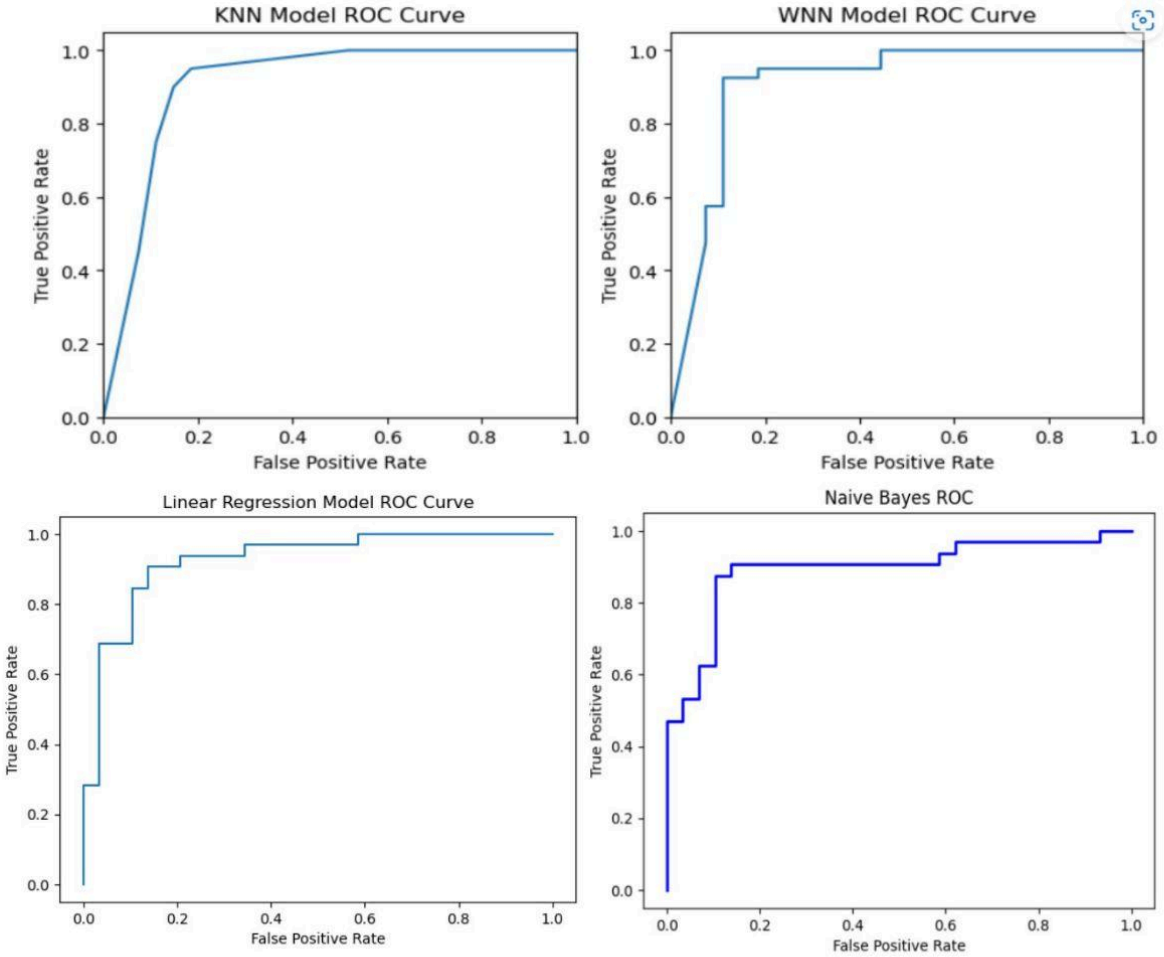
(Fig 3: The barchart comparison between classification models' performance measures)

As shown in Fig 2 and 3, the KNN model performs and shows the best fitness to predict the likelihood of patients experiencing a heart attack out of the other three classification models. Against all 4 models, KNN performs the most consistently, maintaining a high accuracy score of 0.8955, a precision score of 0.8837, and a AUC score of 0.9101. Though being bested by Logistic Regression's AUC score of 0.9256, KNN is able to maintain its overall performance across various evaluation metrics, making it the preferred choice for predicting the likelihood of patients experiencing a heart attack among the four classification models considered.

Meanwhile, the model that performed the worst in terms of overall performance was Naïve Bayes. It achieved an accuracy score of 0.8688, a precision score of 0.9, and an Area Under Curve (AUC) score of 0.8943. Although Naïve Bayes had respectable performance, it had lower accuracy, precision, and AUC scores compared to the other three models, indicating relatively weaker predictive ability in identifying the likelihood of patients experiencing a heart attack.

Naïve Bayes performed worse compared to the other three classification models due to its assumption of attribute independence. This assumption oversimplifies the relationships among the attributes, which may not hold true in real-world scenarios, particularly in predicting the likelihood of heart attacks. The model's inability to capture complex dependencies among the attributes during training contributes to its lower accuracy, precision, and AUC scores. While Naïve Bayes can still be competitive in certain cases, its performance suffers when faced with datasets where attribute independence is violated, leading to suboptimal predictions.

(Fig 4: The ROC Curve compare between classification models' true and false positive rate)



The ROC curve for each classification is plotted in Fig 4. as shown above. The KNN model demonstrates a steeper curve compared to the other three models, suggesting a higher true positive rate and improved performance in predicting the likelihood of heart attacks. This indicates that the KNN classifier operates more efficiently and accurately when utilizing the given input features for heart attack prediction.

5.2 Issues Faced

Naive Bayes 5.2.1

A few potential issues or challenges were encountered during the computation of conditional probabilities, specifically when certain feature values were absent from the training data for a given class, resulting in zero probabilities. To mitigate this problem, methods such as Laplace smoothing or other forms of regularization were employed to reduce its impact. In comparison to other algorithms, naive Bayes models may pose challenges in terms of interpretability. The assumption of feature independence makes it difficult to pinpoint and interpret the exact correlations between features that contribute to the predictions.

Linear Regression 5.2.2

Whilst very few problems were faced due to the simplicity of this algorithm, issues were met particularly in simulation. Integrating the feature simulator was difficult. Multiple instances of errors were met on incorrect

callings of functions which were promptly fixed. Incorrect outputs were also obtained with no indication of being incorrect unless all the bugs of the code were fixed. Resulting in multiple correct and incorrect attempts at fixing the code until eventually its output was verified after all remaining bugs were eliminated.

KNN & WNN 5.2.3

During this project, I met two challenges in the simulation part. Firstly, it was difficult to determine the split value. Initially,I used standard values such as 0.2 and 0.25, but the accuracy and precision scores of the result were not optimal. To address this issue, I used cross-validation, which helped in identifying a more optimally split value. Similarly, in the search for the best K-value, I faced a similar challenge, which was overcome by using a validation curve.

6 Potential for Wider Adoption

The results of this project are promising and suggest a strong potential for wider adoption. Our unique application of machine learning algorithms, specifically KNN, to the Heart Attack Analysis & Prediction Dataset has demonstrated competitive performance against other methods.

Current Market: The current market for this technology is primarily within the healthcare industry, specifically in the prediction and prevention of cardiovascular diseases. Our technology has already shown its effectiveness in predicting the likelihood of heart attacks based on various health factors and personal details.

Potential Market: The potential market for our technology extends beyond the healthcare industry. With further improvements and adjustments, this technology could be applied in various other industries where predictive analysis is needed.

Barriers to Adoption: the assumption of attribute independence in the Naïve Bayes model, which led to suboptimal predictions.

Strategies for Overcoming Barriers: To overcome these barriers, future work might include the development of more sophisticated machine learning models that can handle complex dependencies among the attributes. By resolving these issues, we believe we can develop a more accurate and reliable system that can be demonstrated to potential investors. These would include healthcare providers, insurance companies, and health tech startups.

Improvement:

- Advanced Machine Learning Models: We plan to explore more advanced machine learning models that can handle complex dependencies among the attributes. This could include deep learning models, which have shown excellent performance in various predictive tasks.
- Feature Engineering: We will also focus on feature engineering to improve the predictive power of our models. This could involve creating new features based on the existing ones, selecting the most informative features.
- User Interface and Experience: To make our system more accessible and user-friendly, we plan to develop a user interface where individuals can easily input their personal details and health factors. We also aim to improve the user experience by providing clear and understandable predictions.

Currently, there is lacking of commercial software available that can predict the likelihood of heart attacks. AI is being increasingly used in healthcare, but it is still at an early stage of development. According to a discussion between AI experts Fei-Fei Li and Andrew Ng, AI in healthcare is expected to progress slower

than desired over the next few years. They emphasized the importance of data-centric development and highlighted that quality, privacy, and availability of data pose unique challenges in healthcare settings. They also pointed out that AI could potentially help solve issues related to mental health, diagnostics, and operational aspects of healthcare. However, only 7% of hospitals' AI strategies are fully operational as of the time of the discussion, indicating that the field is still nascent *[8]*.

Mayo Clinic is an example integrating AI into its healthcare, notably in cardiovascular medicine. AI tools help diagnose heart conditions, hasten stroke treatments, and augment radiological diagnoses. One tool accurately identifies risk of left ventricular dysfunction 93% of the time. AI is used to rapidly and accurately analyze data, enhancing decision-making and treatments. They've also built a system that detects heart conditions and predicts future issues using AI's ability to learn from large data sets. Additionally, AI is used in emergency rooms for quick stroke diagnosis and in ECGs for heart failure prediction and atrial fibrillation detection.*[9]*.

With further development and the resolution of the issues encountered, we believe our technology has the potential for wider adoption and could even be commercialized, either through the creation of a startup company or by licensing the technology to an existing company. This could lead to substantial benefits for both the medical industry and society as a whole.

7 Conclusion

Our project showcased a notable application of machine learning models for heart attack prediction. Although certain challenges and deviations from our initial plan arose, as detailed in Section 5.2, the team consistently delivered effort and time towards its success.

The KNN model's superior consistency across key metrics made it the preferred model for heart attack prediction.
Naive Bayes' attribute independence assumption resulted in less optimal predictions, underscoring the significance of considering feature dependencies.
The Logistic Regression model, coupled with a feature simulator function, showed promising capabilities in predicting heart attack likelihood.

Future expansion of this project can explore more comprehensive data and advanced classification methods. Aiming for an accuracy rate closer to 95 percent could unlock commercialization potential, be it through a startup launch or technology licensing to an established company, thereby contributing substantially to healthcare predictive analytics.

References

[1] "Heart Attack Analysis & Prediction Dataset," *www.kaggle.com*. https://www.kaggle.com/datasets/rashikrahmanpritom/heart-attack-analysis-prediction-dataset?resource=download&select=o2Saturation.csv

[2] K. Marks, "Henry Buhl Library: Research Basics: How Do I Know If a Source Is credible?," *hbl.gcc.libguides.com*, 2023.

https://hbl.gcc.libguides.com/research/credible

[3] "Rashik Rahman | Department of Computer Science & Engineering - UAP," *cse.uap-bd.edu*.

https://cse.uap-bd.edu/faculty/faculty_details/43

[4] E. J. Emanuel, A. Glickman, and D. Johnson, "Measuring the Burden of Health Care Costs on US Families," *JAMA*, vol. 318, no. 19, p. 1863, Nov. 2017, doi: https://doi.org/10.1001/jama.2017.15686.

[5] World Health Organization, "Cardiovascular diseases," *World Health Organization*, Jun. 11, 2021. https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[6] World Health Organization, "Cardiovascular Diseases," *World Health Organization*, 2022. https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1

[7] C. Krittanawong, H. Zhang, Z. Wang, M. Aydar, and T. Kitai, "Artificial Intelligence in Precision Cardiovascular Medicine," *Journal of the American College of Cardiology*, vol. 69, no. 21, pp. 2657–2664, May 2017, doi: https://doi.org/10.1016/j.jacc.2017.03.571.

[8] Paul A, F. et al. (2023) Cardiovascular medicine, Mayo Clinic, 26 May 2023. Available at: https://www.mayoclinic.org/departments-centers/ai-cardiology/overview/ovc-20486648

[9] Press, G. (2021) *The future of AI in Healthcare*, *Forbes*, 26 May 2023. Available at: https://www.forbes.com/sites/gilpress/2021/04/29/the-future-of-ai-in-healthcare/?sh=3f7f9e39163b