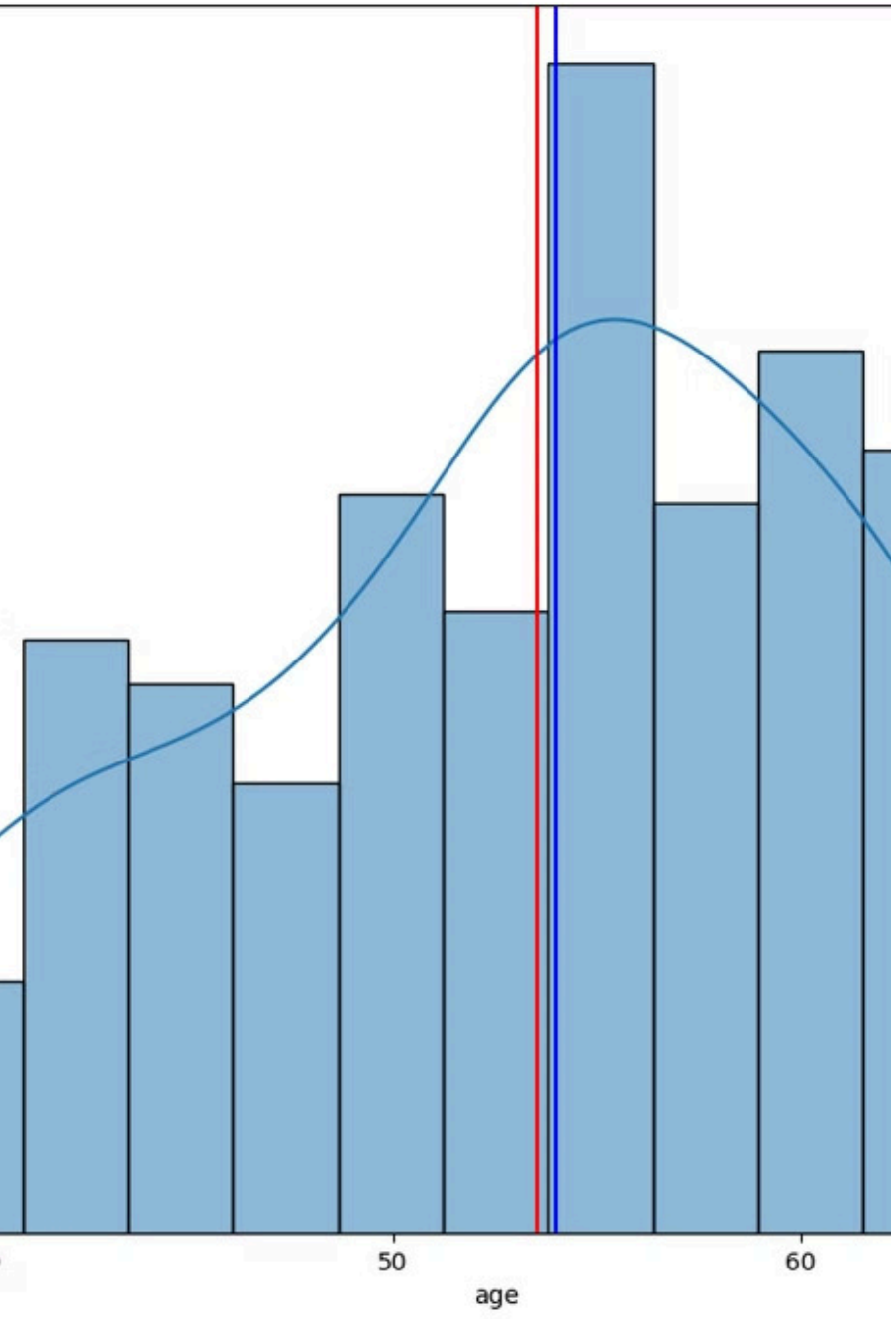
Heart Disease Prediction

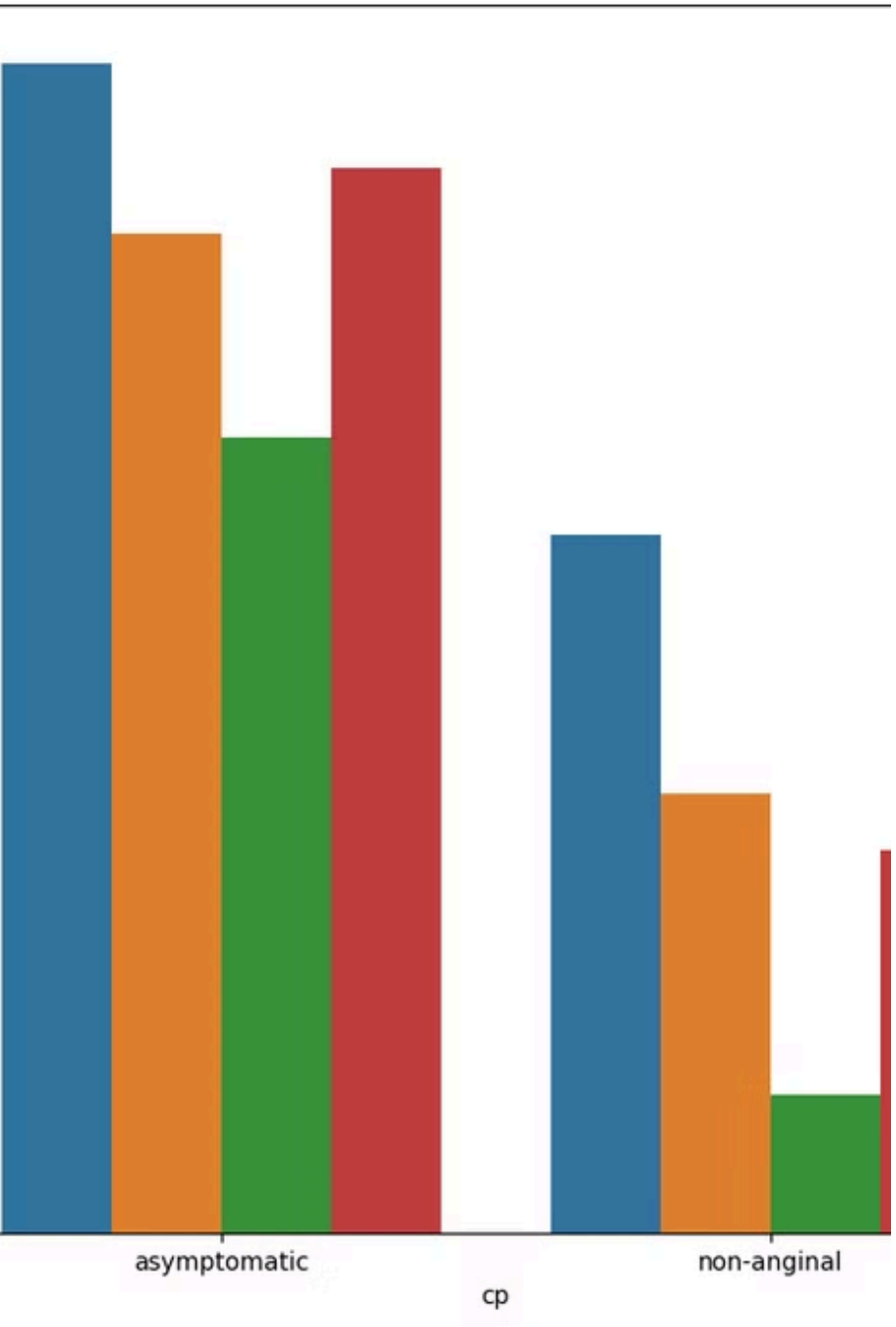
This presentation explores the process of building and deploying a machine learning model to predict heart disease. We will delve into the dataset, its features, and the steps involved in creating a robust and accurate predictive model.

by Adil naeem



Context

This dataset is a multivariate type, meaning it involves multiple variables. It contains 16 attributes, including age, sex, chest pain type, resting blood pressure, serum cholesterol, and more. The dataset was collected from various sources, including the Cleveland Clinic Foundation, the Hungarian Institute of Cardiology, and others. The primary goal is to predict the presence and severity of heart disease based on these attributes.

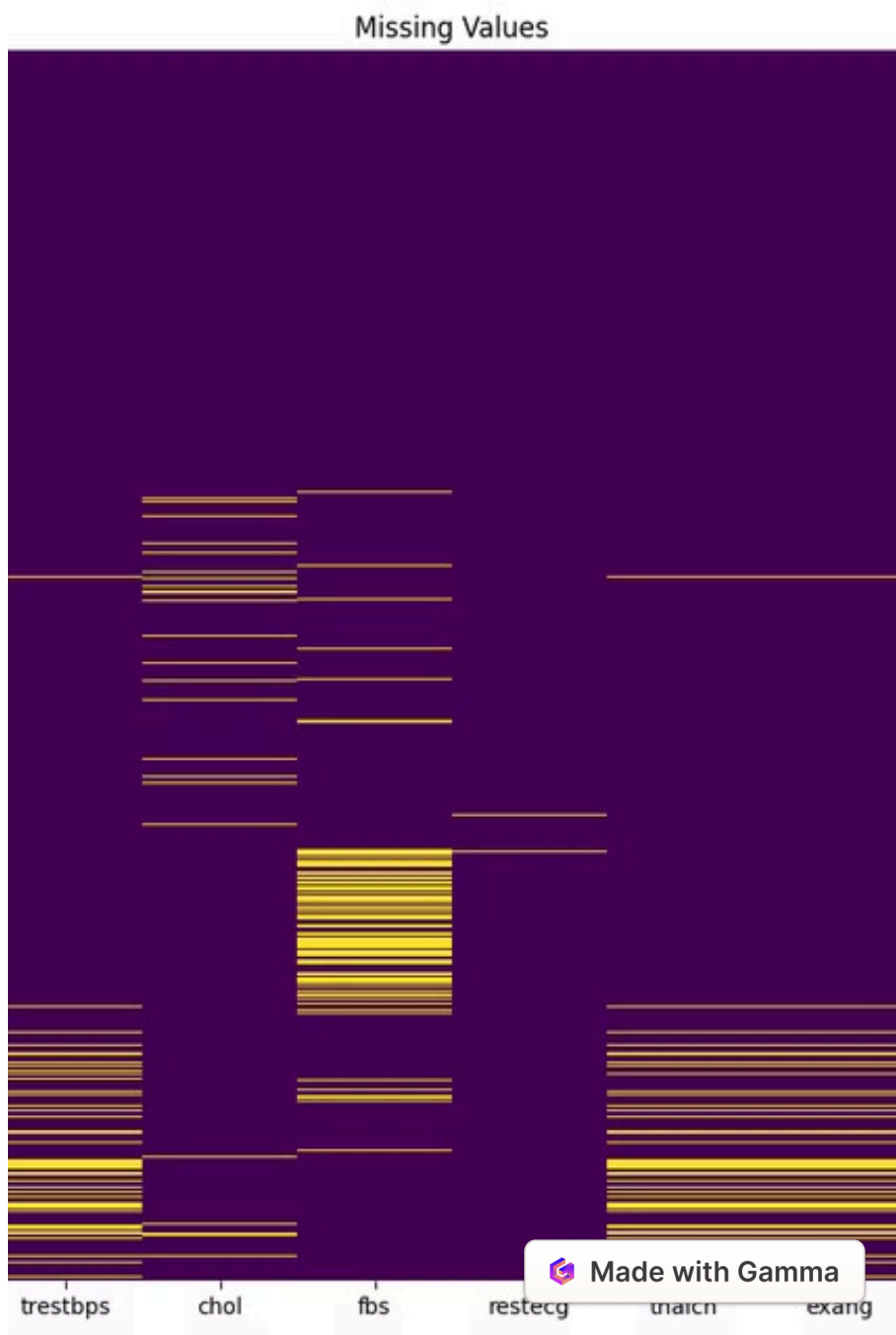


Exploratory Data Analysis (EDA)

The first step is to explore the dataset to understand its structure, patterns, and potential challenges. This includes examining the data types, identifying missing values, and visualizing the distribution of features. The goal is to gain insights that will guide the subsequent steps of data preprocessing and model building.

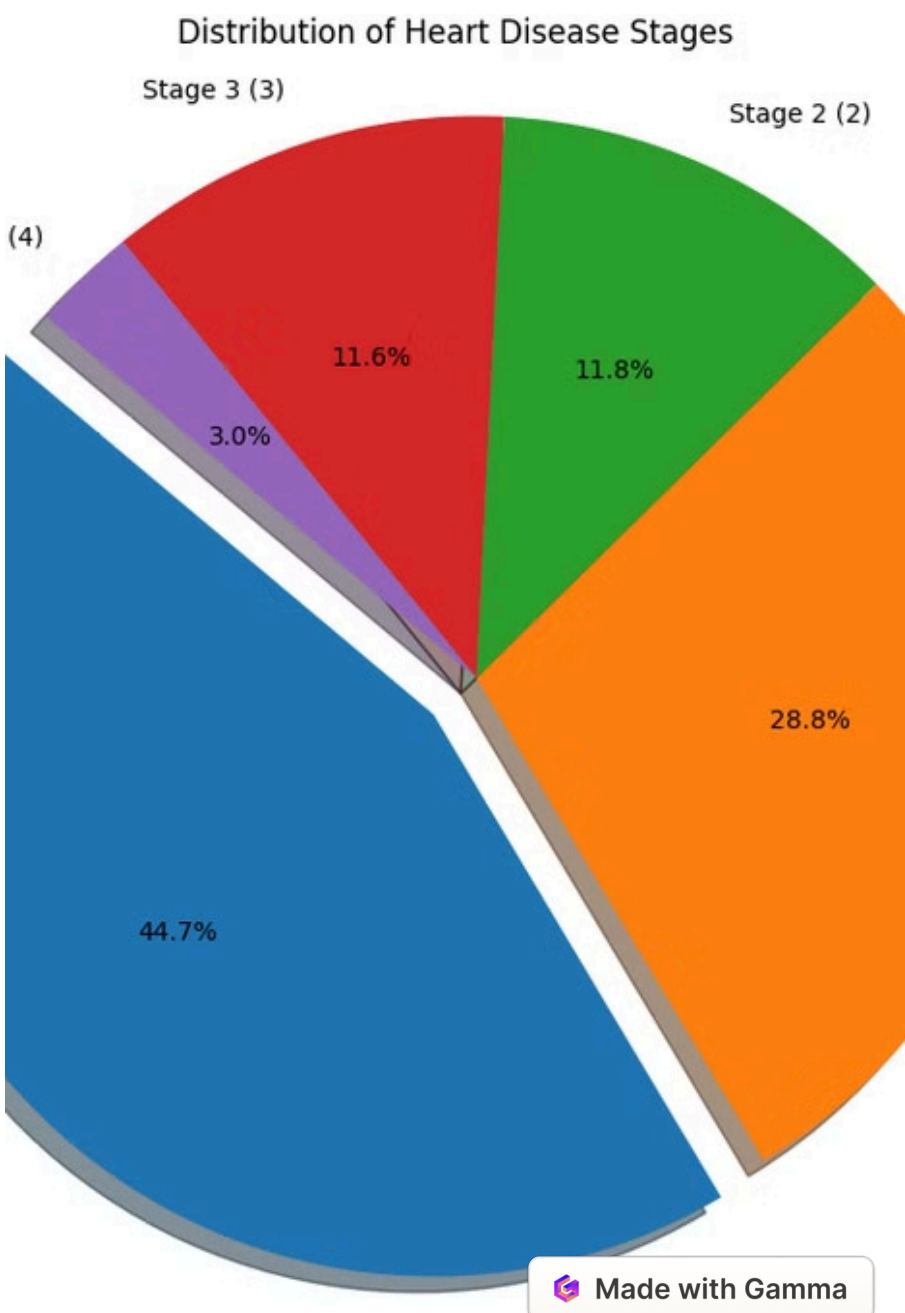
Dealing With Missing Values

The dataset contains missing values in several columns, such as 'trestbps', 'chol', 'fbs', etc. These missing values need to be addressed before building the model. We can use various imputation techniques, such as mean imputation, median imputation, or more advanced methods like KNN imputation or iterative imputation, to fill in the missing values.



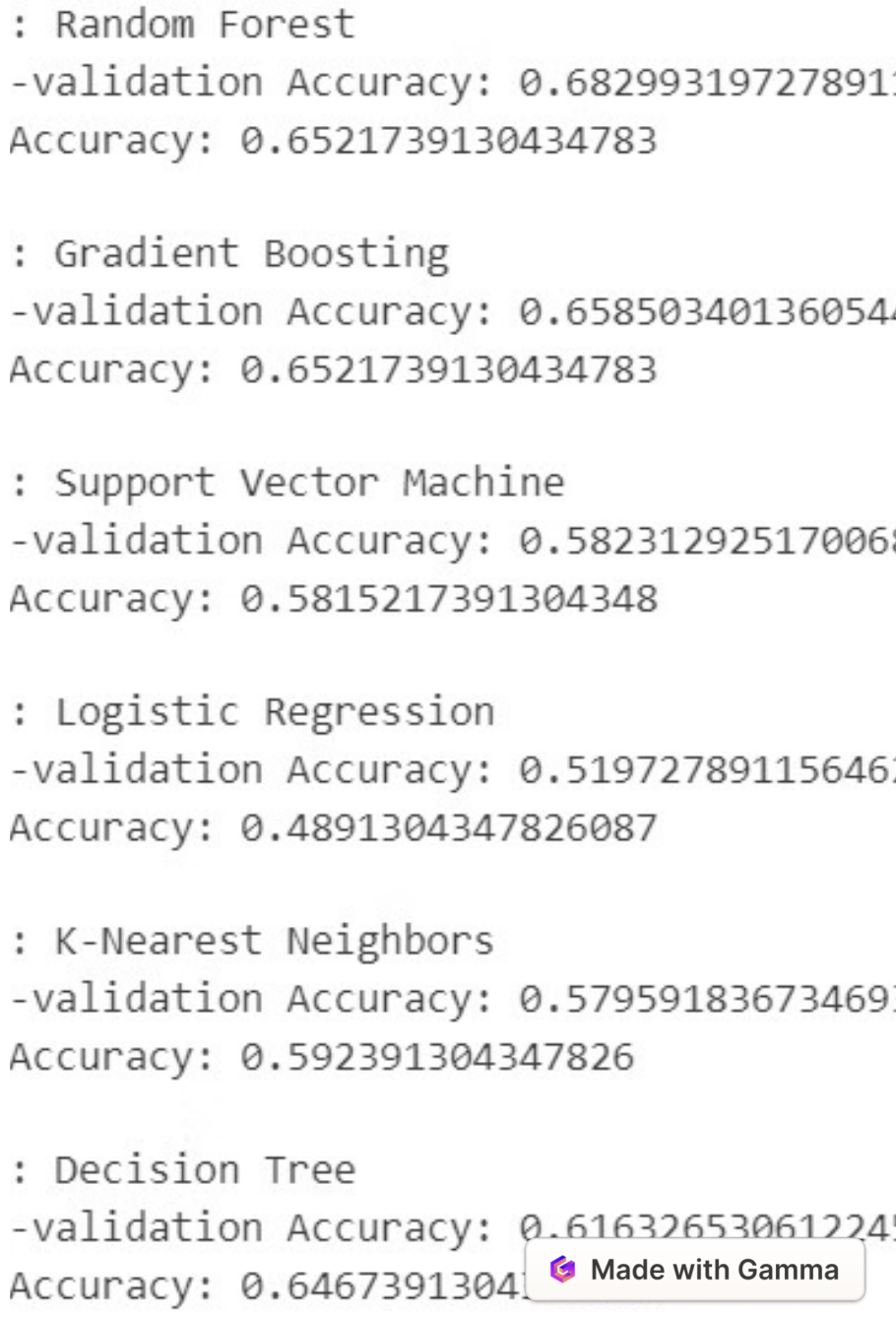
Feature Selection

Feature selection is crucial for building an effective model. We can use statistical tests like the Chi-Square test for categorical features and the ANOVA F-test for numerical features to determine the most relevant features for predicting heart disease. This helps us identify features that have a strong association with the target variable and prioritize them for model building.



Machine Learning

Once we have preprocessed the data and selected the most relevant features, we can start building the machine learning model. We will explore various algorithms, including Random Forest, Gradient Boosting, Support Vector Machine, Logistic Regression, and more. We will evaluate the performance of each model using cross-validation techniques and select the best-performing model based on metrics like accuracy, precision, and recall. The chosen model will then be saved for future use in making predictions on unseen data.



Flask Web App

I built a Flask web app to predict heart disease.

Users can enter specific parameters, like blood pressure and cholesterol levels, to get predictions.

The app uses a trained machine learning model to make these predictions. This project combines my knowledge of machine learning and web development. The app provides a tool for assessing heart disease risk.