PROJECT 4 WINE ANALYSIS

11.25.2024

Predicting Wine Quality

Collaborators:

Chuck Bui

Jack Jeffries

Beau Massie

Christopher Turner

Problem Statement

VinoVista, a renowned winery, is committed to producing consistently exceptional wines. However, they have observed variability in the quality of their white wine batches. To address this challenge, VinoVista seeks a predictive model that can accurately assess the quality of wine batches based on their chemical properties before bottling.

Project Goal:

The primary objective of this project is to develop a robust machine learning model capable of predicting wine quality on a scale of 0-10. This model will utilize a dataset containing various chemical properties of wine, such as acidity, pH, residual sugar, chlorides, free sulfur dioxide, total sulfur dioxide, density, sulfates, and alcohol content.

Resources:

Wine Quality Dataset from UC Irvine Machine Learning Repository with 4899 samples

QUICK TAKE

- This data set was chosen because it contains information about the chemical propterties of white wines, such as acidity, sugar levels, and alcohol content and the data is modeled after the physicochemical wine tests.
- The number of samples in our dataset after filtering is 3,090. We used Quality as our target and there are 11 total features.
- We cleaned our database by checking for null values, dropping any outliers in ‘total sulfur dioxide’ above 150 as that is likely a data entry error. Our quality values range from 1-10, so we created binary classification to use in our machine learning models. We chose 7 and above as good; 6 and below as not good.
- We checked for multi collinearity among our features for possible features that could be dropped to increase our precision and recall values and filtered our data to only keep wines that pH values between 3 and 4, as this is the nominal level for white wines.

LOAD THE DATA AND BEGIN CLEANING

CODING PAGE 2

CODING PAGE 3?

VIZ WALKTHROUGH PAGE 1

VIZ WALKTHROUGH PAGE 1

CONCLUSION

Data Insights

- Residual sugar and density are critical features for predicting wine quality.
- Chemicals like pH, alcohol content, and chlorides also play a significant role

Future Steps

- Incorporate external features like region or grape variety
- Chemicals like pH, alcohol content, and chlorides also play a significant role

WRAP AND SUGGESTIONS TO COMPANY

suggestions

- Suggestions for packaging wine or marketing directly to the consumer
- Teach the consumer with features on the labels and allow them to be part of the process, thus making them feel more loyal to the brand