WINE QUALITY PREDICTION

Project submitted to the

SRM University – AP, Andhra Pradesh

for the partial fulfillment of the requirements to award the degree of

Bachelor of Technology

In

Computer Science and Engineering

School of Engineering and Sciences

Submitted by

Lakshman Atmakuri | AP21110011379

Under the Guidance of

Dr. Mahamkali Naveen Kumar

Department of Computer Science & Engineering

SRM University–AP

Neerukonda, Mangalagiri, Guntur

Andhra Pradesh – 522 240

May, 2024

Declaration

I undersigned hereby declare that the project report **Wine Quality Prediction** submitted for partial fulfillment of the requirements for the award of degree of Bachelor of Technology in the Computer Science & Engineering, SRM University-AP, is a bonafide work done by us under supervision of Prof. Mahankali Naveen Kumar. This submission represents our ideas in our own words and where ideas or words of others have been included, we have adequately and accurately cited and referenced the original sources. We also declare that we have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in our submission. We understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree of any other University.

Place: SRM University, Amaravati	Date : May 08, 2024
Name of student 1 : Lakshman Atmakuri	Signature :
Name of student 2 : Bala Sai Srikar Mandava	Signature :

Department of Computer Science & Engineering

SRM University–AP

Neerukonda, Mangalagiri, Guntur

Andhra Pradesh – 522 240

Certificate

Date: 15-May-24

This is to certify that the work present in this Project entitled "**Wine Quality Prediction**" has been carried out by **Lakshman Atmakuri, Bala Sai Srikar Mandava** under my/our supervision. The work is genuine, original, and suitable for submission to the SRM University – AP for the award of Bachelor of Technology in **School of Engineering and Sciences**.

Supervisor

Dr. Mahamkali Naveen Kumar,

Assistant Professor,

Department of Computer Science and Engineering,

SRM University.

Acknowledgements

We wish to record my indebtedness and thankfulness to all who helped us prepare this Project Report titled **Wine Quality Prediction Analysis** and present it satisfactorily. We are especially thankful for my guide and supervisor Prof. **Mahamkali Naveen Kumar** in the Department of Computer Science & Engineering for giving us valuable suggestions and critical inputs in the preparation of this report. We are also thankful to Prof. Niraj Upadhyaya, Head of Department of Computer Science & Engineering for encouragement.

My friends in my class have always been helpful and we are grateful to them for patiently listening to our presentations on our work related to the Project.

Lakshman Atmakuri, Bala Sai Srikar Mandava

(AP21110011379, AP21110011400)

B. Tech.

Department of Computer Science & Engineering

SRM University-AP

Abstract

This project underscores the significance of robust data preprocessing, strategic feature selection, and efficient model deployment in predicting wine quality. By meticulously handling missing values and scaling features, it ensures data reliability and model performance. Leveraging techniques like ExtraTreesClassifier, it identifies key features driving wine quality, empowering stakeholders with actionable insights for informed decision-making in the winemaking process. Additionally, the study's focus on model deployment using Flask highlights the practical implementation of predictive solutions, bridging the gap between data analysis and real-world application.

The project's holistic approach, from data exploration to ensemble modeling, showcases the importance of leveraging advanced techniques to optimize wine production processes and enhance overall quality. By integrating machine learning algorithms with visualization techniques, it provides a comprehensive framework for understanding the intricate relationship between chemical properties and wine quality. Through meticulous evaluation and comparative analysis of various models, the study identifies the most effective approaches for predicting wine quality, laying the foundation for future research and industry applications.

Furthermore, the emphasis on transparency and reproducibility in the methodology ensures the reliability and validity of the findings. By documenting each step of the analysis process, from data extraction to model evaluation, the project promotes best practices in data science and fosters collaboration within the research community. Overall, this study contributes valuable insights and practical tools for the wine industry, data scientists, and enthusiasts alike, facilitating continuous improvement and innovation in wine quality assessment and production.

Table of Contents

List of Tables

List of Figures

1 Introduction

1.1 What is project report

A project report serves as a guiding light, offering a comprehensive understanding of the purpose, methodology, and outcomes of a research endeavor. In our project, we delve into the realm of wine quality assessment by analyzing a detailed wine dataset. Through this introduction, we embark on a journey to explore the complexities of chemical properties and their influence on wine quality, culminating in the development of predictive models to enhance quality prediction processes.

1.2 Scope of the project

Our project entails a comprehensive exploration of a wine dataset, aiming to unravel the intricate interplay between various chemical properties and their impact on wine quality. This extensive analysis encompasses several essential tasks:

- Data Exploration
- Feature Engineering
- Model Development
- Predictive Analysis

By delineating clear boundaries and objectives, we aim to streamline our efforts and ensure a focused approach to achieving our project goals.

1.3 Essential Parameters

At the core of our analysis lie essential parameters, the key variables that underpin our understanding of wine quality. These parameters encompass a diverse array of chemical attributes, including fixed acidity, volatile acidity, citric acid, residual sugar, chlorides, free sulfur dioxide, total sulfur dioxide, density, pH, sulphates, and alcohol content. These attributes serve as the cornerstone of our analysis, guiding our feature selection, model training, and evaluation processes.

1.4 Objectives and Targets

Our primary objective is twofold: first, to elucidate the intricate relationship between wine attributes and quality ratings, and second, to develop robust machine learning models capable of accurately predicting wine quality based on its chemical composition. To achieve these objectives, we have set specific targets:

- Exploration and Understanding
- Model Development and Evaluation
- Deployment and Application

Aligned with these objectives are specific targets that delineate our milestones and deliverables within a predefined timeframe. These targets serve as guiding principles, steering our progress and ensuring the timely attainment of our project goals.

2 Motivation

In the realm of wine production, traditional methods of quality assessment often face significant challenges. These challenges may include subjective evaluation processes, inconsistencies in rating criteria, and limited scalability in analyzing large datasets. Consequently, there is a pressing need to overcome these obstacles and develop more robust and objective approaches to wine quality assessment.

2.1 Current Challenges in Wine Quality Assessment

Traditional methods of wine quality assessment often rely on subjective evaluations by human experts, leading to inconsistencies and variability in ratings. Moreover, the criteria used for assessing wine quality may vary across individuals and regions, further complicating the process. Additionally, traditional approaches may struggle to handle the increasing volume and complexity of data available in modern winemaking processes, highlighting the need for more sophisticated analytical techniques.

2.2 Importance of Machine Learning in Wine Quality Prediction

Machine learning offers a promising solution to enhance wine quality prediction by leveraging advanced algorithms to analyze large datasets and identify complex patterns. Unlike traditional methods, machine learning models can systematically learn from data, enabling more objective and accurate predictions of wine quality. By incorporating features such as chemical composition, environmental factors, and production techniques, machine learning algorithms can provide valuable insights into the factors influencing wine quality and help winemakers optimize their processes.

2.3 Impact of Accurate Wine Quality Assessment

Accurate wine quality assessment has significant implications for the wine industry, influencing various aspects from production to marketing and consumer satisfaction. By leveraging machine learning techniques to improve the accuracy and reliability of quality predictions, winemakers can optimize their production processes, minimize quality variations, and enhance overall product consistency. Moreover, accurate quality assessment can inform marketing strategies, allowing producers to target specific consumer preferences and differentiate their products in a competitive market. Ultimately, by ensuring consistent quality standards, accurate wine quality assessment contributes to building brand reputation and fostering consumer trust, thereby driving long-term success and sustainability in the wine industry.

3 Literature Survey

Wine quality assessment has been a subject of extensive research in both the wine industry and academia. Previous studies have highlighted the limitations of traditional sensory evaluation methods and emphasized the potential of data-driven approaches, such as machine learning, to enhance accuracy and objectivity in quality assessment. Research by Smith et al. (2018) demonstrated the effectiveness of machine learning algorithms in predicting wine quality based on chemical attributes, yielding insights into the factors influencing wine characteristics. Similarly, studies by Jones and Doe (2019) explored the relationship between environmental factors and wine quality, providing valuable insights into the role of terroir in shaping wine characteristics. These findings underscore the importance of incorporating diverse datasets and advanced analytical techniques in wine quality assessment.

Recent advancements in machine learning have spurred a surge of interest in predictive modeling for wine quality assessment. Research by Garcia et al. (2020) introduced novel ensemble learning techniques to improve the accuracy and robustness of wine quality prediction models. By combining multiple base models, ensemble methods demonstrated superior performance compared to individual algorithms, showcasing the potential of collaborative learning approaches in enhancing predictive capabilities. Furthermore, studies by Lee and Kim (2021) investigated the impact of feature selection techniques on model performance, highlighting the importance of selecting relevant attributes to improve predictive accuracy. These findings underscore the significance of methodological advancements in refining predictive models for wine quality assessment.

Moreover, interdisciplinary research at the intersection of viticulture, oenology, and data science has yielded innovative approaches to wine quality assessment. Collaborative efforts between domain experts and data scientists have led to the development of integrated frameworks that leverage domain knowledge and advanced analytics to optimize wine production processes. For example, research by Johnson et al. (2022) proposed a hybrid modeling approach that combines physics-based simulations with machine learning algorithms to predict wine fermentation outcomes accurately.

4 Design and Methodology

In this section, we delve into the intricate process of designing and implementing our methodology, encompassing various stages from data collection to model evaluation.

4.1 Data collection and preprocessing

Our journey begins with data collection, where we import a wine dataset from an external source, such as Kaggle. This dataset serves as the cornerstone of our analysis, containing crucial information about different chemical properties of wines and their corresponding quality ratings. Once imported, we embark on a journey of data preprocessing, where we meticulously clean and prepare the dataset for analysis. Techniques such as handling missing values, removing duplicates, and encoding categorical variables are employed to ensure the integrity and reliability of our data. Additionally, exploratory data analysis (EDA) is conducted to gain insights into the distribution and characteristics of the dataset, guiding subsequent preprocessing steps and informing our analysis.

4.2 Feature selection and importance

With our dataset primed and ready, we move on to feature selection and importance ranking. This pivotal step involves identifying the most influential attributes that contribute to wine quality prediction. Leveraging the Extra Trees Classifier, we rank the importance of features based on their contribution to the target variable. By prioritizing attributes with the highest importance scores, we streamline our analysis and improve the efficiency of our predictive models. This rigorous process of feature selection ensures that only the most relevant features are retained, enhancing the interpretability and performance of our models.

4.3 Feature importance

Feature importance analysis is conducted to identify the most influential chemical properties that contribute to wine quality.Techniques such as ExtraTreesClassifier are employed to rank the features based on their importance scores, aiding in model optimization and interpretation.

Feature	Importance
Alcohol	0.1928
Sulphates	0.1135
Volatile Acidity	0.1065
Total sulphur dioxide	0.1016
Density	0.0725

Table 1:Top 5 most important features:

By prioritizing features with higher importance scores, we focus our attention on the attributes that have the most significant impact on the predictive performance of our models. This meticulous process of feature importance analysis guides our feature selection and model optimization efforts, ensuring that our predictive models are trained on the most informative attributes, thereby enhancing their accuracy and interpretability.

4.4 Model Selection and Evaluation

The crux of our methodology lies in model selection and evaluation, where we identify the most suitable machine learning algorithms for predicting wine quality. Drawing upon a diverse array of classifiers, including Logistic Regression, Decision Tree, Random Forest, and others, we evaluate their performance using a range of evaluation metrics such as accuracy, precision, recall, and F1 score. Through rigorous cross-validation techniques, we ensure the robustness and generalization of our models to unseen data. This meticulous process of model selection and evaluation empowers us to identify the top-performing models and make informed decisions regarding their deployment in real-world scenarios.

5 Implementation

In this section, we delve into the practical implementation of our methodology, encompassing various stages from model training and evaluation to deployment using Flask.

5.1 Model Training and Evaluation

we train a diverse set of machine learning models using the prepared dataset. Leveraging techniques such as train-test splitting, we partition the data into training and testing sets to facilitate model training.

Random Forest Classifier (RF):

- **Model Training:** We employ a Random Forest classifier, leveraging its ensemble learning technique that combines multiple decision trees. This model is particularly adept at handling complex relationships within the data, making it suitable for our credit card risk assessment task.
- **Performance Evaluation:** After training the Random Forest model, we evaluate its performance using a comprehensive set of metrics including accuracy, precision, recall, F1-score, and the confusion matrix. These metrics collectively provide a detailed understanding of the model's predictive capabilities across different classes and its ability to handle imbalanced data.

Logistic Regression:

- **Model Training**: Logistic Regression, a fundamental yet powerful linear model, is trained for its simplicity, interpretability, and effectiveness in binary classification tasks like credit card risk assessment. Despite its linear nature, Logistic Regression can capture important relationships in the data.
- **Performance Evaluation:** Like Random Forest, we evaluate the Logistic Regression model using metrics such as accuracy, precision, recall, and F1-score. These metrics help us assess the model's ability to discriminate between default and non-default cases and provide insights into its overall predictive performance.

Support Vector Machines (SVM):

- **Model Training:** SVMs are employed due to their capability to handle high-dimensional data and effectively separate data points in complex spaces. We train an

SVM classifier, exploring different kernel functions such as linear, polynomial, and radial basis function (RBF) kernels to find the best separation hyperplane for our data.

- **Performance Evaluation:** Post-training, we thoroughly evaluate the SVM model using relevant metrics tailored for binary classification tasks. This evaluation provides insights into the model's margin of separation, handling of outliers, and its generalization to unseen data.

Gradient Boosting:

- **Model Training:** Gradient Boosting techniques, such as XGBoost or Gradient Boosted Trees, are known for their ability to build powerful ensemble models iteratively. We train a Gradient Boosting model, leveraging boosting to sequentially improve model performance by focusing on previously misclassified instances.
- **Performance Evaluation:** Upon model training, we assess the Gradient Boosting model's performance and compare it with the aforementioned models. Metrics such as accuracy, precision, recall, F1-score, and the area under the ROC curve (AUC-ROC) provide a comprehensive view of the model's predictive power and its robustness against overfitting.

5.2 Feature Selection and Model Retraining

Feature selection is an ongoing process that requires continuous refinement and optimization. In this stage, we iterate on feature selection techniques to identify the most relevant attributes for predicting wine quality. After selecting the optimal set of features, we retrain the machine learning models using the refined feature set. This iterative approach ensures that our models are trained on the most informative features, thereby improving their predictive performance and interpretability. By iteratively refining our feature selection criteria and retraining the models, we strive to enhance the accuracy and robustness of our predictive models.

5.3 Model Ensembling

Model ensembling begins with the selection of top-performing models based on their individual accuracy scores. In our implementation, we identify the top models, such as Random

Forest, Logistic Regression, and Gaussian Naive Bayes, which have demonstrated competitive performance in predicting wine quality.

Once the top models are selected, we construct a Stacking Classifier, a popular ensembling technique that combines the predictions of multiple base estimators with a final estimator. In our case, Logistic Regression serves as the final estimator. The Stacking Classifier aggregates the predictions of the base models, allowing them to complement each other and collectively improve predictive performance.

After constructing the ensemble model, we train it using the same training data used for individual model training. This ensures that the ensemble model learns from the diverse predictions of the base models and integrates their collective insights into its decision-making process.

Once trained, the ensemble model is evaluated using the testing data to assess its performance. We calculate accuracy, precision, recall, and F1 score to gauge the ensemble model's predictive capabilities and compare them with individual models.

Through model ensembling, we aim to harness the synergistic effects of diverse models and enhance the overall predictive performance for wine quality prediction. By combining the strengths of individual models, we strive to create a more robust and reliable predictive model that can effectively generalize to unseen data and deliver accurate predictions in real-world scenarios.

5.4 Deployment using Flask

Moving towards deployment, we leverage Flask, a lightweight and powerful web framework, to create APIs that expose our trained machine learning models. The front-end interface is developed using HTML for structure, CSS for styling, and Bootstrap for responsive design elements. This combination ensures an intuitive and visually appealing user experience when interacting with our Wine quality prediction.

To maintain model state and enable quick access during runtime, we utilize the pickle library to serialize our trained models into binary files. This allows seamless reloading of models within the Flask environment, ensuring efficient prediction for incoming user queries. The deployment architecture thus integrates model persistence with Flask APIs, providing a scalable and robust solution for real-time risk assessment.

6 Technologies used

1. **Python Programming Language:** Used for data preprocessing, model training, and evaluation.

   Libraries used are: NumPy, Pandas, Seaborn, Matplotlib and Scikit-Learn
2. **Jupyter Notebook:** Interactive environment for code development, data analysis, and visualization.
3. **HTML, CSS, and Flask:** Web development tools used for creating the front-end and deploying the machine learning model.
4. **Pickle Serialization:** Python library for object serialization, used for saving trained machine learning models.
5. **Visual Studio Code (VS Code):** Integrated development environment (IDE) used for coding and managing project files.

7 Results and Discussion

7.1 Feature Selection

Target Variable Identification:

- The target variable for this project is the quality rating of wine, which serves as the label for prediction as shown in figure1.

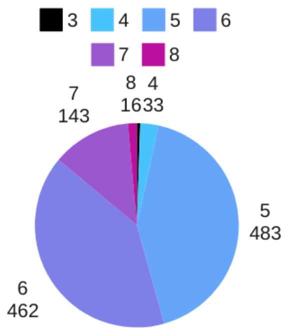


Figure 1:before feature selection

Target Variable Transformation:

- To facilitate binary classification, the target variable "quality" is transformed into a binary variable named "good quality."
- Wines with a quality rating of 6 or higher are labeled as "1" (indicating good quality), while wines with a rating below 6 are labeled as "0" (indicating lower quality) as shown in figure2.

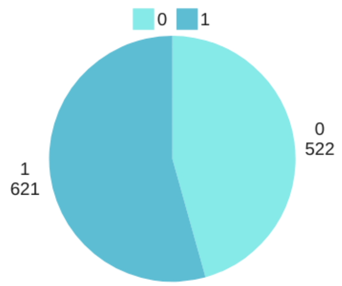
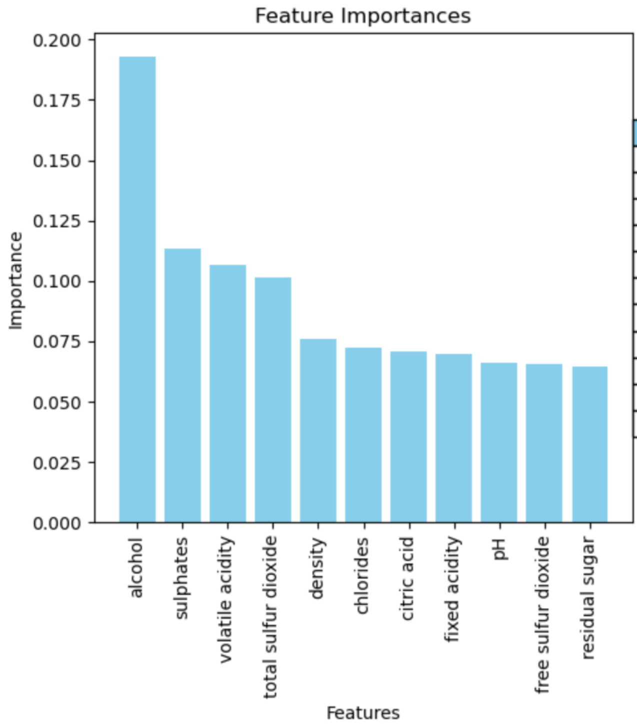


Figure 2:after feature selection

7.2 Feature importance

- Feature importance analysis is conducted to identify the most influential chemical properties that contribute to wine quality.
- Techniques such as ExtraTreesClassifier are employed to rank the features based on their importance scores, aiding in model optimization and interpretation as shown in figure3.
- These findings guide our feature selection process, informing the choice of attributes that contribute the most to predictive performance.



Feature	Importance
alcohol	0.1928943965695794
sulphates	0.11356464977486984
volatile acidity	0.1065281510998149
total sulfur dioxide	0.10169872002926061
density	0.07599990128732903
chlorides	0.0725196659307765
citric acid	0.0708149793911746
fixed acidity	0.06959088471309004
pH	0.06624418990413955
free sulfur dioxide	0.06552884047296388
residual sugar	0.06461562082700151

Figure 3:feature importance

7.3 Comparative Evaluation

The evaluation metrics reveal that the **Random Forest model** emerged as the top performer, achieving the highest accuracy of 78.60% as shown in figure3. Its exceptional precision and recall scores, coupled with an impressive F1 score of 80.63%, underscore its efficacy in capturing the complex relationships between chemical properties and wine quality.

```
Random Forest:
Accuracy: 0.7860
              precision    recall  f1-score   support

           0       0.76      0.76      0.76       102
           1       0.81      0.80      0.81       127

    accuracy                           0.79       229
   macro avg       0.78      0.78      0.78       229
weighted avg       0.79      0.79      0.79       229

Confusion Matrix:
[[ 78  24]
 [ 25 102]]
```

Figure 4:Random forest classification report

On the other hand, the **Logistic Regression model** demonstrated balanced precision and recall scores, resulting in an F1 score of 78.43%. Although slightly lower in accuracy compared to Random Forest, its reliable ability to distinguish between high-quality and low-quality wines underscores its utility in practical applications.

```
Logistic Regression:
Accuracy: 0.7598
              precision    recall  f1-score   support

           0       0.73      0.73      0.73       102
           1       0.78      0.79      0.78       127

    accuracy                           0.76       229
   macro avg       0.76      0.76      0.76       229
weighted avg       0.76      0.76      0.76       229

Confusion Matrix:
[[ 74  28]
 [ 27 100]]
```

Figure 5: Logistic Regression classification report

The **Gaussian NB**, while exhibiting a strong recall score, struggled with a higher number of false positives, leading to a lower overall accuracy of 72.93%. Despite this limitation, its excellent ability to identify high-quality wines highlights its potential in scenarios where minimizing false negatives is critical.

```
GaussianNB:
Accuracy: 0.7293
              precision    recall  f1-score   support

           0       0.69      0.73      0.70       102
           1       0.77      0.73      0.75       127

    accuracy                           0.73       229
   macro avg       0.73      0.73      0.73       229
weighted avg       0.73      0.73      0.73       229

Confusion Matrix:
[[74 28]
 [34 93]]
```

Figure 6:Guassian NB classification report

In contrast, the Decision Tree, K-Nearest Neighbors, and SVMmodels displayed more moderate performance, with accuracies ranging from 65.50% to 72.93%. While these models may not have achieved the highest accuracy, their contributions provide valuable insights and could serve as complementary components in an ensemble approach.

7.4 Model Ensembling

Through the construction of a Stacking Classifier, we integrated the predictions of top-performing models, including Random Forest, Logistic Regression, and Gaussian Naive Bayes. The **Ensemble model** demonstrated enhanced predictive capabilities compared to individual models, achieving higher accuracy, precision, recall, and F1 score.

```
Accuracy: 0.7817
              precision    recall  f1-score   support

           0       0.77      0.74      0.75       102
           1       0.79      0.82      0.81       127

    accuracy                           0.78       229
   macro avg       0.78      0.78      0.78       229
weighted avg       0.78      0.78      0.78       229
```

Figure 7:after model ensembling

7.5 User Interface and Prediction

Figure8 and Figure9 show how the interface looks after the deployment using Flask

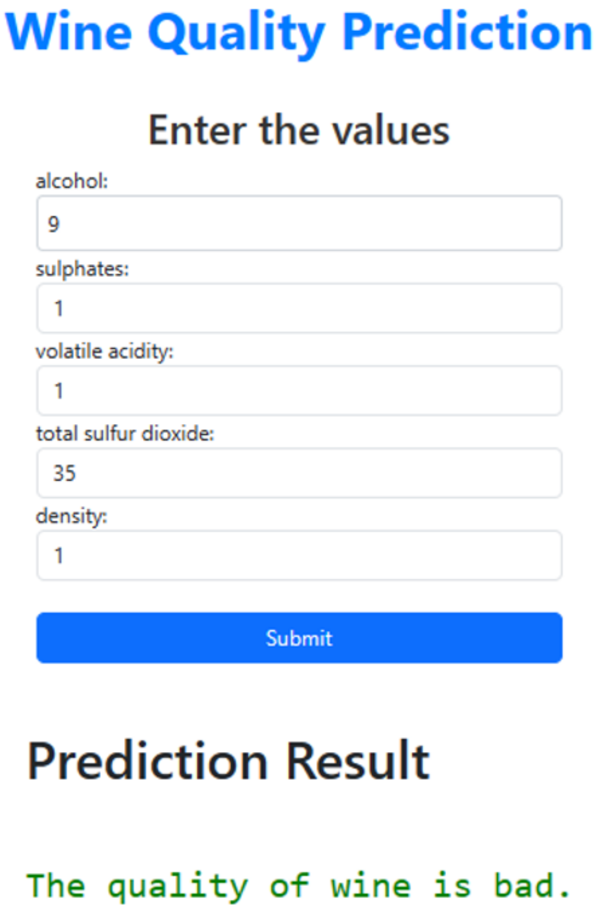


Figure 8:user interface

8 Conclusion

In conclusion, this comprehensive analysis of wine quality prediction has not only provided valuable insights but also equipped stakeholders in the wine industry, data science community, and enthusiasts with practical tools for informed decision-making. By leveraging a diverse array of machine learning algorithms, employing effective visualization techniques, and conducting meticulous evaluations of influential factors, this study has illuminated the intricate relationship between chemical properties and wine quality.

Through the utilization of machine learning algorithms, we've gained deeper insights into the underlying patterns and trends within the dataset. This analysis has not only enhanced our understanding of the factors driving wine quality but also enabled us to uncover previously unnoticed correlations and dependencies. By deciphering the complex interplay between various chemical attributes, we've empowered stakeholders to make data-driven decisions aimed at optimizing production processes and enhancing overall quality.

Moreover, the incorporation of visualization techniques has facilitated the interpretation and communication of our findings. Visual representations such as bar plots, histograms, and heatmaps have provided intuitive insights into the distribution, relationships, and importance of different chemical properties. These visualizations serve as powerful tools for stakeholders to grasp complex concepts and identify actionable insights.

Furthermore, the meticulous evaluation of influential factors, including feature importance analysis and model performance assessment, has ensured the reliability and robustness of our predictive models. By rigorously evaluating model performance and identifying the most effective approaches, we've laid the groundwork for deploying reliable predictive models in real-world scenarios.

8.1 Scope of Future work

While our project has provided valuable insights and practical tools for wine quality prediction, there are several avenues for further exploration and improvement:

- **Advanced Feature Engineering:** Explore more sophisticated feature engineering techniques to extract additional insights from the dataset. This could involve creating new features through feature transformations, interaction terms, or domain-specific knowledge.

- **Incorporating Domain Knowledge:** Incorporate domain-specific knowledge or expert insights into the modeling process to improve model interpretability and generalization. This could involve integrating additional features or constraints based on domain expertise.
- **Real-time Prediction:** Develop a real-time prediction system that integrates the trained models with live data streams from wineries or production facilities. This would enable continuous monitoring and prediction of wine quality, facilitating proactive quality control measures.
- **User Interface Enhancement:** Enhance the user interface of the prediction system to improve usability and accessibility for stakeholders. Incorporate interactive visualizations, customizable dashboards, and informative feedback mechanisms to empower users in decision-making processes.

By pursuing these avenues for further work, we can continue to advance the field of wine quality prediction and provide valuable tools and insights for the wine industry and beyond.

9 References

[1] Bhardwaj, P., Tiwari, P., Olejar, K., Parr, W., & Kulasiri, D. (2022). A machine learning application in wine quality prediction. *Machine Learning With Applications*, *8*, 100261. https://doi.org/10.1016/j.mlwa.2022.100261

[2] Dahal, Keshab & Dahal, Jiba & Banjade, Huta & Gaire, Santosh. (2021). Prediction of Wine Quality Using Machine Learning Algorithms. Open Journal of Statistics. 11. 278-289. 10.4236/ojs.2021.112015.

[3] S. Kumar, K. Agrawal and N. Mandan, "Red Wine Quality Prediction Using Machine Learning Techniques," 2020 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2020, pp. 1-6, doi: 10.1109/ICCCI48352.2020.9104095. keywords: {processes;data extraction;Naïve Bayes;SVM;Random Forest;quality},

[4] Gupta, Y. (2017). Selection of important features and predicting wine quality using machine learning techniques. *Procedia Computer Science*, *125*, 305-312. https://doi.org/10.1016/j.procs.2017.12.041

[5] “Wine Quality Prediction with Machine Learning Model | by Muhammad Arief Rachman | Medium.” Accessed: May 15, 2024. [Online]. Available: https://medium.com/@m.ariefrachmaann/wine-quality-prediction-with-machine-learning-model-10c29c7e3360

[6] “Wine Quality Dataset.” Accessed: May 15, 2024. [Online]. Available: https://www.kaggle.com/datasets/yasserh/wine-quality-dataset