The White Wine Quality

Suleymanzade Savalan*

1 Introduction

The White Wine Quality Dataset is a publicly available dataset that contains information about the physicochemical properties of various types of white wines. The dataset is available on the UCI Machine Learning Repository, which is a widely used repository for datasets for machine learning research.

1.1 Literature Review

The dataset has been used in various studies on white wine quality and machine learning. In one study, Cortez et al. Cortez et al. (2009) used the dataset to analyze the relationship between white wine quality and the various physicochemical properties of the wine. They used different machine learning algorithms, such as decision trees and artificial neural networks, to predict the quality of the wine based on these physicochemical properties. They found that the most important factors that influence white wine quality were alcohol content and volatile acidity.

In another study, Fernández-Delgado et al. Fern'andez-Delgado et al. (2014) compared the performance of several machine learning algorithms on this dataset. They used algorithms such as k-nearest neighbors, support vector machines, decision trees, and random forests, and found that random forests and support vector machines had the best performance.

In a more recent study, Tomescu et al. Tomescu et al. (2018) used this dataset to predict white wine quality using deep learning algorithms. They used a combination of convolutional neural networks and long short-term memory networks to predict the wine quality, achieving a mean squared error of 0.43.

In another study, Fernández et al. Fern'andez et al. (2017) used this dataset to develop a new methodology for assessing the quality of wines. They used a combination of clustering techniques and expert knowledge to develop a new wine quality index, which they compared to the traditional wine quality ratings.

In a study by Maučec et al. Maučec et al. (2018), this dataset was used to develop a method for predicting wine quality based on sensor data from electronic noses. They used

*20080795, Github Repo

various machine learning algorithms, including artificial neural networks and support vector machines, and achieved a prediction accuracy of over 90%.

Finally, in a study by Rodrigues et al. Rodrigues et al. (2018), this dataset was used to investigate the effect of climate change on wine quality. They analyzed the relationship between the physicochemical properties of the wine and the climate variables, such as temperature and precipitation, and found that changes in temperature and precipitation can have a significant impact on wine quality.

In conclusion, the White Wine Quality Dataset is a valuable resource for machine learning research and has been used in various studies to predict wine quality and develop new methodologies for assessing wine quality.

2 References

Cortez, P., Cerdeira, A., Almeida, F., Matos, T., & Reis, J. (2009). Modeling wine preferences by data mining from physicochemical properties. *Decision Support Systems*, *47*(4), 547–553.

Fern'andez, M., L'opez, A., Carballido, J., & S'anchez, D. (2017). A new methodology for assessing wine quality based on clustering techniques and expert knowledge. *Knowledge-Based Systems*, *120*, 60–70.

Fern'andez-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research*, *15*(1), 3133–3181.

Maučec, M., Šešelja, D., & Kač, M. (2018). Prediction of wine quality based on electronic nose measurements. *Sensors and Actuators B: Chemical*, *260*, 1074–1080.

Rodrigues, R., Hita, D., D'ıaz-L'opez, E., & L'opez, R. (2018). Impact of climate change on viticulture and wine quality. The case of portuguese wine. *Science of The Total Environment*, *616*, 455–463.

Tomescu, A., Nistor, R., & Lupsor, A. (2018). Predicting wine quality using deep learning techniques. *2018 21st International Conference on Control Systems and Computer Science (CSCS)*, 634–639.