Wine Reviews

Elizabeth Webster
November 2022

Introduction

This project explores the use of Natural Language Processing in order to understand how well a wine will be rated.

I will also be utilizing a Recommendation System to understand recommended wines, giving an idea of wines to create and who to market to.

Outline

- Business Problem
- Dataset
- Multinomial Naive Bayes Model
- Description Keywords
- Wine Varieties
- Recommendation System
- Recommendations
- Next Steps

Business Problem

Stakeholder:

- Small Walla Walla winery
- Just starting out
- Only producing a few wines currently

Deliverables:

- Model to assess current wines
- Keywords and varieties that generate high scores
- Recommendation System for tasters

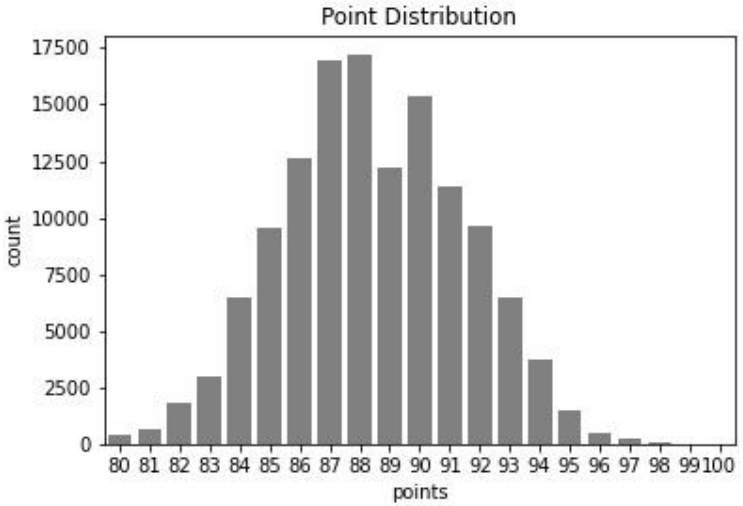
Dataset

Dataset from Wine Enthusiast

- Information on 130k wines including winery, variety, taster name, and description.

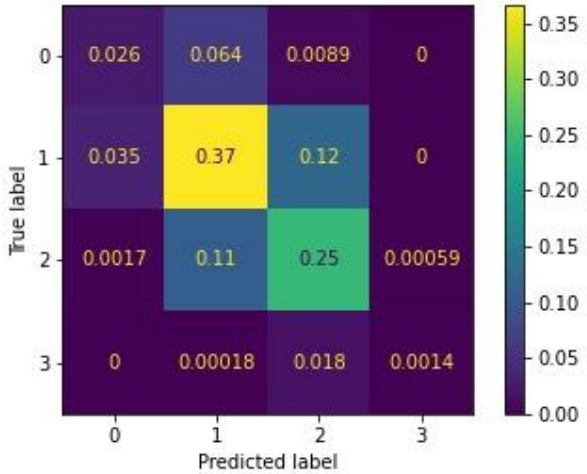
Target Value - Points

- 80-84 points: Acceptable
- 85-89 points: Good
- 90-94 points: Very Good
- 95-100 points: Outstanding

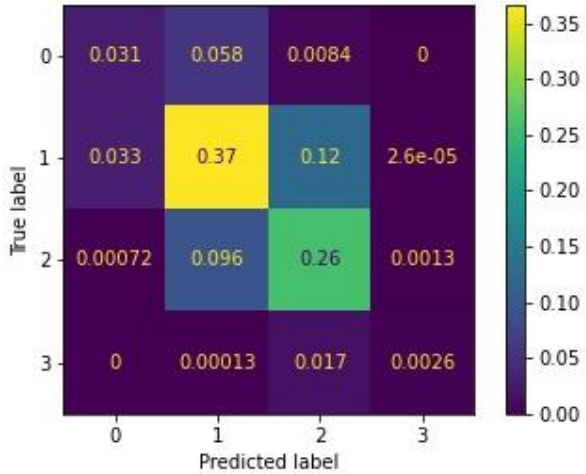


Multinomial Naive Bayes Model

Baseline Model:



Final Model:



- Not much difference in Matrices
- Model has difficulty predicting outlier classes - Acceptable and Outstanding

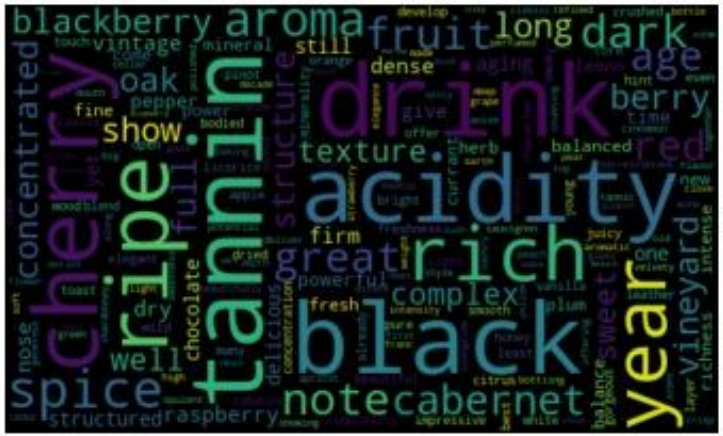
Description Keywords

Acceptable - Green, simple, bitter

Good - Fresh, crisp, berry

Description Keywords

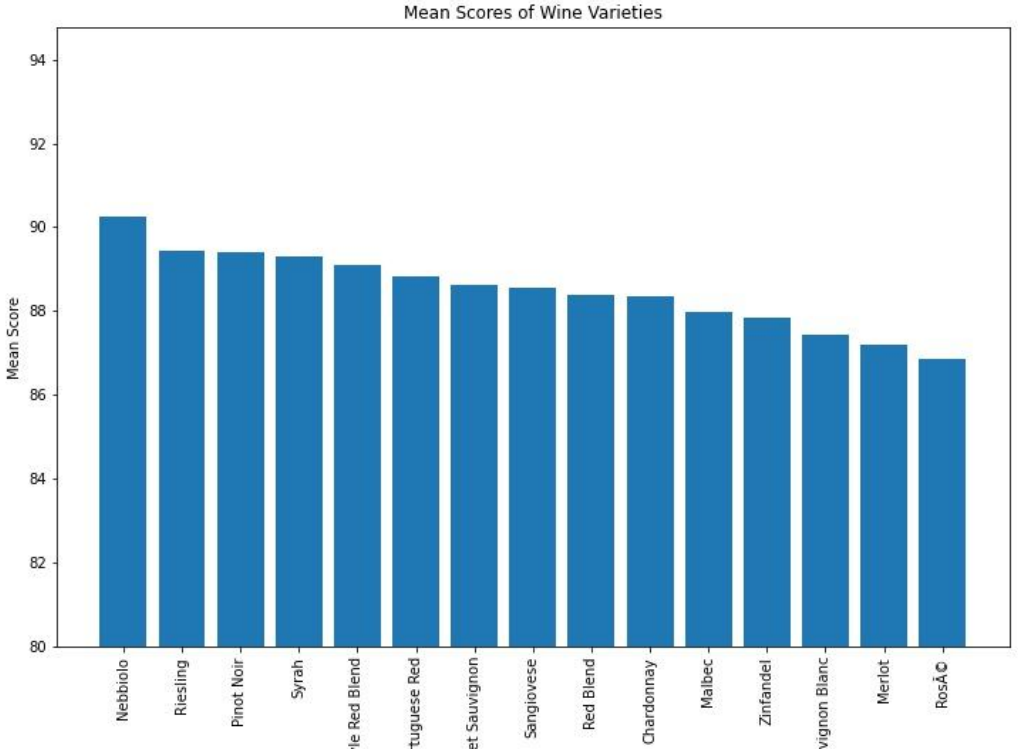
Very Good - Texture, rich, ripe



Outstanding - Complex, spice, aging

Wine Varieties

Highest Mean Scores:



Wine Varieties

Most Outstanding Scores:

- Pinot Noir - 402 times
- Cabernet Sauvignon - 274 times
- Bordeaux-style Red - 256 times
- Chardonnay - 235 times
- Riesling - 194 times

Perfect Scores:

- Bordeaux-style Red - 5 times
- Syrah - 2 times

Best Overall Potential:

- Bordeaux-style Red Blend
- Pinot Noir
- Syrah

Recommendation System

Steps I took:

- Cleaned Data
- Re-formatted Data
- Compared Models
- Built SVD Model
- Made Predictions

Findings:

Most common varieties:

- Pinot Noir
- Cabernet Sauvignon
- Syrah
- Riesling
- Chardonnay

Tasters for our region:

- Sean Sullivan
- Paul Gregutt
- Virginie Boone

Recommendations

- Naive Bayes Model - Use this model to understand who well current wines would be received.

- Keywords - Produce complex, aged red wines with deep flavor profiles. Explore flavors such as spice, pepper, and chocolate.

- Varieties to Produce:
  - Bordeaux-style Red Blend
  - Pinot Noir
  - Syrah

- Recommendation Systems - Produce Pinot Noirs, Cabernets, Syrahs, Rieslings, and Chardonnays and send them to Sean, Paul, or Virginie for tasting.

Next Steps

- Putting Model into Action - Gather the data from our current wines to predict scores.
- Understanding Successful Wineries - Explore which wineries are producing highly rated wines.
- Cold Start Problem - Making recommendations for new users

Thank You!

Email: eaw524@gmail.com

GitHub: https://github.com/elizabeth524