## Artificial

## Intelligence and Machine Learning

Project Report

Semester-IV (Batch-2022)

# HEART DISEASE PREDICTION

A red and white sign

Description automatically generated with low confidence

**Supervised By: Submitted By:**

Tushar Khitoliya PURAB (2210990692)

PRANAV (2210990667)

PRITAM (2210990683)

ARYAN (2210990174)

**Department of Computer Science and Engineering**

Chitkara University Institute of Engineering & Technology,

Chitkara University, Punjab

|  |  |  |
| --- | --- | --- |
| **S.NO** | **CONTENTS** | **PAGE NO** |
| **1** | **INTRODUCRION** | **1** |
| **2** | **PROCEDURE** | **2-3** |
| **3** | **FEATURES** | **4-5** |
| **4** | **CODE SNIPPETS** | **6** |
| **5** | **FUTURE SCOPE** | **7** |
| **6** | **CONCLUSION** | **8** |

**Introduction: Predicting Heart Disease using Machine Learning**

Heart disease remains one of the leading causes of mortality worldwide, emphasizing the importance of effective diagnostic tools and preventive measures. Leveraging the power of artificial intelligence and machine learning, this project aims to develop a predictive model capable of accurately identifying individuals at risk of heart disease based on their health metrics.

By utilizing a diverse dataset encompassing factors such as age, sex, cholesterol levels, blood pressure, and presence of diabetes, this project employs a range of machine learning algorithms including Logistic Regression, Support Vector Machines (SVM), Decision Trees, and Random Forest Classifiers. Through comprehensive exploratory data analysis, model building, and evaluation, this endeavor seeks to not only accurately predict heart disease but also provide insights into the most influential factors contributing to its onset.

Ultimately, the developed predictive model holds the potential to assist healthcare professionals in early identification and intervention, thereby improving patient outcomes and reducing the burden of heart disease on global health systems.

**PROCEDURE**

**1.Data Collection:**

Gathered a dataset containing various health metrics such as age, sex, cholesterol levels, blood pressure, presence of diabetes, etc., along with a label indicating whether the person has heart disease or not. Datasets for heart disease prediction are available from sources like Kaggle or UCI Machine Learning Repository.

**2.Exploratory Data Analysis (EDA):**

Used pandas to load and explore the dataset.

Checked for missing values and handle them appropriately.

Explored basic statistics of the dataset.

Visualized the distributions of features using matplotlib and seaborn.

**3.Data Preprocessing:**

Handled any missing or inconsistent data.

Encoded categorical variables if necessary.

Normalized or scaled the features if necessary.

Splitted the dataset into training and testing sets.

**4.Model Building:**

Implemented Logistic Regression, SVM, Decision Tree, and Random Forest Classifier using scikit-learn.

Trained each model on the training data.

**5.Model Evaluation:**

Evaluated each model's performance using metrics like accuracy, precision, recall, and F1-score.

Plotted confusion matrices for each model using matplotlib.

Compared the performance of different models.

**6.Hyperparameter Tuning:**

Used techniques like GridSearchCV to tune the hyperparameters of each model for better performance.

**7.Results Visualization:**

Visualized the performance of each model using matplotlib.

Plot the ROC curve and calculate the AUC score for each model.

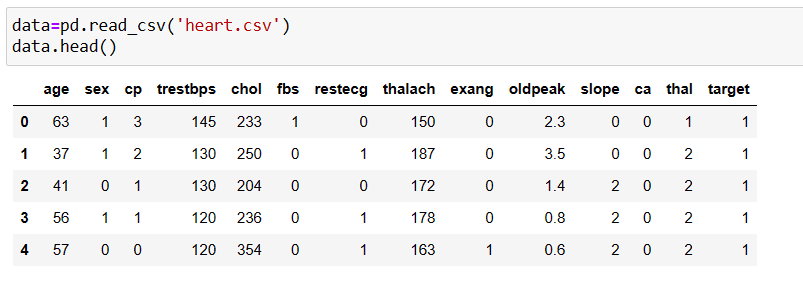
**FEATURES**

**We have use MATPLOTLIB AND SEABORN**

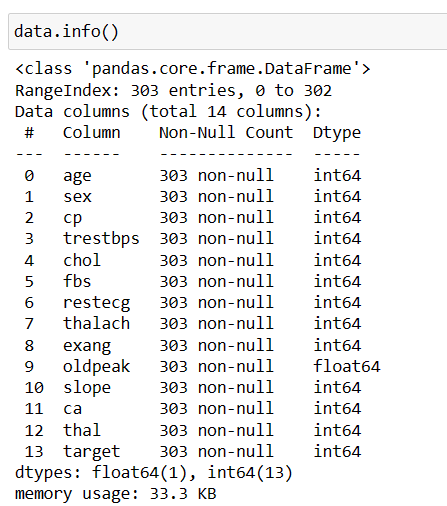
Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. It provides a wide range of functionalities for producing high-quality plots and figures, including line plots, scatter plots, bar charts, histograms, 3D plots, and more. Matplotlib is highly customizable, allowing users to fine-tune every aspect of their visualizations, such as colors, labels, axes, and annotations. It also integrates well with other Python libraries and tools, making it a popular choice for data visualization tasks in various fields, including scientific research, data analysis, and machine learning.

Seaborn is a Python data visualization library based on Matplotlib, offering a high-level interface for creating attractive and informative statistical graphics. It provides a simple yet powerful API for generating various types of plots, including scatter plots, line plots, bar plots, histograms, box plots, and heatmaps. Seaborn is designed to work seamlessly with Pandas DataFrames and arrays, making it easy to visualize data from different sources. One of Seaborn's key features is its ability to produce complex visualizations with minimal code, thanks to its built-in support for statistical estimation and automatic plot styling. It also offers advanced features for exploring relationships between variables, such as categorical data visualization and linear regression modeling. Seaborn is widely used in data analysis, research, and visualization tasks, particularly in fields like data science, social sciences, and biology.

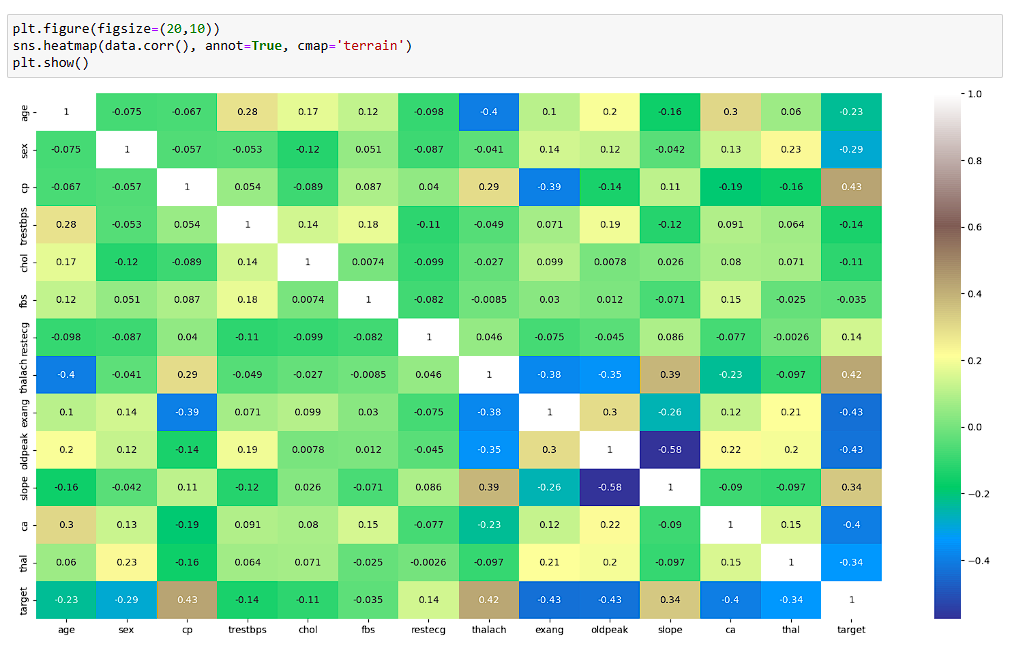
**Reading The Dataset**

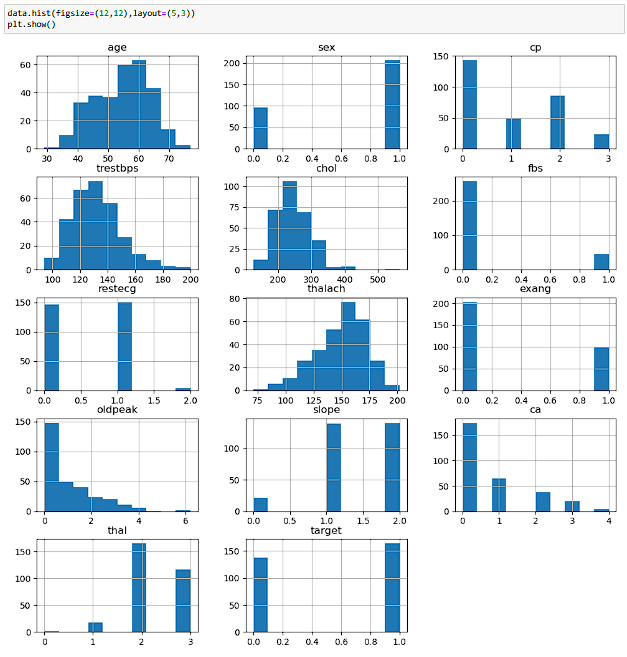


**Dataset Information**



**Some Project Snippets**





**FUTURE SCOPE**

The heart disease prediction project holds immense potential for real-world applications in the healthcare sector. The developed machine learning models can be integrated into clinical decision support systems, enabling healthcare professionals to quickly and accurately identify individuals at risk of heart disease. This early detection can lead to timely interventions, such as lifestyle modifications, medication management, or referrals to specialists, ultimately improving patient outcomes and reducing the burden on healthcare systems.

**CONCLUSION**

In conclusion, the heart disease prediction project has demonstrated the immense potential of artificial intelligence and machine learning in the field of healthcare. By developing accurate predictive models and gaining valuable insights into the key risk factors for heart disease, this endeavor has laid the foundation for a more proactive and personalized approach to cardiovascular health management. The project's findings can empower healthcare professionals, policymakers, and individuals to take informed actions, ultimately leading to improved patient outcomes and a reduction in the global burden of heart disease. As the field of AI and machine learning continues to evolve, the team remains committed to exploring new frontiers and pushing the boundaries of what is possible in the pursuit of a healthier and more resilient future.

**THANK YOU**