Wine Quality Analysis

NANCY LOPEZ

Summary

The project involved a comprehensive analysis of key features in wines to create a predictive model for categorizing wine quality as either "good" or "bad."

- Alcohol Content
- Sulphates
- Citric Acid
- Total Sulfur Dioxide
- Volatile Acidity

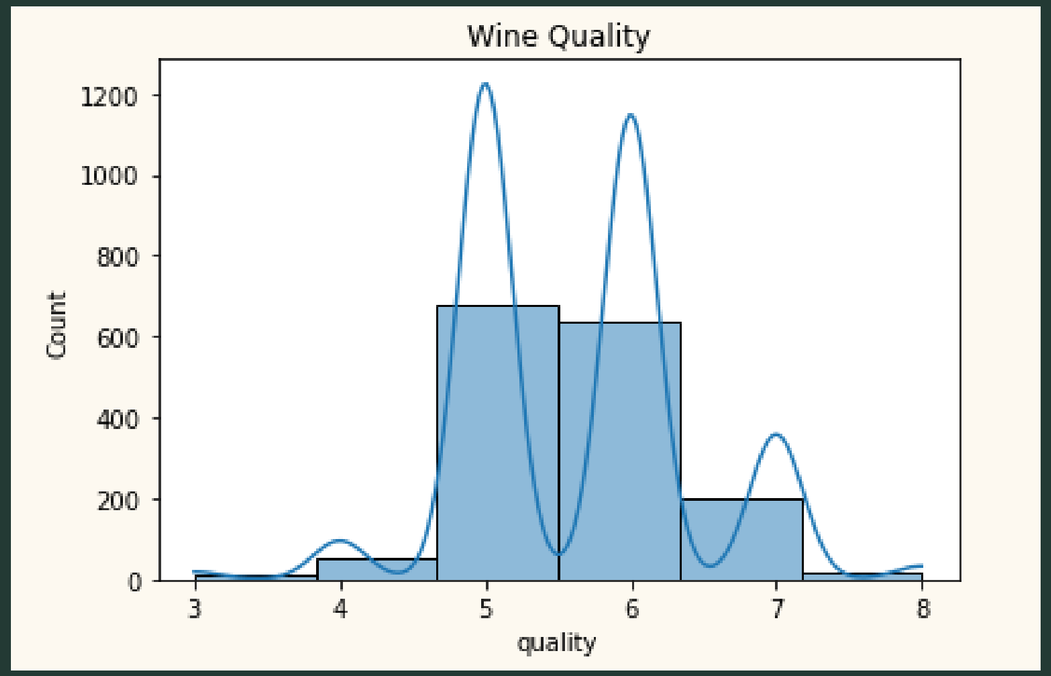
Outline

1. Business Problem
2. Data
3. The Approach
4. Selecting the Best Model
5. The Results
6. Recommendations

Business Problem

- Does the chemical composition of the wine, including levels of alcohol, acidity, sulfates, and other components could be key factors in quality assessment?
- Is there a way to predict if the wine is good or bad?

The Data



Red Wine Quality Dataset - Kaggle: contains data on the chemical composition of red wines, including alcohol content, sulphates, citric acid, total sulfur dioxide, and volatile acidity. The dataset also includes quality scores, with wines scoring 7 or above considered good quality.

The Approach

- Data Analysis
- Predictive Modeling
- Wine Quality Prediction
- Feature Importance

Selecting the Best Model

Our model selection process involved comparing the performance of several models:

Logistic Regression

Accuracy: 87.8%
Precision: 68%
Recall: 0.274
F1 Score: 0.391

Incorrect Predictions: 53

Support Vector Classifier

Accuracy: 88.3%
Precision: 87%
Recall: 0.884
F1 Score: 0.856

Incorrect Predictions: 51

K-Nearest Neighbors

Accuracy: 87.8%
Precision: 86%
Recall: 0.87
F1 Score: 0.862

Incorrect Perdictions: 58

Random Forest Classifier

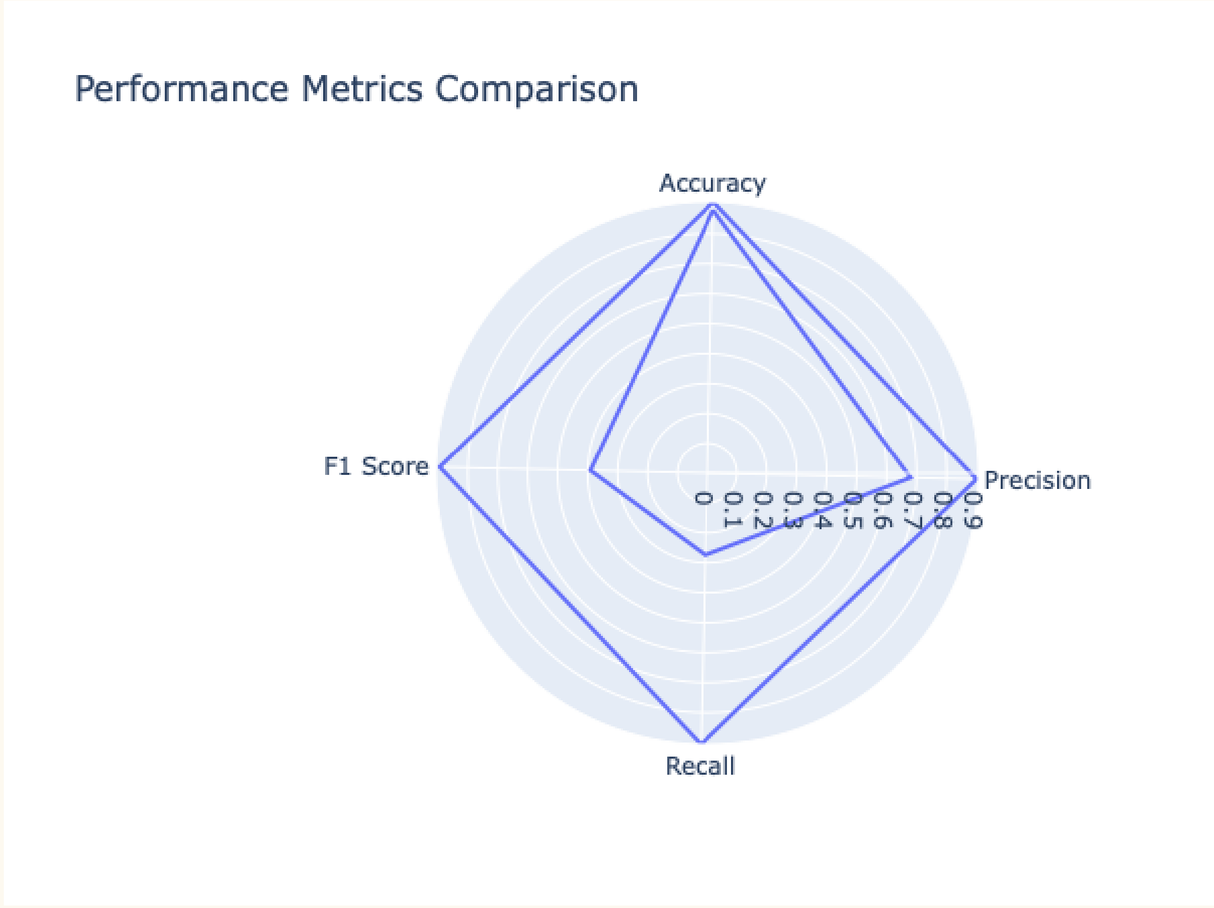
Accuracy: 89.9%
Precision: 089%
Recall: 0.8995
F1 Score: 0.885

Incorrect Predictions: 44

The Results

After optimizing tuned Random Forest Classifier it demonstrated remarkable improvements in key performance metrics:

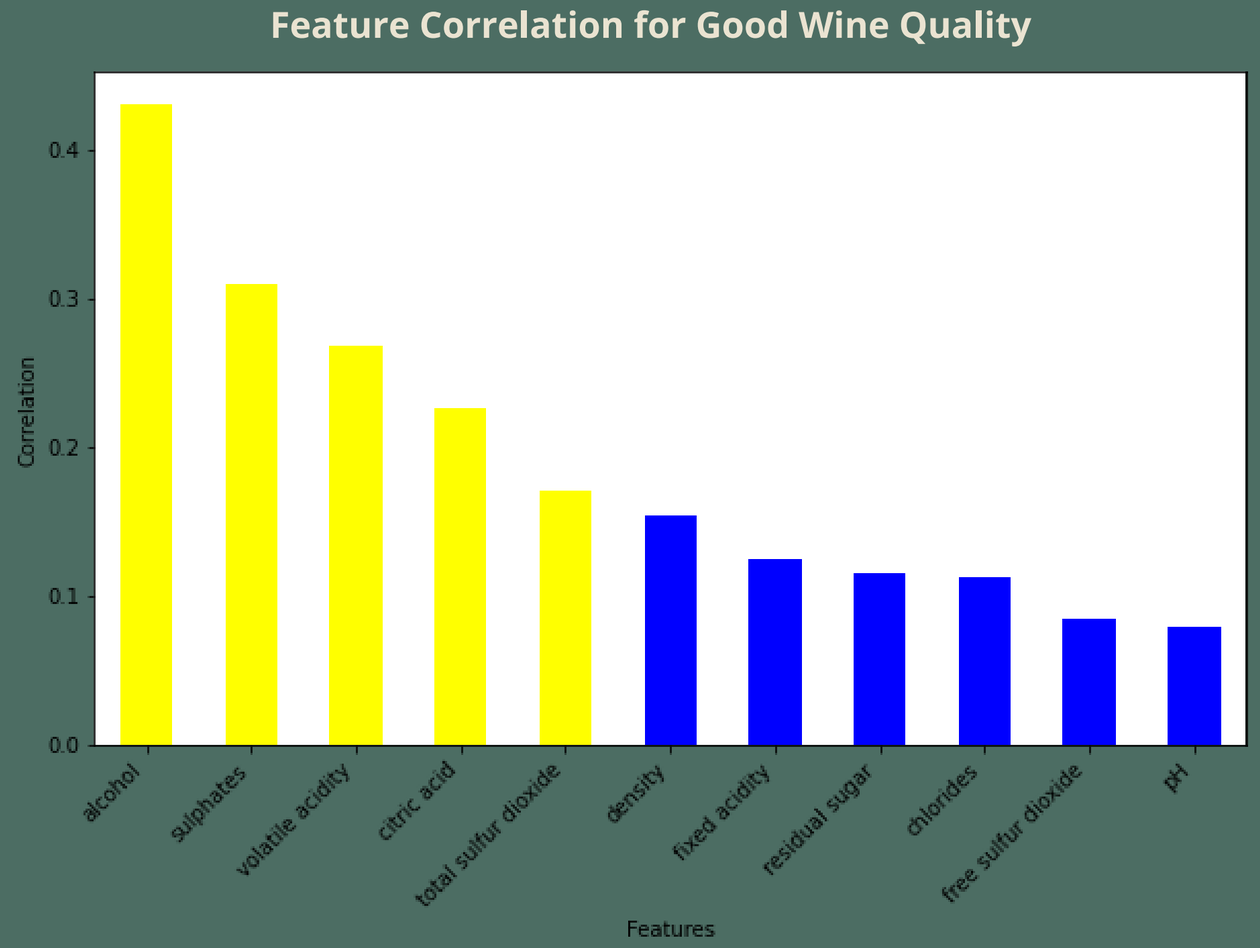
- Accuracy: Improved to 91% from the baseline model's 88%.
- Precision: Enhanced to 90% from the baseline model's 68%.
- Recall: Improved to 0.91 from the baseline model's 0.27.
- F1 Score: Increased to 0.89 from the baseline model's 0.39.



Recommendations

Based on the model's insights, we recommend the following steps for enhancing wine quality:

- Maintain alcohol levels around **11.5%** to maximize quality.
- Manage sulphate levels above .**62** for wine preservation.
- Utilize citric acid strategically for freshness and tartness.
- Monitor total sulfur dioxide to maintain chemical stability **below 40**.
- Keep volatile acidity below **0.5** to avoid off-flavors.



Next Steps:

Data Monitoring: Regularly update the model with new data to maintain its accuracy.

Data Collection: Continue collecting and maintaining data on chemical properties and quality ratings.

Further Analysis: Explore additional factors affecting wine quality, such as environmental conditions and winemaking techniques.

Thank you!

email: nancyvl593@gmail.com
GitHub: nv593