This project aims to predict wine quality based on physicochemical properties using various machine learning models. The dataset used in this project includes features such as acidity, sulfur dioxide levels, density, pH, and alcohol content. By applying data preprocessing, exploratory data analysis (EDA), and multiple machine learning algorithms, we evaluate which model best predicts wine quality.

**Dataset**

The dataset used in this project is the Wine Quality dataset, which contains the following columns:

* **Features**:
  + fixed acidity
  + volatile acidity
  + citric acid
  + residual sugar
  + chlorides
  + free sulfur dioxide
  + total sulfur dioxide
  + density
  + pH
  + sulphates
  + alcohol
* **Target Variable**:
  + quality (an integer representing the wine quality)
* **Identifier**:
  + Id (unique identifier for each sample)

**Project Workflow**

1. **Data Preprocessing**:
   * Checked for missing values and corrected data types.
   * Standardized the features to ensure they are on a similar scale for better model performance.
2. **Exploratory Data Analysis (EDA)**:
   * Conducted summary statistics to understand feature distributions.
   * Created histograms, box plots, and a correlation heatmap to analyze relationships between features and detect outliers.
3. **Modeling**:
   * Trained and evaluated several machine learning models:
     + Logistic Regression
     + Decision Tree
     + Random Forest
     + Support Vector Machine (SVM)
     + Gradient Boosting
     + Neural Network (Multi-layer Perceptron)
   * Evaluated models based on classification metrics, including accuracy, precision, recall, and F1-score.
4. **Model Comparison**:
   * Compared models using a bar plot of accuracy scores to select the best-performing model for predicting wine quality.

**Results**

The models were evaluated based on accuracy and other classification metrics. The Random Forest and Gradient Boosting models performed best on this dataset, with high accuracy and balanced precision-recall.

* **Best Model**: Random Forest
* **Evaluation Metrics**:
  + Accuracy, F1-score, Precision, and Recall

**Acknowledgments**

* The Wine Quality dataset was sourced from the Kaggle Repository.
* This project was developed as a machine learning practice for predictive modeling.