DSCI 310: Predicting Wine Cultivars

Zhibek Dzhunusova, Andrea Jackman, Kaylan Wallace & Chuxuan Zhou

Table of contents

Summary

In this project, we will predict what cultivar a wine was derived from based on its chemical properties.

The data was sourced from the UCI Machine Learning Repository (Aeberhard and Forina 1991). It contains data about various wines from Italy derived from three different cultivars. Each row represents the chemical and physical properties of a different wine, such as its concentration of alcohol, magnesium level and hue.

Introduction

Wine is a beverage that has been enjoyed by humans for thousands of years (Fehér, Lengyel, and Lugasi 2007). Consequently, humans have a long agricultural history with the grape plant which has led to the development of many different cultivars: grape plants selected and breed for their desirable characteristics (Harutyunyan and Malfeito-Ferreira 2022). Our dataset contains information about twelve chemical properties of 178 red wines made from three grape cultivars in Italy.

The recorded chemical properties include:

1. Alcohol content
2. Malic acid (gives the wine a fruity flavour)
3. Ash (left over inorganic matter from the wine-making process)
4. Alkalinity of ash (ability to resist acidification)
5. Magnesium, total phenols (contribute to bitter flavour of wine)
6. Flavanoids (antioxidants that contribute to bitter flavour and aroma of wine)
7. Nonflavanoid phenols (weakly acidic)
8. Proanthocyanins (bitter smell)