




Data Science & Heart Disease

The background is a light pink color. In the top left corner, there are two blue-outlined leaf shapes. In the top right corner, there is a large, abstract shape in light blue and purple. In the bottom left corner, there is a large, abstract shape in light orange. In the bottom right corner, there are two pink-outlined leaf shapes. Several small, solid-colored circles (orange, red, and blue) are scattered across the background.

Introduction

Data Science plays a critical role in healthcare. This presentation explores how data science aids in heart disease prediction and prevention.



Overview of Heart Disease


Heart disease is a leading cause of death worldwide. Early detection and preventive measures are crucial in reducing its impact.



The Role of Data Science in Heart Disease

Data science allows for better prediction models, personalized treatment plans, and improved patient outcomes in heart disease.

Dataset Description :

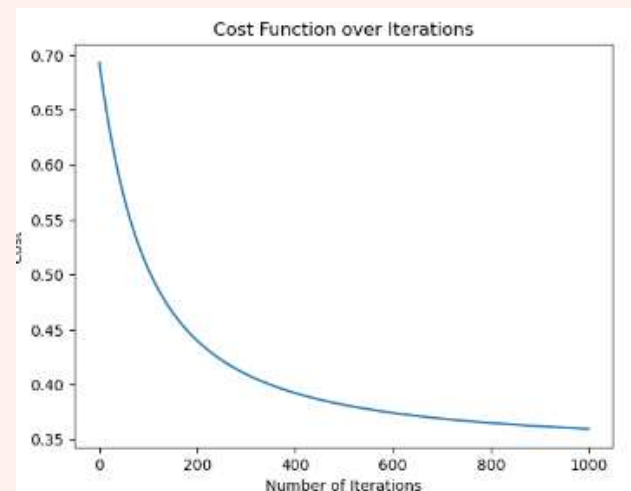


The dataset contains features such as cholesterol levels, age, and blood pressure that are used for model training.

Logistic Regression Overview

Logistic Regression is a classification model that predicts the likelihood of heart disease based on input features.

- Accuracy: 81.42%
- Confusion Matrix shown.

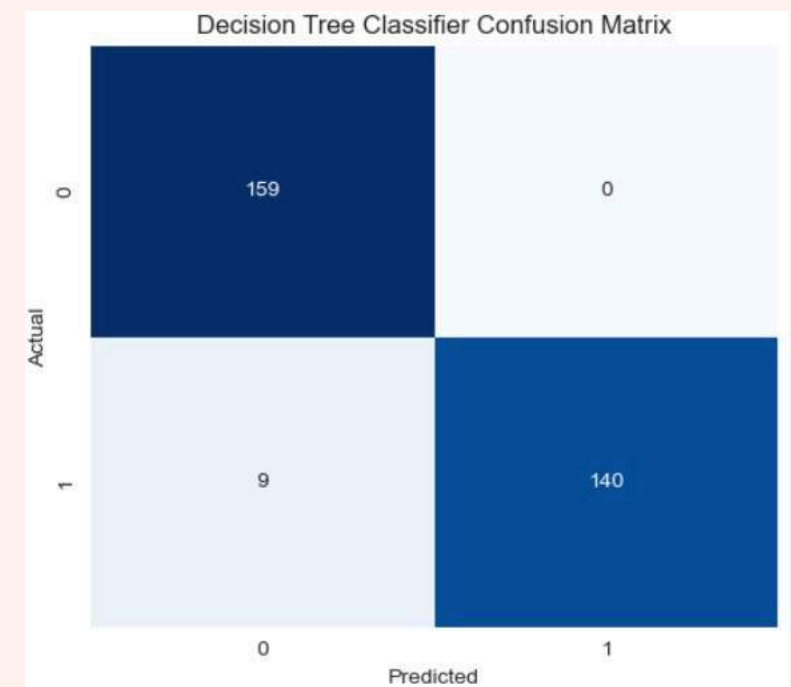


Logistic Regression Confusion Matrix		
Actual	0	1
0	119	40
1	20	129
	0	1
		Predicted

Decision Tree Overview

Decision Trees classify data by splitting it into branches, making predictions based on input features.

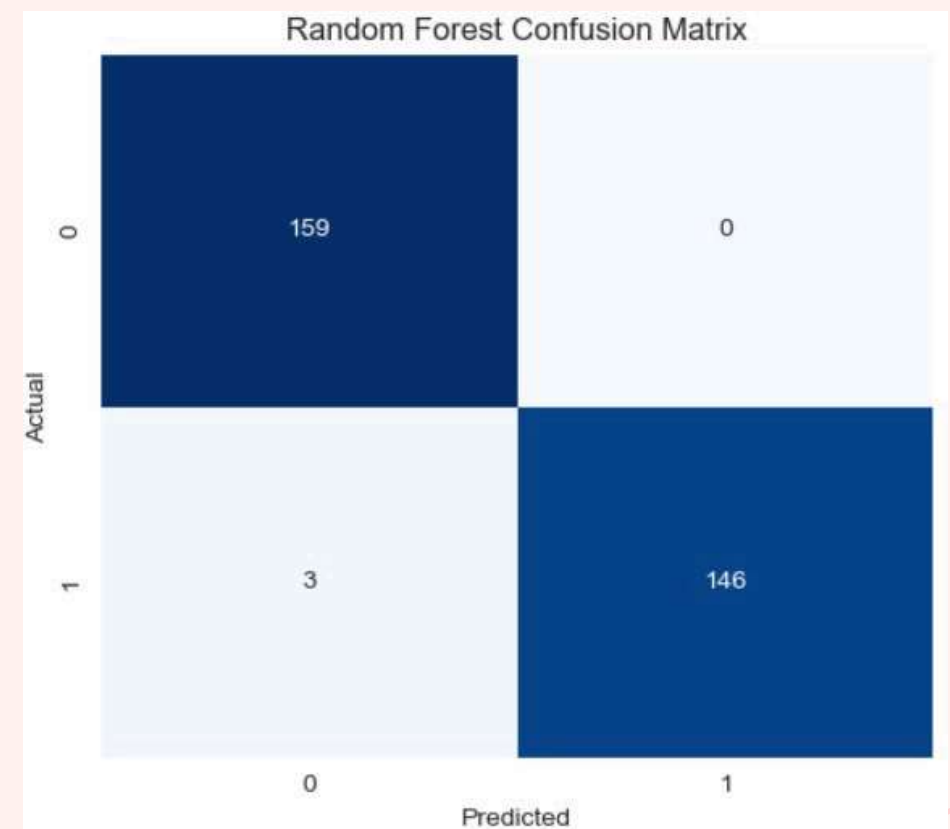
- Accuracy: 81.86%
- Confusion Matrix shown below.



Random Forest Overview

Random Forest is an ensemble learning method that builds multiple decision trees to improve model accuracy.

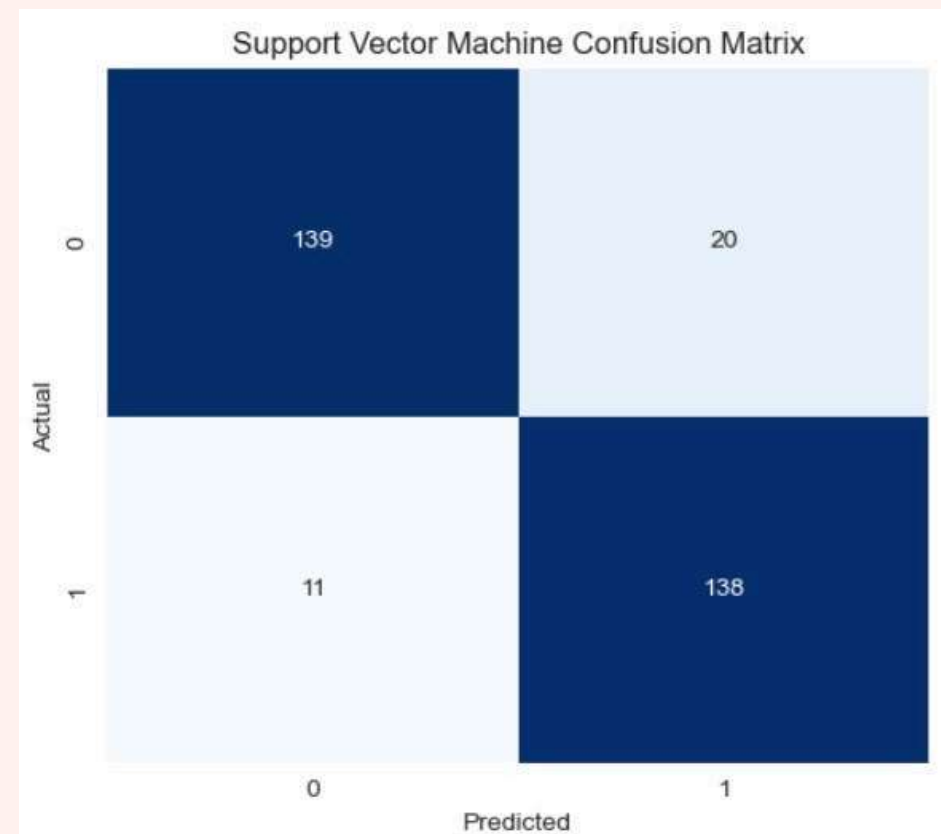
- Accuracy: 85.40%
- Confusion Matrix shown below.



Support Vector Machine (SVM)

SVM is a powerful model for binary classification that can handle both linear and non-linear data.

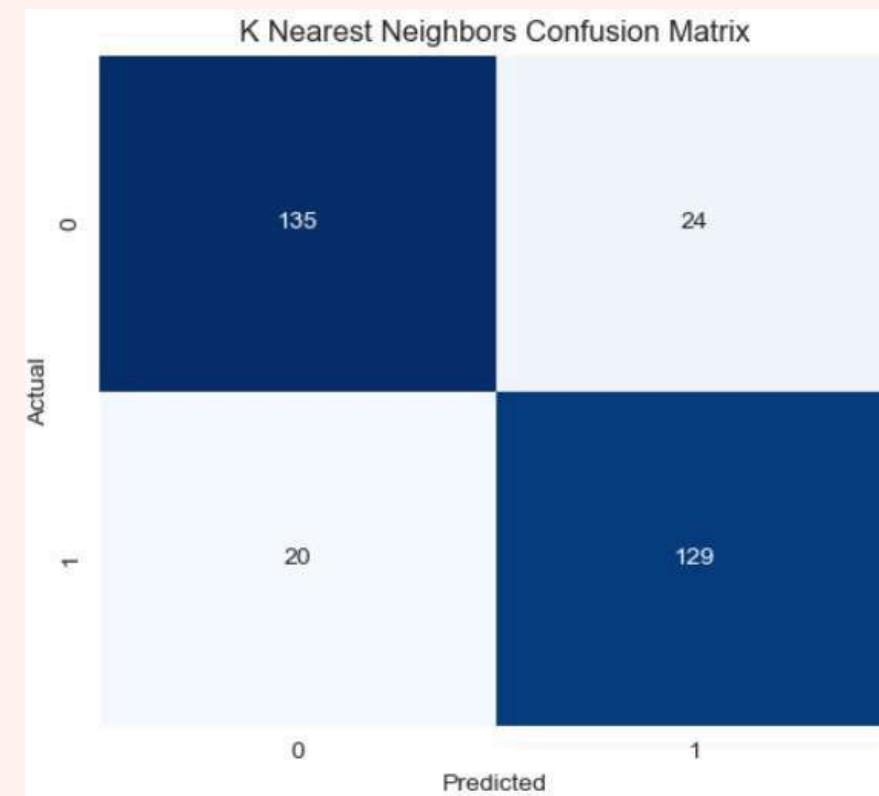
- Accuracy: 80.97%
- Confusion Matrix shown.



K-Nearest Neighbors (KNN)

KNN is a simple, instance-based learning algorithm used for classification and regression.

- Accuracy: 79.53%
- Confusion Matrix shown.



Naive Bayes

Naive Bayes is a probabilistic classifier based on applying Bayes' theorem with strong independence assumptions.

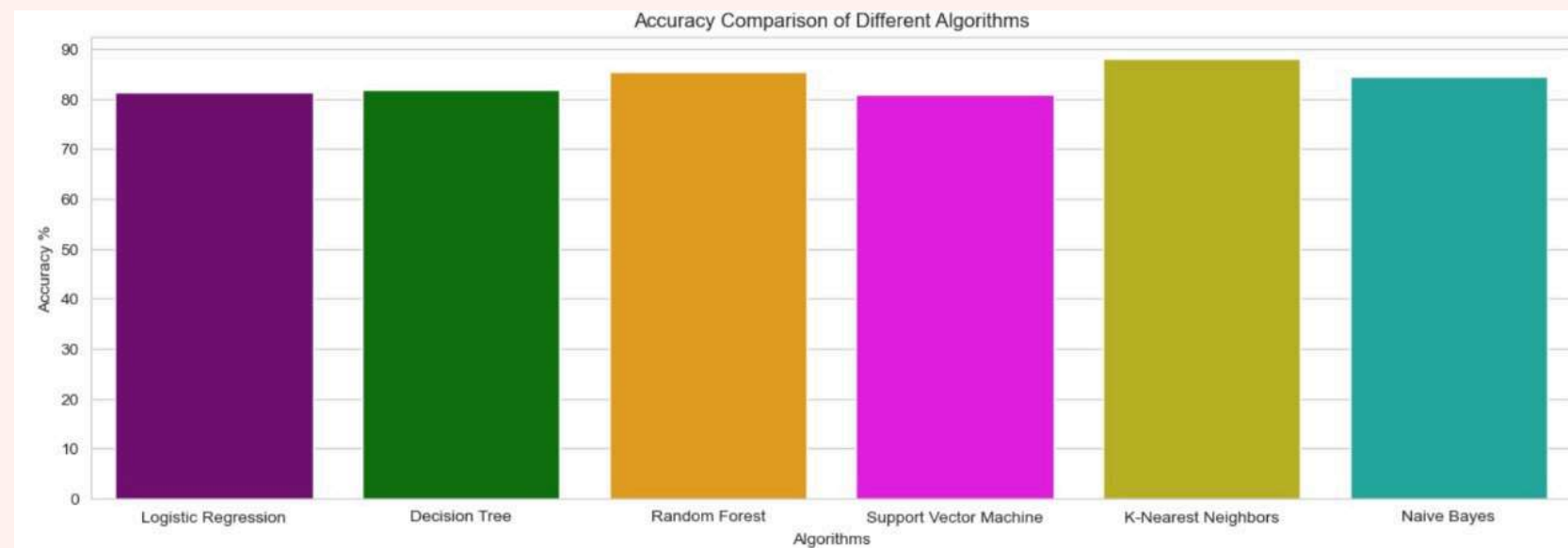
- Accuracy: 84.51%
- Confusion Matrix shown.

Naive Bayes Confusion Matrix

Actual \ Predicted	0	1
0	118	41
1	16	133

CONCLUSION

This study compared various models for heart disease prediction. The K-Nearest Neighbors model showed the highest accuracy.



The image features a central orange rounded rectangle containing the text "Thank You" in a dark brown, rounded, bubbly font. Below the text is a dark brown decorative flourish. The background is a light cream color with large, wavy shapes in light blue and light orange. In the top-left and bottom-right corners, there are illustrations of autumn leaves and branches in a dark brown line-art style.

Thank You