|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sl.NO | Author | Year | Methodology | Evaluation |
| 1 | Mahima, Ujjawal Gupta, Yatindra Patidar, Abhishek Agarwal and Kushall Pal Sing | 03 April 2020 | They collected and cleaned red and white wine datasets, selected influential parameters such as pH and fixed acidity, and applied dimensionality reduction techniques.Their predictive modeling algorithms utilized Random Forest and K-Nearest Neighbors (KNN) and they designed a user interface to provide real-time quality predictions.. | The models were assessed using accuracy, precision, recall, and F1-score, with the confusion matrix used to gauge performance across quality categories. Random Forest's performance was compared to other algorithms, showcasing its strengths and limitations in extreme quality classification. Future work aims to refine the dataset and improve the user interface for enhanced usability. |
| 2 | Siniša Ilić Stefan Pitulić | 16 November 2019 | The study utilized the vinho verde wine database from Portugal, containing 6497 samples of white and red wines. Six classification algorithms, including Decision Tree (J48), Random Forest (RF), Multilayer Perceptron (MLP), Support Vector Machine (SVM), Naïve Bayes Classifier (NB), and Algorithm k star (k\*), were applied to categorize wines based on their physicochemical properties | The performance of the algorithms was assessed using metrics like accuracy, Kappa statistics, Mean Absolute Error, and Recall. Random Forest emerged as the most effective algorithm, achieving over 85% accuracy in wine categorization. The k\* algorithm followed with slightly lower accuracy, while SVM performed the poorest. The study demonstrated the potential of using classification algorithms for wine quality evaluation, with Random Forest offering promising results. |
| 3 | Shruthi P | 2019 | The study collected wine samples and their attributes for quality assurance, applying data mining classification algorithms such as Naive Bayes, Simple Logistic | Using a dataset of 178 wine samples, the study compared the efficiency of the classification algorithms. Naive Bayes emerged as the most accurate classifier, offering potential benefits in consumer decision-making, fraud reduction, and government quality certification. Future research could explore additional algorithms for wine quality assessment and cost prediction. |
| 4 | Sunny Kumar, Kanika Agrawal, Nelshan Mandan | JAN 2022 | The study used data mining techniques to predict red wine quality, employing a dataset sourced from reliable sources. Three classification algorithms—Random Forest, Support Vector Machine (SVM), and Naïve Bayes—were applied and evaluated based on accuracy, precision, and other metrics. | upport Vector Machine (SVM) achieved the highest accuracy of 67.25%, followed by Random Forest at 65.83%, and Naïve Bayes at 55.91%. SVM performed best, suggesting its effectiveness in red wine quality prediction. Future research may enhance accuracy by combining features from different techniques and optimizing hyperparameters. |
| 5 | Nitin Khilari1, Pravin Hadawale 2, Hasan Shaikh3, Sachine Kolase4 | DEC 2021 | The study used machine learning algorithms, including Logistic Regression, Decision Tree, Random Forest, SVM, Ada Boost Classifier, and Gradient Boosting Classifier, to predict wine quality. Both fundamental and technical analysis approaches were compared using visualizations. | Random Forest Algorithm performed best, achieving 92% accuracy in predicting wine quality. The study highlighted key features associated with good quality wines and suggested focusing on essential features for improved prediction accuracy. Future research could explore larger datasets and alternative algorithms for wine quality prediction. |
| 6 | Zhou Tingwei\* | 2021 | The study used active learning with K-Nearest Neighbor and ranked batch-mode sampling to predict red wine quality, starting with an initial training set of 60 labeled data points and 15 iterations | The active learning method achieved an 88.2% prediction accuracy for red wine quality. Compared to the iris flower dataset, accuracy was slightly lower due to factors like low correlation and incomplete data. Future research should optimize parameters and collect more data to improve accuracy. |
| 7 | Mohit Gupta, Vanmathi C | MAY 2021 | The study used Decision Tree and Random Forest algorithms to predict wine quality based on physicochemical properties, comparing 1599 red wine and 4898 white wine samples. | he improved Multiple Regression Model (MP5) performed best, outperforming both Decision Tree and Random Forest algorithms, particularly for red wine. The study highlighted the importance of alcohol content and fermentation methods in enhancing wine quality. Future research will explore broader datasets and additional techniques to improve predictive accuracy. |