

Predicting the Likelihood of a Heart Attack: A Machine Learning Approach

Neeharika Singh

September 23, 2023

Abstract

Heart diseases are a leading cause of mortality globally. Early detection of individuals at high-risk of heart attacks can substantially aid in implementing preventive measures. This paper presents a machine learning approach to predict the likelihood of heart attacks. The study adheres to the Knowledge Discovery in Databases (KDD) methodology and utilizes several machine learning algorithms for predictive modeling. The K-Nearest Neighbors (KNN) algorithm demonstrated the best performance, achieving an accuracy of approximately 87.9%.

1 Introduction

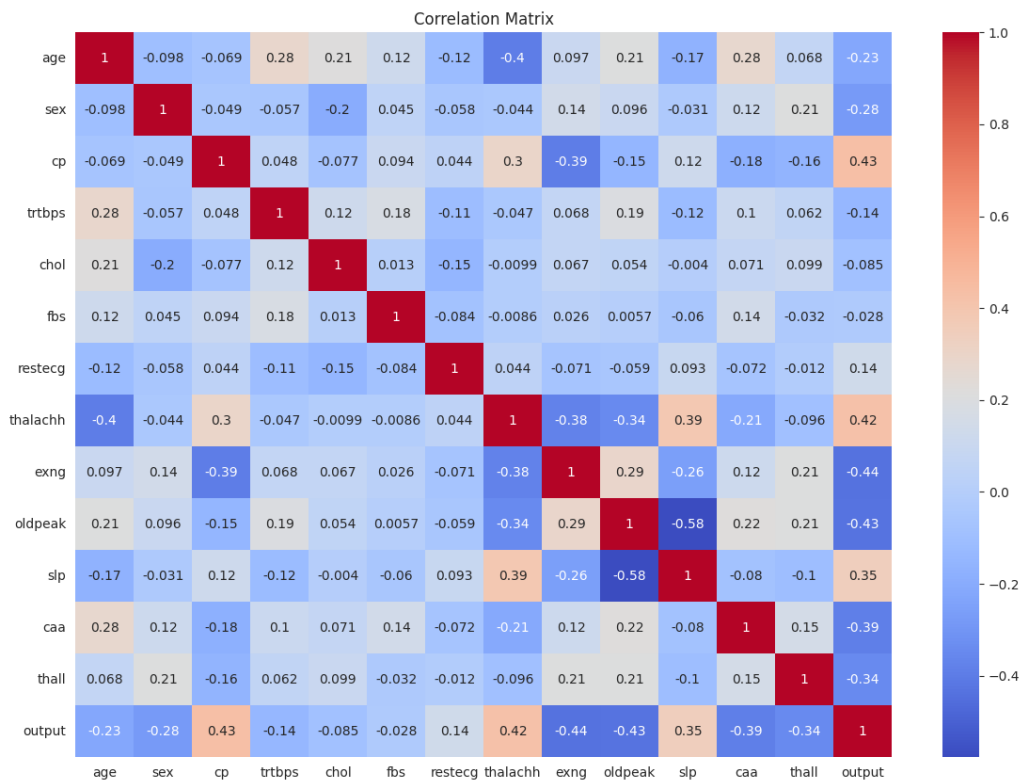
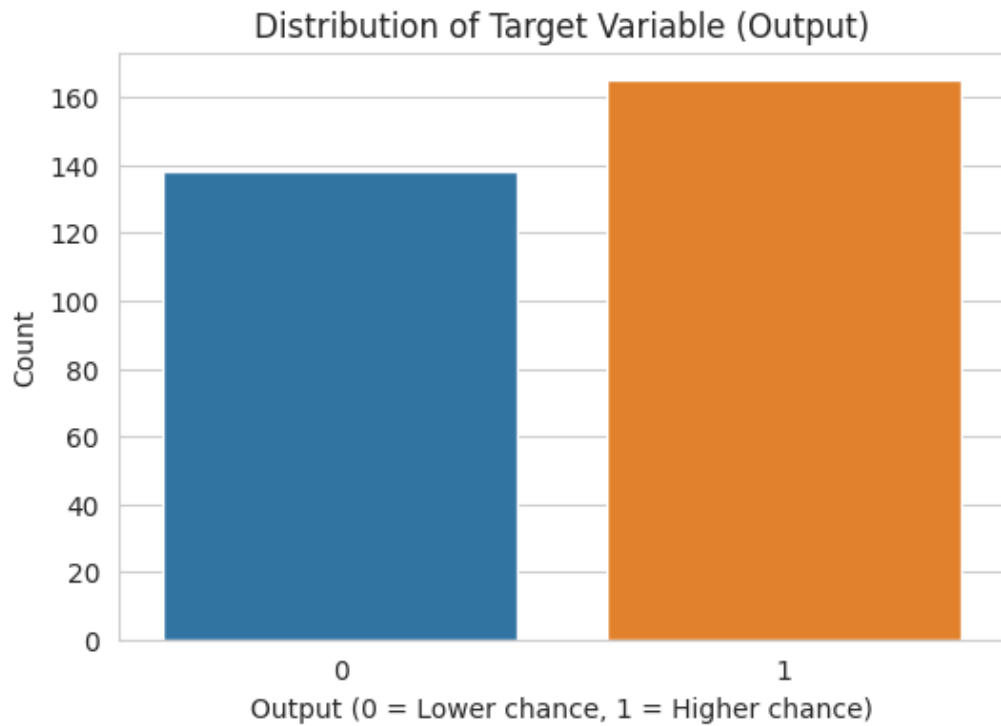
Cardiovascular diseases, particularly heart attacks, are among the leading causes of death worldwide. Preventive care is highly dependent on identifying high-risk individuals as early as possible. This study aims to apply machine learning algorithms to predict the likelihood of heart attacks based on various health metrics.

2 Methodology

The Knowledge Discovery in Databases (KDD) methodology was employed in this study. It encompasses several key steps including understanding the domain, data selection, data pre-processing, data mining, pattern evaluation, and knowledge presentation.

2.1 Dataset

The dataset consists of 303 samples with 13 features related to heart health and a target variable indicating the likelihood of a heart attack.



2.2 Data Preprocessing

Features were scaled to standardize their ranges, and the dataset was split into 70% training and 30% test sets.

2.3 Machine Learning Algorithms

Four machine learning algorithms were applied in this study:

- Logistic Regression

- Random Forest Classifier
- Support Vector Machine (SVM)
- K-Nearest Neighbors (KNN)

3 Results

The KNN model outperformed other models in terms of accuracy, precision, recall, and F1 score. The model achieved an accuracy of approximately 87.9%, a precision of 88.2%, a recall of 90%, and an F1 score of 89.1%.

4 Discussion

4.1 Insights

- Features like chest pain type, maximum heart rate, and slope of the peak exercise are strong indicators for a higher chance of heart attack.
- The dataset was fairly balanced, facilitating better model training.
- The KNN model showed high reliability and robustness.

4.2 Recommendations

- This predictive model can be used for early interventions in healthcare settings.
- Medical professionals can focus on specific features for more detailed risk assessment.
- The KNN model can be integrated into healthcare systems after further validation.
- The model should be continuously updated for real-world applications.
- Future models can incorporate lifestyle factors for improved prediction.

5 Conclusion

The study demonstrates the potential of machine learning algorithms in predicting the likelihood of heart attacks. The KNN algorithm, in particular, showed promising results. Continuous refinement and real-world validation are essential for practical applications.