

Wine Quality Prediction Using Machine Learning

MEMBERS:-

- | | |
|----------------------|--------------|
| • VAIBHAV T. GORME | 202301060034 |
| • SAYLI MUNDE | 202301060045 |
| • RAJU PISAL | 202301060052 |
| • ATHARVA RAJANDEKAR | 202301060027 |

GUIDED BY:-

Dr. Abhilasha Joshi

INTRODUCTION

In this project, we built a machine learning model that predicts wine quality based on different chemical parameters. We performed EDA, trained several algorithms, compared results, and deployed the best model.



PROBLEM STATEMENT AND OBJECTIVES

To develop a machine learning model that predicts wine quality from its chemical features.

- **Perform Data Preprocessing**
Clean, analyze, and prepare the Wine Quality dataset for model building.
- **Build & Train ML Models**
Train multiple machine learning algorithms and compare their performance
- **Evaluate Model Performance**
Use accuracy, precision, recall, F1 score, and confusion matrix to select the best model.
- **Deploy the Best Model**
Implement a simple deployment (Streamlit) to make the prediction model usable.



Project Overview



Implement end-to-end ML pipeline



**Dataset → Preprocessing → Training
→ Evaluation → Deployment**



Real-world ML workflow experience





Dataset & Preprocessing

- **Wine Quality Dataset (Red/White)**
- **Handled missing values**
- **Scaling numeric features**
- **Train-test split performed**
- **EDA: correlation heatmap, pairplots, stats**

Model Training

Tried multiple models:

- **Logistic Regression**
- **SVM (LINEAR, RBF , Poly)**
- **XGBoost Classifier**

Hyperparameter tuning for best accuracy



Model Results

Logistic Regression :-

- **Training Accuracy: 69%**
- **Testing Accuracy: 70%**

SVM :-

- LINEAR :**
- **Training Accuracy : 69%**
 - **Testing Accuracy : 71%**

RBF:

- **Training Accuracy : 72%**
- **Testing Accuracy : 73%**

Poly:

- **Training Accuracy: 72%**
- **Testing Accuracy: 74%**

XGBoost :-

- **Training Accuracy: 97%**
- **Testing Accuracy: 82- 83%**





Model Evaluation

Metrics used:

Accuracy, Precision, Recall,
F1-Score
Confusion Matrix to compare
models

Metric	Class 0 Value	Class 1 Value	Overall Average/Accuracy
Precision	0.77	0.88	Macro Avg: 0.82, Weighted Avg: 0.84
Recall	0.78	0.87	Macro Avg: 0.83, Weighted Avg: 0.84
F1-Score	0.78	0.87	Macro Avg: 0.82, Weighted Avg: 0.84
Support	468	832	Total Samples: 1300
Accuracy	N/A	N/A	0.84

Results

- **XGBoost achieved highest accuracy with 84%**
- **Model predicts wine quality effectively**
- **Confusion matrix showed strong classification ability**



Model Deployment



**User inputs features → model predicts
quality**

**Shows real-time prediction
output**

Streamlit web app concept

Conclusion

Our project successfully implemented an machine learning model for predicting wine quality based on chemical properties. Through data preprocessing, EDA, and testing multiple algorithms, XGBoost emerged as the most accurate and reliable model. The results show that machine learning can effectively support fast and consistent wine quality evaluation.

REFERENCES

- [1] P. Cortez, A. Cerdeira, F. Almeida, et al., “Modeling wine preferences,” Decision Support Systems, 2009.**
- [2] UCI Machine Learning Repository, Wine Quality Dataset.**
- [3] Géron, A., Hands-On Machine Learning, O’Reilly, 2020.**
- [4] Streamlit Documentation, 2024.**

