Prediction of Wine Quality Using Machine Learning

Puneetha M R, Ananya P, Ashritha M, Bhavana H S, Chethana B M

[1]Assistant Professor,[2]Student,[3]Student, [4]Student, [5]Student

[1]Department of Computer Science and Engineering,

[1]K S School of Engineering and Management, Bengaluru, Karnataka, India

Abstract - This study employs ML algorithms, including Ridge regression, Decision trees, Random Forest, and Support Vector, to predict wine quality based on a dataset encompassing various chemical properties. The automated system replaces traditional human-driven assessments, reducing time and effort. With a user-friendly interface, individuals can input specific parameters to swiftly generate a predicted quality score for a designated wine, offering quick and informed insights into its potential quality.

Index Terms - Ridge regression(RR), Machine Learning(ML), Decision trees(DT), Random Forest Algorithm(RFA), Support Vector Machine(SVM)

I. INTRODUCTION

Wine, a fermented beverage derived from grapes, undergoes the fermentation process without the addition of sugars, acids, enzymes, water, or other nutrients. Wine is not just a drink it is an essential part of laws and traditions in many religions like Jewish and Christianity. The pivotal element in this context is the quality of wine, which has a significant importance for both consumers and the manufacturing industry. Certification of product quality proves beneficial for businesses in boosting revenue. In the past, assessing product quality was a time-consuming process conducted at the final step of the manufacturing process, demanding substantial resources. This included the need for numerous human experts to evaluate product quality, making the procedure quite costly. Determining wine quality based on human specialists' opinions is a challenging task due to varying perspectives. Whereas the prediction of wine quality through machine learning, based on user input collected from different instruments, represents a groundbreaking approach in the realm of winemaking. In this innovative methodology, data obtained from diverse instruments serves as input to powerful machine learning algorithms, allowing for the accurate anticipation of wine quality. This approach transcends traditional subjective assessments by harnessing objective measurements and analytical insights derived from sophisticated instruments.

II. LITERATURE SURVEY

The study by Nithin Khilari, Pravin Hadawale, Hasan Shaikh & Sachine Kolase employ machine learning to predict wine quality through a systematic process. The initial dataset undergoes pre-processing. Then the required algorithms are trained on the former, while the latter is used to evaluate and compare model accuracies. Conclusions are drawn to select the optimal algorithm for wine quality prediction based on the testing set's results. The correctness of the chosen model is then validated, and the study assesses and compares the accuracy of different algorithms to determine the optimum one for predicting wine quality. However, the dataset is considerably limited. The authors could contact the wine experts and add more data to the dataset. The provided information lacks details regarding direct interactions with wine experts and the dataset's origin. It is essential to understand how the dataset was obtained and whether insights from wine experts played a role in shaping the study. Meeting with wine experts could offer valuable domain knowledge, influencing feature selection and enhancing the dataset's relevance. Additionally, knowing how experts assess wine quality and incorporating their insights into model development could strengthen the predictive capabilities. Clarity on data collection methods and the involvement of domain experts is crucial for ensuring the dataset's representativeness and the robustness of the machine learning approach.

III. METHODOLOGY

Wine quality prediction through machine learning involves a systematic approach in five key steps. The initial step is data collection, where information from measuring instruments is gathered to serve as input for the predictive model. This data should encompass a variety of features such as alcohol, pH, sugar content, and more, which play crucial roles in determining wine quality. the second step is to gather a diverse and comprehensive dataset with various wine characteristics to train the machine learning model. This dataset should represent a wide range of wine samples. Subsequently, the third step involves a comprehensive data analysis, involving preprocessing techniques to handle missing values, and normalization. Additionally, the dataset is typically split into training and testing sets to evaluate the model's performance effectively. The fourth step involves of selecting an appropriate algorithm for model building. Based on the complexity of the dataset and the desired level of accuracy, various algorithms of machine learning such as Decision trees, Support vector machines, Ridge regression are considered. The algorithm should align with the nature of the wine quality prediction problem and the characteristics of the dataset. Once the model is trained and validated, the fifth step is to provide the user-collected input to the system. Users can provide the input of specific wine characteristics into the model, and the algorithm will generate predictions regarding the quality of the wine based on the patterns from the training dataset. This interactive element allows users to obtain real-time predictions for specific wines they are interested in. Finally, deploy the trained model into a production environment. This deployment ensures that the predictive capabilities of the model can be accessed by end-users seamlessly.

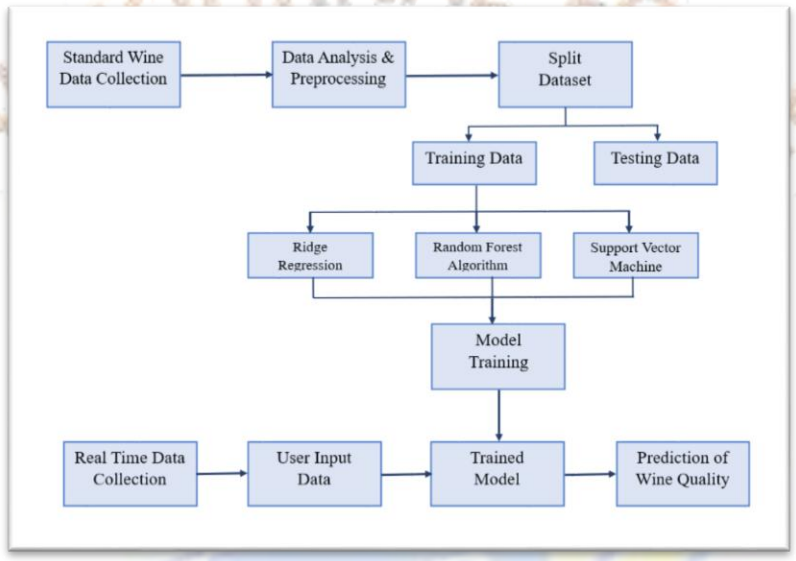

Fig 1. Implementation of Flow chart

IV. CONCLUSIONS

In conclusion, the forecast of wine quality using machine learning has proven to be a valuable and promising approach in the field of viticulture and oenology. Through the utilization of effective algorithms such as Ridge regression, Decision trees, Random Forests algorithm, Support Vector algorithm, and the analysis of diverse features, such as chemical composition and sensory attributes, machine learning models can effectively assess and predict the quality of wines. This technology provides many advantages, including increased efficiency in quality control, enhanced precision in grading, and the potential for optimizing production processes. The ability to anticipate wine quality allows for proactive decision-making, enabling winemakers to make timely adjustments in the production process, ultimately leading to improved overall quality and consumer satisfaction. Moreover, the integration of machine learning in the wine industry can contribute to personalized recommendations for consumers based on their preferences, fostering a more tailored and enjoyable wine selection experience.

V. APPLICATIONS

Anticipating wine quality using machine learning offers a broad range of applications across the wine industry and related sectors. One key area is quality assurance in winemaking, where machine learning prototype can detect patterns indicative of potential issues at various stages of production. This early detection allows winemakers to intervene promptly, ensuring the overall quality of the wine. Machine learning enhances authentication and fraud detection in the wine industry by verifying the authenticity of wines and reducing the occurrence of fraudulent products. Predictive models can also assist in optimizing the supply chain by predicting demand and ensuring that the right quantity of high-quality wine is distributed to different markets. Furthermore, these models aid in inventory management, supporting production planning to meet demand and maintain a balance between quality and quantity. Market segmentation and targeting benefit from predictive models, enabling wineries to understand and leverage factors influencing quality for strategic marketing. E-commerce platforms and wine retailers can use machine learning prototypes to analyse customer preferences and recommend wines based on predicted quality. This enhances the customer experience and increases the likelihood of customer satisfaction. It also can analyse historical data on weather conditions, soil composition, and grape characteristics to predict the optimal time for harvesting. This ensures that grapes are picked at the peak of ripeness, contributing to better wine quality.

VI. REFERENCES

[1] Dahal K.R., Dahal J.N., Banjade H., Gaire S. Prediction of wine quality using machine learning algorithms Open Journal of Statistics, 11 (2021), pp. 278-289, 10.4236/ojs.2021.112015

[2] Larkin, T. and McManus, D. (2020) An Analytical Toast to Wine: Using Stacked Generalization to Predict Wine Preference. Statistical Analysis and Data Mining: The ASA Data Science Journal, 13, 451-464.

[3] Yunhui Zeng1, Yingxia Liu1 , Lubin Wu1 , Hanjiang Dong1. "Evaluation and Analysis Model of Wine Quality Based on Mathematical Model ISSN 2330-2038 E-ISSN 2330-2046, Jinan University, Zhuhai, China

[4] Saini, A., & Sharma, A. (2019), Predicting the Unpredictable: An Application of Machine Learning Algorithms in Indian Stock Market, Annals of Data Science, doi:10.1007/s40745-019-00230-7

[5] Shin, T. (2020, May 8), Predicting Wine Quality with Several Classification Techniques Medium, https://towardsdatascience.com/predicting-wine-qualitywith-several-classification-techniques-179038e.

[6] Mohit Gupta, Vanmathi C, "A Study and Analysis of Machine Learning Techniques in Predicting Wine Quality", International Journal of Recent Technology and Engineering, Volume-10 Issue-1, May 2021.

[7] Shuhao Zhang, Caixing Shao, Wei Xiao, "Research on Red Wine Quality Based on Data Visualization", 2020 3rd International Conference on Artificial Intelligence and Big Data.

[8] Devika Pawar, Aakanksha Mahajan, Sachin Bhoithe, "Wine Quality Prediction using ML Algorithms", International Journal of Computer Applications Technology and Research Volume 8–Issue 09, 2019.

[9] Swati R. Khokale, Bhushan, Diksha , Divya Bhamare, Harshada Pawar, "Review on Wine Quality Testing using MachineLearning", JETIR May 2022.