RED AND WHITE WINE QUALITY ANALYSIS

Pietro Bogani
Politecnico di Milano
Valencia

ABSTRACT

This study centers on the development and training of Neural Networks (NNs) using data on the chemical properties of various red and white wines, alongside an aggregated quality rating based on multiple assessments. In the first phase, an NN model is constructed to predict the wine quality grade using its chemical features, while also examining potential differences between red and white wines. The accuracy achieved by the NN is not entirely satisfactory, suggesting opportunities for enhancement through the inclusion of additional covariates. This limitation may be attributed to the complexity and subjectivity of wine preferences, which are challenging to capture through a limited set of chemical attributes. In the second phase, the focus shifts to classifying wines as red or white based on the available chemical features. The proposed model achieves high accuracy in this classification task. A feature selection analysis is also conducted to identify the chemical components that contribute most significantly to distinguishing between red and white wines.

Keywords Wine, Neural Network, Classification

1 Introduction

1.1 Data

The dataset, sourced from the UCI Machine Learning Repository [1], comprises two files, `wine_red.csv` and `wine_white.csv`, each containing 12 attributes. These include 11 chemical features (e.g., residual sugar and free sulfur dioxide) and a quality score (rated 1–10). The red wine dataset has 1599 observations, while the white wine dataset includes 4898 entries. This class imbalance requires consideration in model development.

1.2 Descriptive Analysis

We examine the distribution of wine quality by type, producing bar plots to visualize differences. Initial observations suggest that white wines received higher average quality ratings. At first glance, the two distributions appear different, suggesting that white wines are more appreciated than red wines by survey participants. To confirm this observation, a hypothesis test is performed. In order to have a robust result, which does not depend on normality assumptions, a permutational test is chosen. A permutational test is a non-parametric statistical test that assesses the significance of an observed effect by randomly permuting the data labels and comparing the observed statistic to a null distribution generated from the permutations.

The test statistic is defined as:

$$T_0 = |X_w - X_r|$$

where X_w is the empirical mean quality of white wine and X_r is the empirical mean quality of red wine.

The p-value is computed as:

$$p\text{-value} = \frac{\#(T^* \geq T_0)}{N}$$

where T^* is the value of the test statistic after a single permutation, and N is the total number of permutations (set to 1000 for this test).

The p-value obtained is 0, indicating a statistically significant difference in the quality distributions between the two wine types.

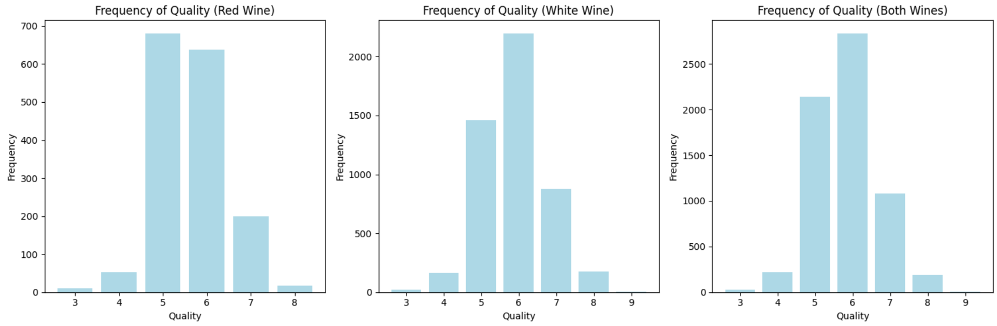

Figure 1: Histogram of Quality for red and white wines

2 Neural Network for Quality Prediction

2.1 Red Wine

Several neural networks with varying architectures, activation functions, and optimizers were tested to predict red wine quality. Results are summarized in Table 1. Notably, adding batch normalization layers did not improve accuracy. Despite adjustments in architecture and epochs, the accuracy did not exceed 61%.

Table 1: Neural Network Architectures and Performance for Red Wine Quality Prediction

Architecture	Activation	Optimizer	Epoch	Accuracy
Dense(100) + Dense(50) + Dense(30) + Dropout	**ReLU**	**Adam**	**200**	**0.60**
Dense(100) + Dense(50) + Dense(30) + Dropout	ReLU	SGD	200	0.56
Dense(100) + Dense(50) + Dense(30) + Dropout	ReLU	Adagrad	200	0.53
Dense(100) + Dense(50) + Dense(30) + Dropout	LeakyReLU	Adam	200	0.59
Dense(100) + Dense(50) + Dense(30) + Dropout	ELU	Adam	200	0.58
Dense(100) + Dense(50) + Dense(30) + Dropout	selu	Adam	200	0.59

The addition of BatchNormalization layers do not lead to an increase in accuracy. Given the best performing model, we focus on that and increase the number of epoch and the complexity of the architecture, trying to improve the accuracy.

Table 2: Again Neural Network Architectures and Performance for Red Wine Quality Prediction

Architecture	Activation	Optimizer	Epoch	Accuracy
Dense(100) + Dense(50) + Dense(30) + Dropout	**ReLU**	**Adam**	**1000**	**0.61**
Dense(500) + Dense(300) + Dense(100) + Dropout	ReLU	Adam	200	0.60
Dense(500) + Dense(300) + Dense(100) + Dropout	**ReLU**	**Adam**	**500**	**0.61**
Dense(500) + Dense(300) + Dense(100) + Dense(100) + Dropout	ReLU	Adam	200	0.60

2.2 White Wine

Similar neural network configurations were applied to predict white wine quality. Table 3 summarizes the results. Like with red wine, increasing model complexity and computational load had little effect on accuracy, indicating that the available chemical features may not sufficiently capture wine quality.

Table 3: Neural Network Architectures and Performance for White Wine Quality Prediction

Architecture	Activation	Optimizer	Epoch	Accuracy
Dense(100) + Dense(50) + Dense(30) + Dropout	ReLU	Adam	200	0.52
Dense(100) + Dense(50) + Dense(30) + Dropout	ReLU	SGD	200	0.48
Dense(100) + Dense(50) + Dense(30) + Dropout	ReLU	Adagrad	200	0.44
Dense(100) + Dense(50) + Dense(30) + Dropout	**LeakyReLU**	**Adam**	**200**	**0.54**
Dense(100) + Dense(50) + Dense(30) + Dropout	ELU	Adam	200	0.52
Dense(100) + Dense(50) + Dense(30) + Dropout	selu	Adam	200	0.53

The addition of BatchNormalization layers do not lead to an increase in accuracy. Given the best performing model, we focus on that and increase the number of epoch and the complexity of the architecture, trying to improve the accuracy.

Table 4: Again Neural Network Architectures and Performance for Combined Wine Quality Prediction

Architecture	Activation	Optimizer	Epoch	Accuracy
Dense(100) + Dense(50) + Dense(30) + Dropout	**LeakyReLU**	**Adam**	**1000**	**0.55**
Dense(500) + Dense(300) + Dense(100) + Dropout	LeakyReLU	Adam	200	0.53
Dense(500) + Dense(300) + Dense(100) + Dropout	LeakyReLU	Adam	500	0.54
Dense(500) + Dense(300) + Dense(100) + Dense(100) + Dropout	LeakyReLU	Adam	200	0.53

2.3 Both Wines

Table 5: Neural Network Architectures and Performance for Combined Wine Quality Prediction

Architecture	Activation	Optimizer	Epoch	Accuracy
Dense(100) + Dense(50) + Dense(30) + Dropout	ReLU	Adam	200	0.57
Dense(100) + Dense(50) + Dense(30) + Dropout	ReLU	SGD	200	0.51
Dense(100) + Dense(50) + Dense(30) + Dropout	ReLU	Adagrad	200	0.45
Dense(100) + Dense(50) + Dense(30) + Dropout	**LeakyReLU**	**Adam**	**200**	**0.58**
Dense(100) + Dense(50) + Dense(30) + Dropout	ELU	Adam	200	0.56
Dense(100) + Dense(50) + Dense(30) + Dropout	selu	Adam	200	0.56

Table 6: Again Neural Network Architectures and Performance for Combined Wine Quality Prediction

Architecture	Activation	Optimizer	Epoch	Accuracy
Dense(100) + Dense(50) + Dense(30) + Dropout	LeakyReLU	Adam	1000	0.56
Dense(500) + Dense(300) + Dense(100) + Dropout	**LeakyReLU**	**Adam**	**200**	**0.58**
Dense(500) + Dense(300) + Dense(100) + Dropout	**LeakyReLU**	**Adam**	**500**	**0.58**
Dense(500) + Dense(300) + Dense(100) + Dense(100) + Dropout	LeakyReLU	Adam	200	0.56

No significant changes when using both red and white wines together.

3 Prediction of Wine Color

3.1 Neural Network for Color Classification

Several neural networks (NNs) were developed to perform binary classification, specifically predicting whether a given wine is red or white based on its 11 chemical features and quality score. Notably, there is a significant class imbalance in the dataset, with 75% of wines classified as white and only 25% as red. Due to this imbalance, accuracy alone may not adequately reflect model performance, so the AUC-ROC score is also reported. The advantage of the AUC-ROC score is that it remains unaffected by class imbalance, providing a more reliable metric for model evaluation. The constructed NNs vary in architecture, activation functions, and optimizers. The results are summarized in the table below:

Table 7: Neural Network Architectures and Performance for Wine Color Prediction (Initial Models)

Architecture	Activation	Optimizer	Epoch	Accuracy	AUC-ROC
Dense(100) + Dense(50) + Dense(30) + Dropout	ReLU	Adam	200	0.96	0.994
Dense(100) + Dense(50) + Dense(30) + Dropout	ReLU	SGD	200	0.95	0.980
Dense(100) + Dense(50) + Dense(30) + Dropout	ReLU	Adagrad	200	0.88	0.914
Dense(100) + Dense(50) + Dense(30) + Dropout	**LeakyReLU**	**Adam**	**200**	**0.97**	**0.994**
Dense(100) + Dense(50) + Dense(30) + Dropout	**ELU**	**Adam**	**200**	**0.98**	**0.9965**
Dense(100) + Dense(50) + Dense(30) + Dropout	selu	Adam	200	0.98	0.9963

The addition of BatchNormalization layers does not lead to an increase in accuracy. Given the best performing model, we focus on that and increase the number of epochs and the complexity of the architecture, aiming to improve both the accuracy and the AUC-ROC score.

Table 8: Neural Network Architectures and Performance for Wine Color Prediction (Enhanced Models)

Architecture	Activation	Optimizer	Epoch	Accuracy	AUC-ROC
Dense(100) + Dense(50) + Dense(30) + Dropout	**ELU**	**Adam**	**1000**	**0.986**	**0.997**
Dense(500) + Dense(300) + Dense(100) + Dropout	ELU	Adam	200	0.978	0.9946
Dense(500) + Dense(300) + Dense(100) + Dropout	ELU	Adam	500	0.978	0.9929
Dense(500) + Dense(300) + Dense(100) + Dense(100) + Dropout	ELU	Adam	200	0.975	0.9928

The introduction of a richer and more sophisticated model brings negligible benefits to both accuracy and ROC-AUC score.

3.2 Classification Tree for Color Prediction

In this section another technique to perform binary classification it's explored: the classification tree. This is useful to have an idea of the performance of a simpler model, but also to perform feature selection and understanding which covariates are more useful to distinguish between red and white wine. The model achieves a 98% accuracy and ROC-AUC score of 0.972, which are comparable to results obtained in the previous neural network models. The confusion matrix is reported below:

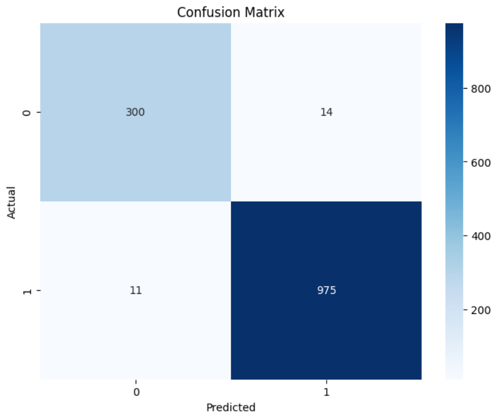

Figure 2: Confusion Matrix

We now aim to perform feature selection. The feature importances according the classification tree algorithm are reported below:

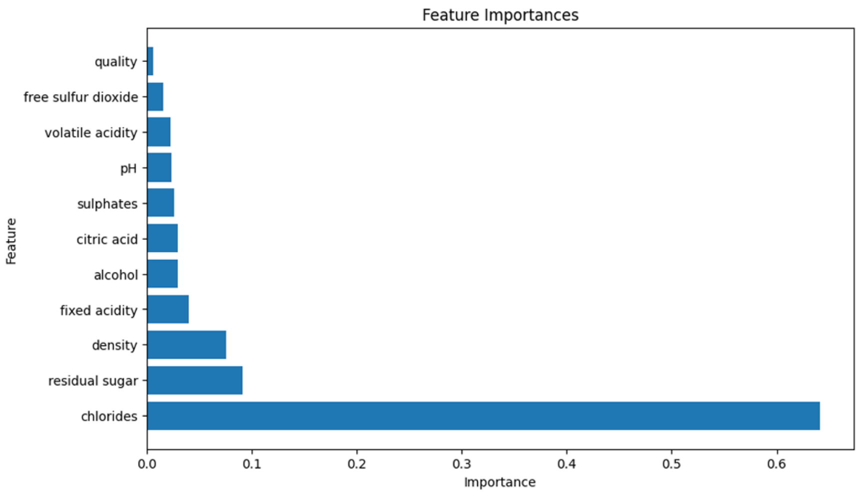

Figure 3: Feature Importances

Chlorides is considered the most significant and useful discriminator between red and white wines. To further confirm that chlorides is a useful variable for classifying red and white wine, I perform a hypothesis test:

$$H_0 : \mu_0 = \mu_1 \quad \text{vs} \quad H_1 : \mu_0 \neq \mu_1$$

Where: μ_0 is the mean chloride level in red wines and μ_1 is the mean chloride level in white wines. A non-parametric Mann-Whitney test is performed to avoid making assumptions about normality. The test yields a p-value of exactly 0, providing strong evidence in support of rejecting H_0, and concluding that chloride is an effective discriminator between red and white wines.

Lastly, a plot of the empirical distributions of chlorides for the two classes is produced:

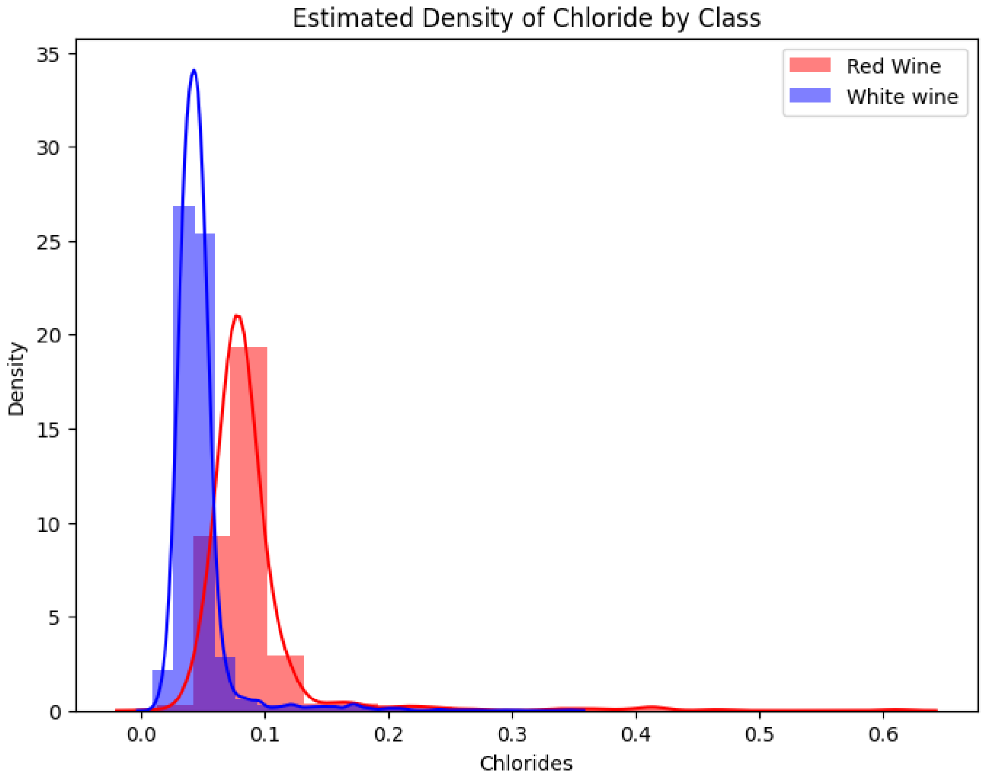

Figure 4: Chloride in the two wines

The plot intuitively denotes a significant difference both in mean and distribution.

4 Conclusion

The study is divided into two main parts. In the first part, we aimed to investigate the predictive capability of chemical features in determining the quality of wine. Despite extensive experimentation with various neural network architectures, optimizers, and activation functions, the models generally failed to achieve satisfactory performance, with accuracy rarely exceeding 60%. This suggests that the available covariates, which primarily capture the chemical characteristics of the wines, are insufficient to explain the subjective quality ratings given by consumers. This limitation may stem from the chosen covariates not adequately representing the chemical attributes that make a wine appealing or from the complexity of wine quality itself, which—being an average of multiple reviews—may require additional information beyond chemical properties. In the second part, the objective was to classify the type of wine (red or white) based on its chemical properties. Here, the models demonstrated promising results, with the best-performing models achieving accuracy and ROC-AUC scores close to perfection. This indicates a strong correlation between chemical composition and wine color. Simple models, such as classification trees or shallow neural networks, were sufficient to achieve high accuracy without requiring additional model complexity. Feature selection analysis further revealed that chlorides are the most effective discriminator between red and white wines among all chemical properties tested.

References

[1] UCI Machine Learning Repository, `https://archive.ics.uci.edu/`.