Pouring Insights: Predicting Wine Quality

Introduction:

Today, machine learning is becoming more and more popular for predicting wine quality. Basically, there are some algorithms that can determine the wine quality. So, we made the decision to dissect this into a more thorough and technical synopsis, and we chose "Wine Quality Prediction" as our research topic. From the dataset that we chose, we will analyze the ingredients of 32485 wine instances to predict the quality and type of wine. According to the dataset, the input variables for the machine learning-based prediction of wine quality are derived from laboratory-based physicochemical tests. Therefore, an accurate understanding of the physicochemical characteristics of both red and white wine is essential to the success of a machine learning model that predicts wine quality. So, we set some questions to help us analyze the data accurately.

Smart questions:

1. Do certain types of wine (red or white) tend to have higher quality scores on average?
2. Can machine learning models accurately predict wine quality based on its chemical composition, and if yes, which algorithms perform the best?
3. Can we quantitatively measure the correlations between all attributes and wine quality ratings?
4. What is the range of wine quality scores, and how can we improve this range through analysis and recommendations?
5. How will understanding and improving wine quality benefit winemakers, distributors, and wine consumers?

Modeling techniques:

1. Logistic Regression
2. K-Nearest Neighbours
3. Decision trees

Dataset link:

https://www.kaggle.com/code/nicoletacilibiu/wine-quality-prediction/input

Github link:

https://github.com/sanika71118/team_4