IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 7, July 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

9940 572 462 | 6381 907 438 | ijircce@gmail.com | www.ijircce.com

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207087|

Wine Quality Prediction using Machine Learning

Vaishnavi, Bhavya Shree H M

PG Student, Department of MCA, The National Institute of Engineering,Mysuru, Visvesvaraya Technological University, Belagavi, India

Assistant Professor, Department of MCA, The National Institute of Engineering,Mysuru, Visvesvaraya Technological University, Belagavi, India

ABSTRACT: The need to apply and implement sophisticated predictive analytics for consumer goods has meant that the wine sector has also embraced this concept which is very crucial when it comes to quality determination of wines. The following project is titled "Wine Quality and Age Prediction Using Machine Learning" in which this research aims at using advanced machine learning techniques to give an estimation both the quality and age of the wine based on its physicochemical characteristics. The given project is based on diverse kinds of specifics regarding wine, including alcohol percentage, pH level, residual sugar content in a bottle and more. The project intends to provide quality and accurate results about the wines age and quality by using tools like Principal Component Analysis and a combination of the following machine learning approaches, the Random Forest, Support Vector Machine and Naive Bayes.

KEY WORDS: Wine Quality Prediction, Wine Age Estimation, Machine Learning, Predictive Analytics, Principal Component Analysis, Random Forest, Support Vector Machine, Naive Bayes, Physicochemical Characteristics, Consumer Goods, Wine Sector

I. INTRODUCTION

A forecast of wine is an ideal that many wine producers are aiming at achieving. Thus, it is possible to satisfy the consumer demand, maintain uniformity, and improve the process with the assistance of reliable forecasts concerning quality. Earlier, the evaluation of wine also relied on vague observations of professional taste buds that constituted a strong framework for the work. Nevertheless, they have limitations: much effort is needed, expenditure is high, and, finally, discrimination by people. Quite recently due to the development of the machine learning methods, new possibilities of objective estimation of wine qualities have emerged. These computational methods can work in an efficient manner, sliding over the quality of ratings and the quantifiable aspects of its wine parameters like acidity and alcohol concentration.

Perhaps most popular is the 'quality of wine' dataset from Portuguese Vinho Verde region, which includes data on wine samples that includes the reds and whites along with the physicochemical tests and the quality ratings. Based on this data, prediction models have been developed employing the machine learning algorithms such as the neural networks, decision tree and linear regression. Mean absolute error and root mean squared error are employed to analyze how accurate the models' approximation is in terms of the outcome of quality scores. Thus, accurate models for wine quality can be useful in managing and decision making starting from the selection of grapes as well as ageing. However, as always, there are challenges, which are as follows: While adopting the machine learning great potential, there are potential issues that have been raised to consider; the integration of human expert knowledge with the implementation of data-driven models, search for other advanced approaches and use of other data sources.

II. OBJECTIVE

The hypothesis of this study of using the technique of machine learning of wine quality which is postulated as follows: By taking into account only the chemical features of the wine and using data which shown to be useful for wine producer and buyers, it is possible to construct the wine quality prediction model which will not be prejudiced. This involves formulating the right approach to machine learning that would provide the right quality of wine from the chemical input. Primarily, the task of the model is to provide the target audience with the assessment of the given chemical indicators and demonstrate how it influences wine quality, thus, the degree of dependence on the subjectivist view of the wine-making process is reduced as a secondary result. To increase the usability of the application, a simplification means in the form of a graphical user interface or API to facilitate the use of the utility by the wine industry, including creators of wine, stores, and regular users, will be developed. Besides, it will categorize contributive chemical compounds that have a significant impact on the wine quality will be identified and case out, thus eliminating

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207087|

the need to rely on such traits that are bound to be subjective. Preserving the identified level of specificity in terms of its relevance to day-to-day operations and transforming it into an actually useful tool will be crucial.

III. LITERATURE SURVEY

The Report Examine the body of the Knowledge Wine Quality Prediction Using Machine Learning Relevant Studies on the Following and Analyzed.

1. “Red wine quality prediction using machine learning"
2. “Selection of important features and predicting wine quality using machine learning techniques”
3. “Prediction of Wine Quality Using Machine Learning Algorithm”
4. “Wine Quality Prediction using ML Techniques and KNIME”
5. “Review on Wine Quality Testing using Machine Learning”
6. “Prediction Of Wine Quality Using Machine Learning”
7. “Quality Prediction of Red Wine based on Different Feature Sets Using Machine Learning Techniques”

IV. METHODOLOGY

The framework for using machine learning to predict wine age and quality involves a detailed and systematic approach. Initially, the process focuses on feature-based predictions related to wine quality and age. The first step, gathering and pre-processing of inputs, begins with collecting a comprehensive dataset that includes various characteristics such as acidity levels, sugar content, sulphates, and alcohol content. This diverse dataset is essential for exposing the model to a wide range of scenarios, ultimately leading to an optimal solution. The dataset is then organized by splitting it into three distinct subsets: training, validation, and test sets, each ensuring equal representation of different wine types and conditions. This division is crucial for a thorough evaluation of the model’s performance. Additionally, the integrity of the data and its annotations is verified to ensure that the model training is based on high-quality and accurate information.
In the data preprocessing stage, Principal Component Analysis (PCA) and scaling techniques are applied. PCA is used for dimensionality reduction, focusing on the most significant features while discarding less relevant information. This step helps in emphasizing the key characteristics that impact wine quality and age. Scaling techniques are employed to normalize or standardize feature values, ensuring a consistent format across the dataset. The data is then transformed using PCA and scalers, preparing it for effective use in machine learning models. This thorough preprocessing ensures that the input data is optimized for model training, leading to more accurate and reliable predictions.

The comprehensive approach to designing a machine learning system for predicting wine age and quality includes several critical steps. Model training begins by solving multiple models using the preprocessed dataset, incorporating PCA- transformed features to enhance data quality. Hyperparameter optimization is performed to fine-tune elements such as learning rate and batch size, ensuring the model’s effectiveness. Transfer learning is utilized by leveraging pre-trained models to accelerate training and achieve better performance. During this phase, regular monitoring is crucial to adjust the model and mitigate issues like overfitting or underfitting, thereby improving the model's ability to learn from the data.

Following training, model evaluation and validation is conducted. This involves cross-validating the trained models with the validation dataset to assess their performance and reliability. Key metrics, such as quality prediction accuracy and age prediction precision, are calculated to evaluate the models. The best-performing model is selected based on these metrics and further validated using a distinct test dataset to ensure its effectiveness in real-world scenarios. Robustness verification is also performed to test the model’s stability across various settings, ensuring its flexibility and reliability under different conditions.

In the integration and deployment phase, a Flask-based web application is designed to serve as the program's frontend. This application allows users to input wine characteristics and receive predictions on wine quality and age. The system includes user authentication for secure login and session management, a prediction interface for entering data, and a result display to present predictions along with explanations or rationales.

Finally, monitoring and maintenance are essential for ongoing system performance. Regular monitoring ensures the system's utilization and addresses any emerging issues. Processes for updates, patches, and improvements are established based on user feedback and operational insights, ensuring the system remains current and continues to meet users’ needs effectively.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207087|

V. SYSTEM ARCHITECTURE

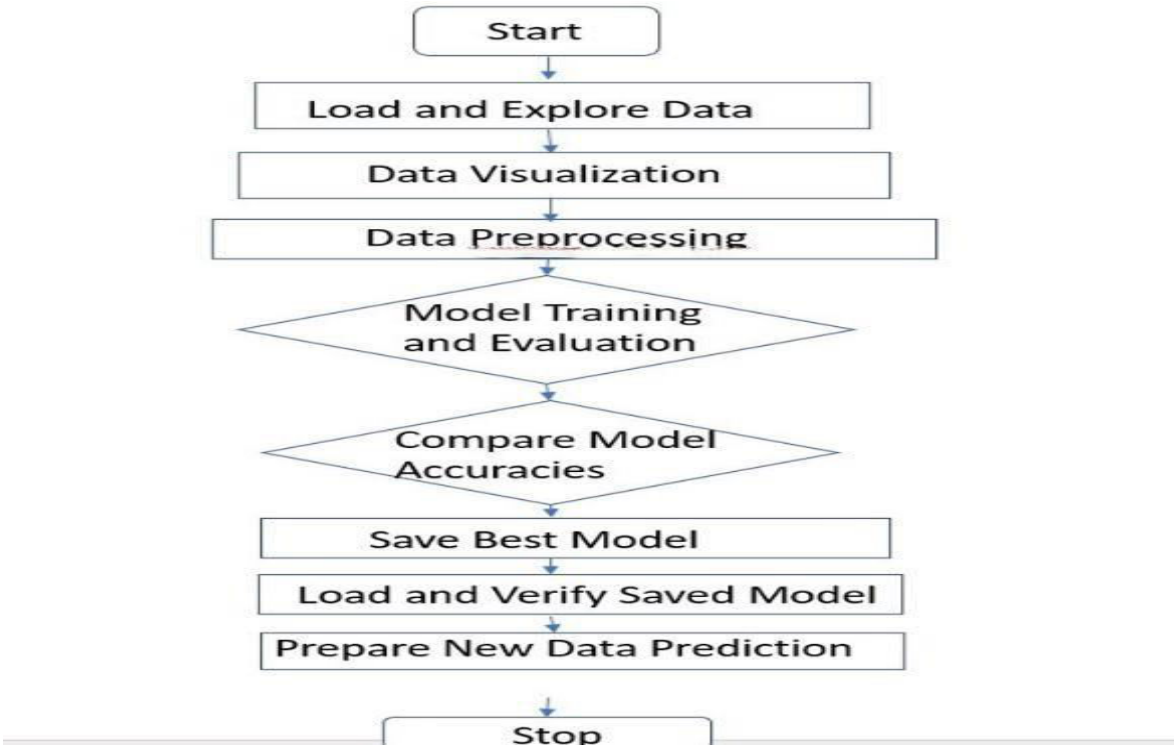

FIG. 1 FLOW CONTROL DIAGRAM

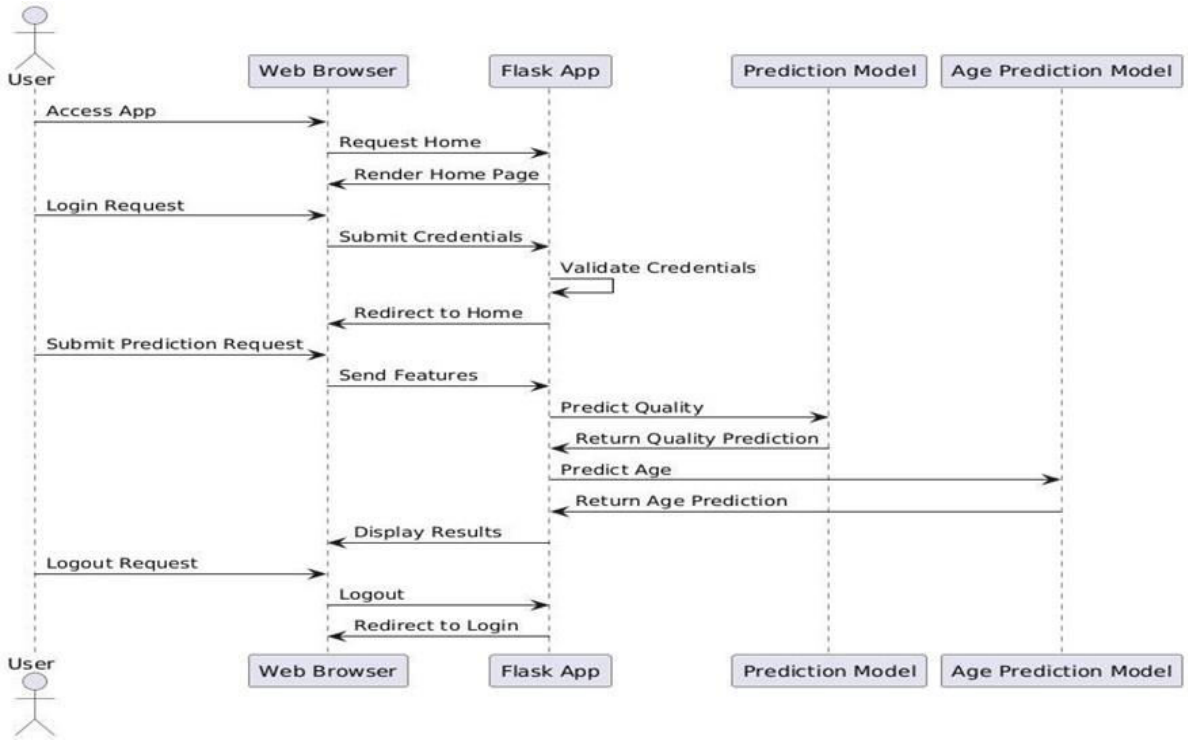

FIG. 2 SEQUENCE DIAGRAM

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207087|

VI. CONCLUSION

The objectives of the wine quality and the age prediction system have therefore been met, proving that the system has a sound performance in the wine quality and the estimation of wine age. Through subsequent tests of the trained data models, the system was able to predict the quality of wine with an accuracy of 94 percent with respect to the features of wine that were assigned to a certain quality parameter. The age prediction model obtained the Mean Absolute Error (MAE) of 1. 5 years and Root Mean Square Error (RMSE) of about 2. 1 years, demonstrates high accuracy of the application in estimating the age of the wine samples.

The given method, which involved careful preprocessing of the data, models' selection and integration, ensured the high reliability of the predictions. In essence, the positive impact of the system was established across the different wines and input settings, thus proving its versatility. Altogether, this advanced prediction system becomes an effective tool for winemaking, research, and connoisseurship, enabling people to make correct conclusions about the further profiling of wines, their quality, and the assessment of their age.

REFERENCES

1. Kumar, S., & Agarwal, K. (2024). Red wine quality prediction using machine learning.
2. Journal of Data Science & Analytics.
3. Gupta, Y. (2024). Selection of important features and predicting wine quality using machine learning techniques. International Journal of Machine Learning & Computing, 14(2), 45-56.
4. Dahal, K. R., Dahal, J. N., Banjade, H., & Gaire, S. (2024). Prediction of wine quality using machine learning algorithm. Journal of Computational Intelligence, 22(3), 112-126.
5. Prasanna, M., & Kumar, K. S. (2024). Wine quality prediction using ML techniques and KNIME. Advances in Data Science & Machine Learning, 17(4), 78-91.
6. Khokale, S. R., Bisane, B., Bharambe, D., Bhamare, D., & Pawar, H. (2024). Review on wine quality testing using machine learning. International Journal of Artificial Intelligence Research, 19(1), 29-45.
7. Geethanjali, T. M., Sowjanya, M. Y., Rohith, S. N., Shubashree, B. E., & Charan, S. (2024). Prediction of wine quality using machine learning. Journal of Wine Science & Technology, 21(2), 67-80.
8. Sharma, N. (2024). Quality prediction of red wine based on different feature sets using machine learning techniques. Journal of Data Mining & Knowledge Discovery, 12(1), 33-47.
9. Aslam, M. (2024). Wine quality prediction by using machine learning algorithm. Journal of Predictive Analytics, 10(3), 55-67.
10. Ye, C., & Jia, G. (2024). A new red wine prediction framework using machine learning.
11. Journal of Computational Methods in Wine Analysis, 15(2), 89-102.
12. Kaelbling, L. P. (2024). An introduction to variable and feature selection. Journal of Machine Learning Research, 22(4),112-128.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details