

A photograph of a long, narrow wine cellar aisle. The walls are lined with wooden racks filled with numerous wine bottles, their labels facing outwards. The floor is made of polished wood, reflecting the light from a single, glowing pendant lamp hanging from the ceiling. The perspective leads the eye down the center of the aisle towards a vanishing point in the distance.

Wine Ratings: **Science or Bullshit?**

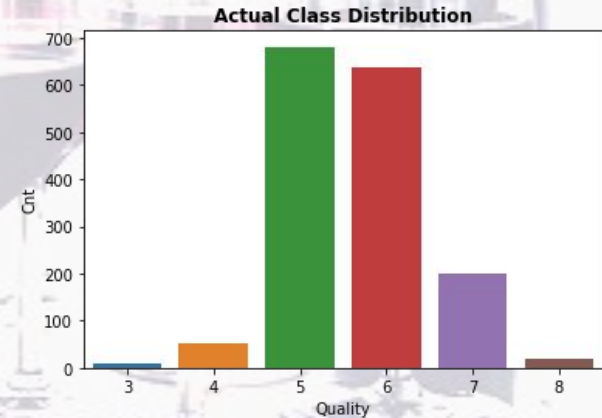
Data Summary

Interval Features (based on physicochemical tests):

- Free Sulfur Dioxide
- Total Sulfur Dioxide
- Density
- pH
- Sulphates
- Alcohol
- Fixed Acidity
- Volatile Acidity
- Citric Acid
- Residual Sugar
- Chlorides

Ordinal Target (based on sensory data):

- Quality (score between 0 and 10)

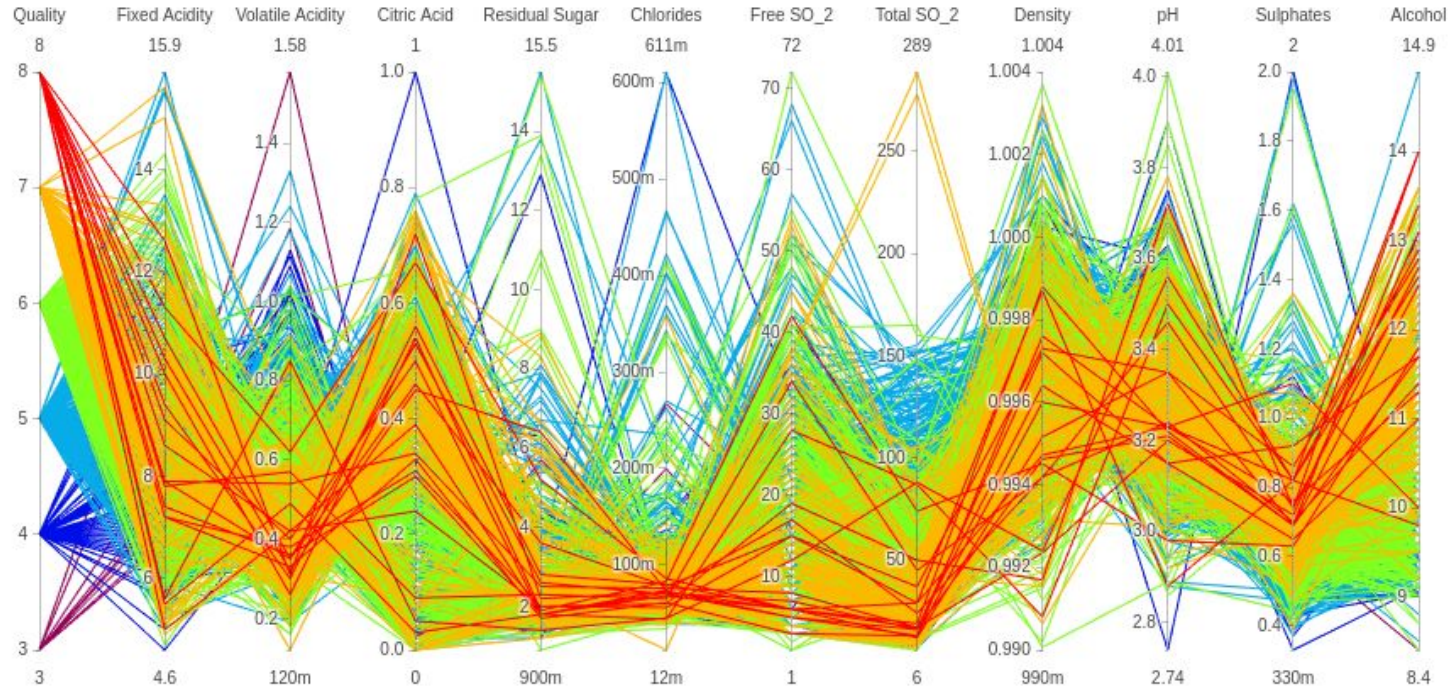


P. Cortez, A. Cerdeira, F. Almeida, T. Matos and J. Reis. (2009)

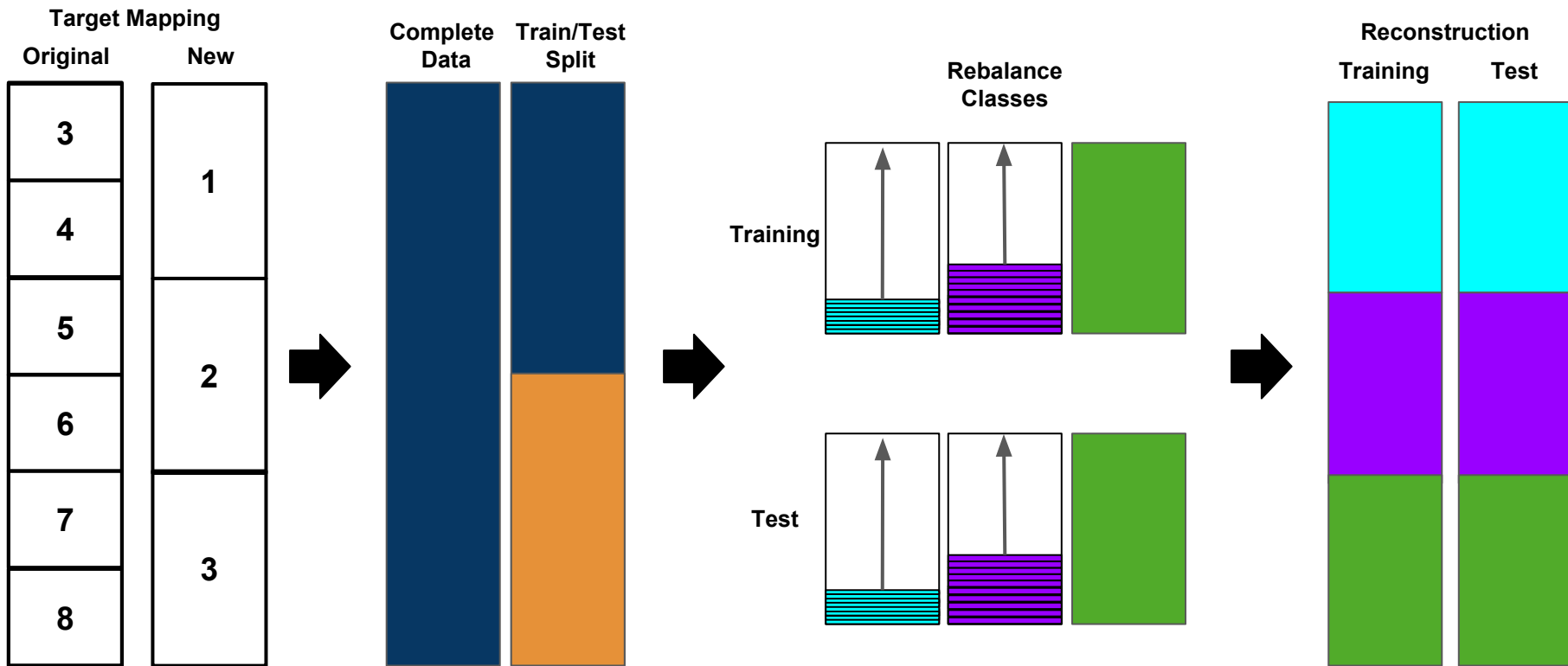
Modeling wine preferences by data mining from physicochemical properties. In Decision Support Systems, Elsevier, 47(4):547-553, 2009.

UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml/datasets/Wine+Quality>]. Irvine, CA: University of California, School of Information and Computer Science.

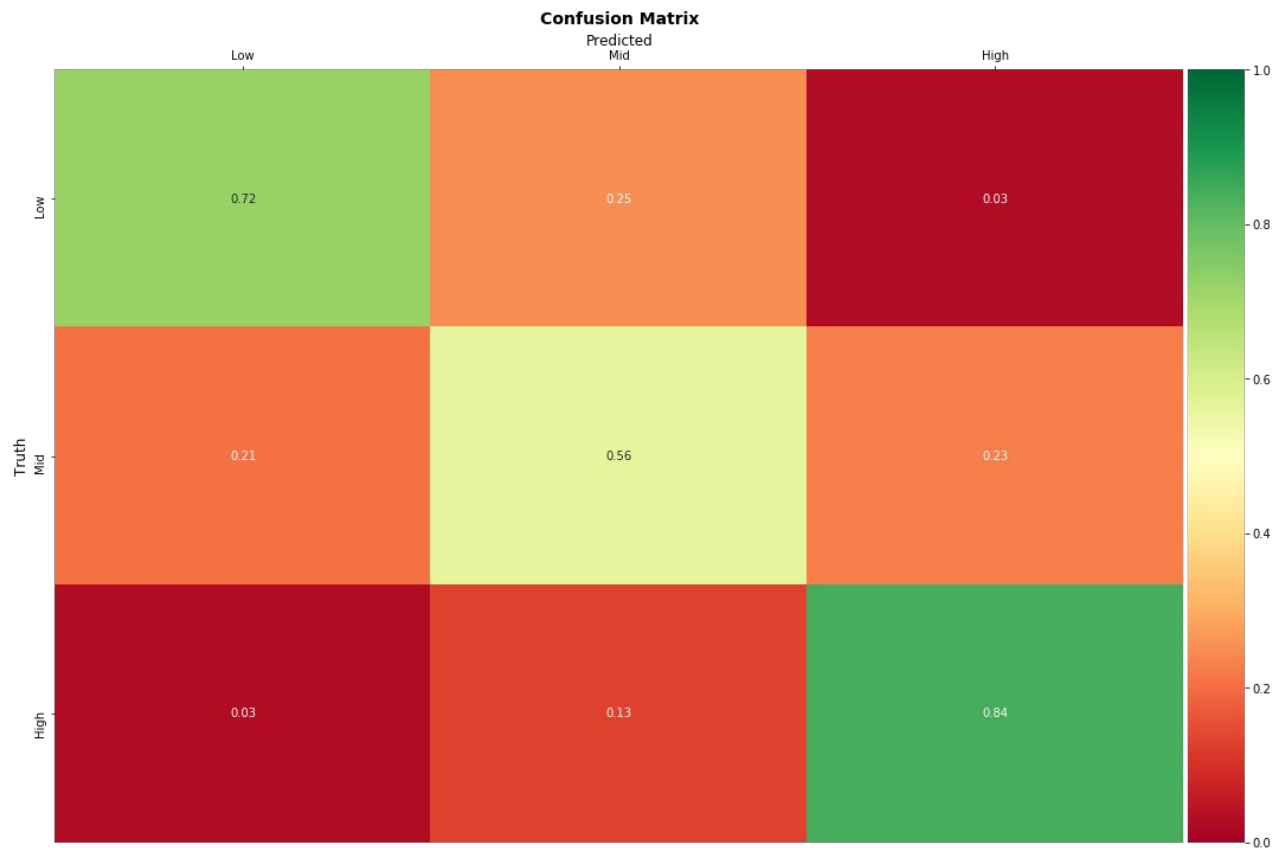
Exploratory Analysis



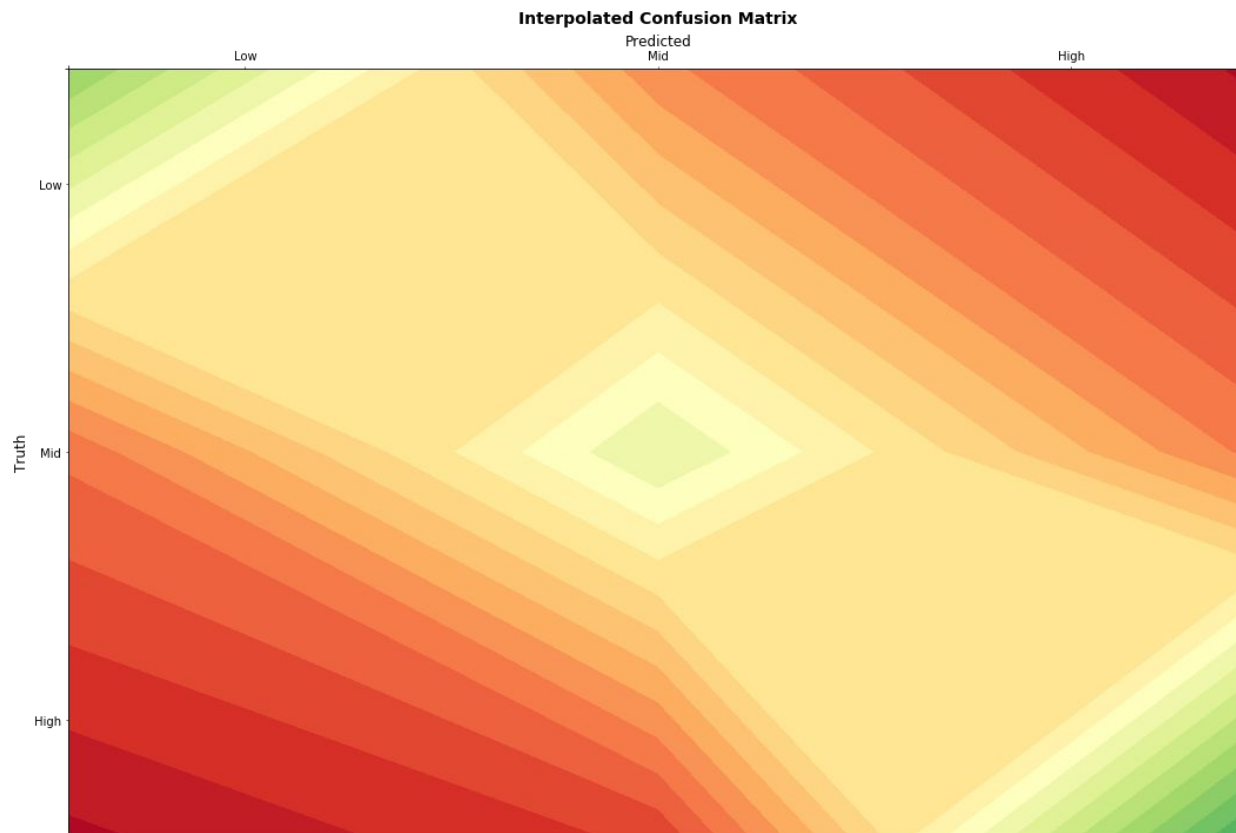
Data Preprocessing



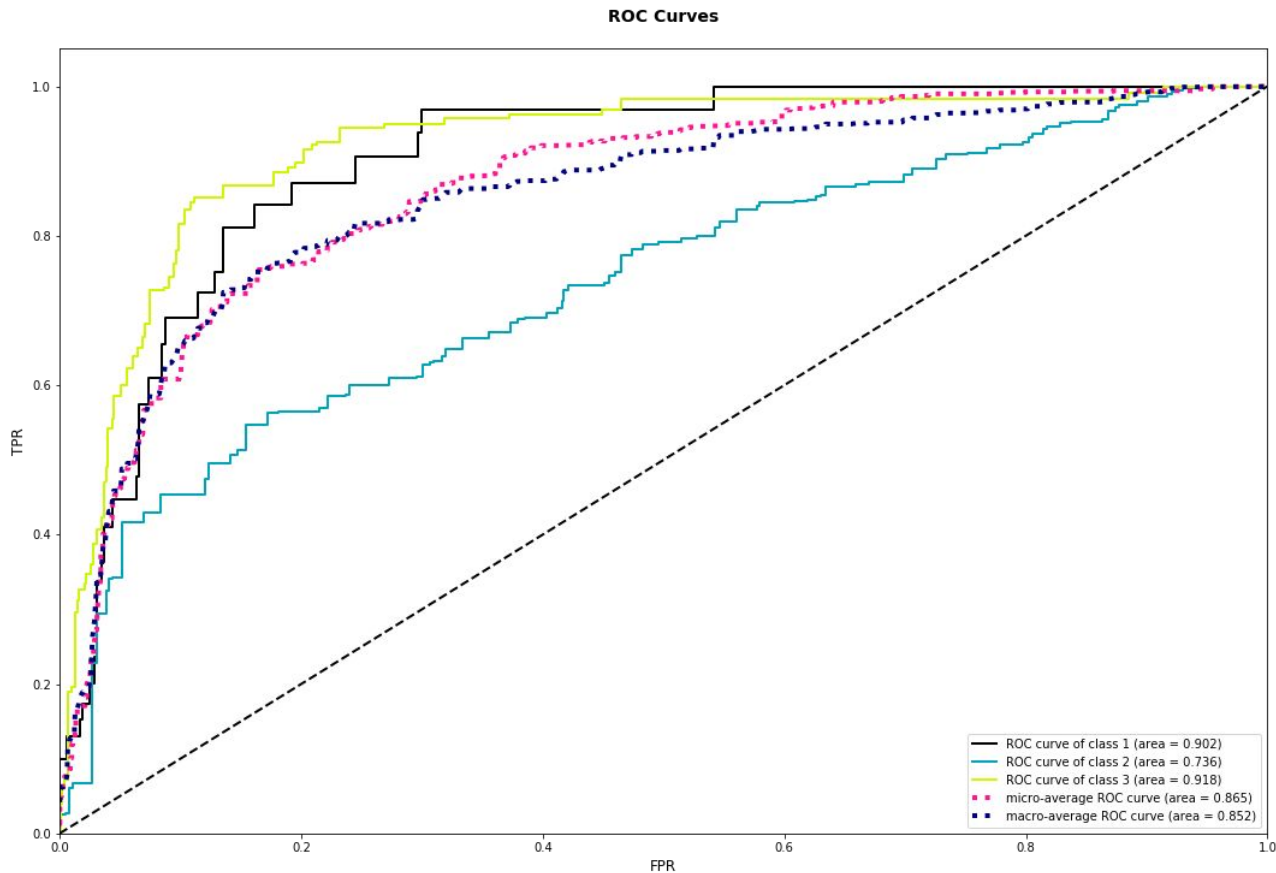
Classification Results



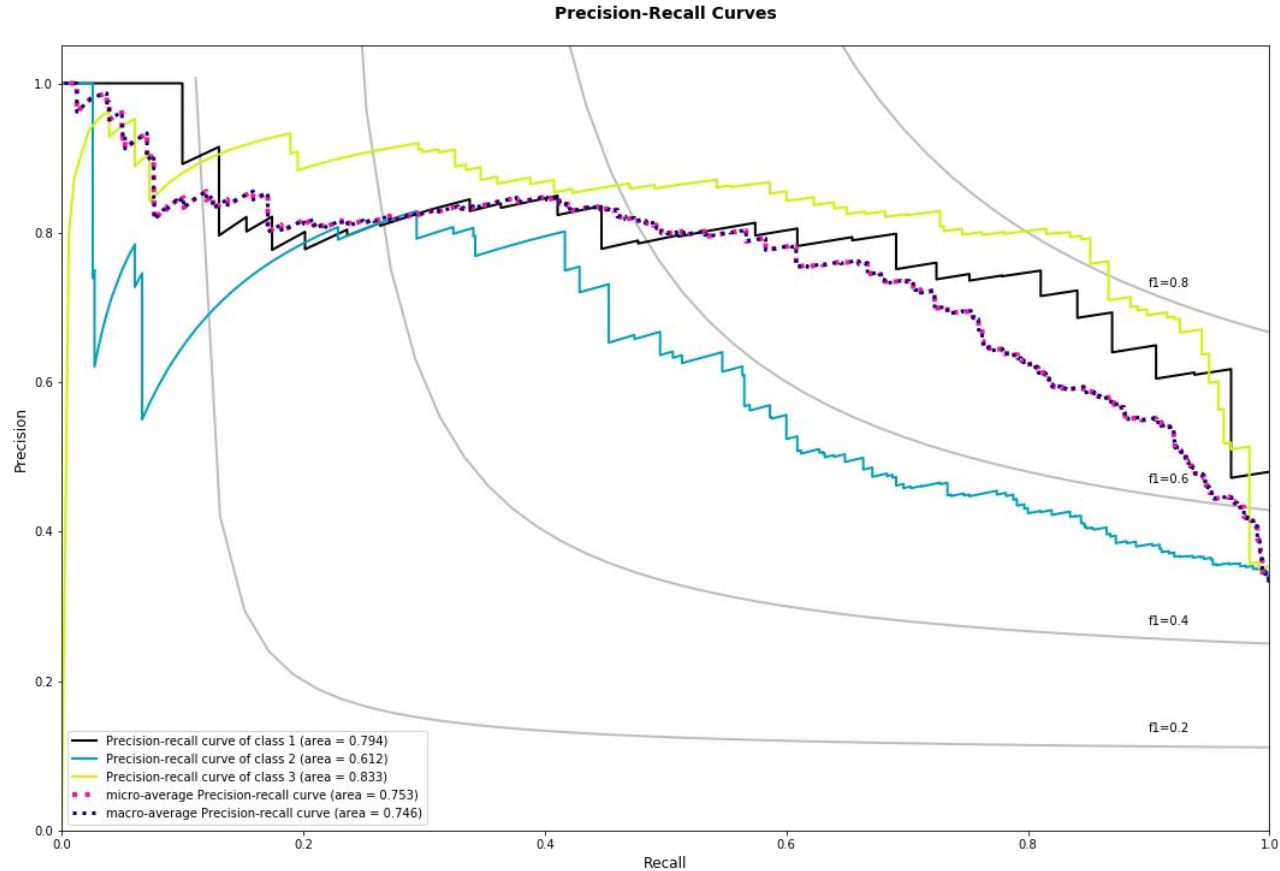
Are Not Black & White



Don't Evaluate ROC Curves in Isolation



Pair with P-R Curve Analysis





Flask App Demo

Wine Tasting is Junk Science

- **Because it is almost entirely subjective**, and it can be influenced by so many other factors. We taste with our eyes, ears, noses, and even our sense of touch. We taste with our emotions, and our state of mind. This has been demonstrated time after time after time.
- **There is no hard science involved in reviewing wine**, no real way to quantify results, no test cases, and certainly no verifiable set of standards that everyone adheres to. Everyone makes up their own processes for reviewing from Wine Spectator all of the way down to the most recent person who just discovered how easy it is to set up a blog of their own.



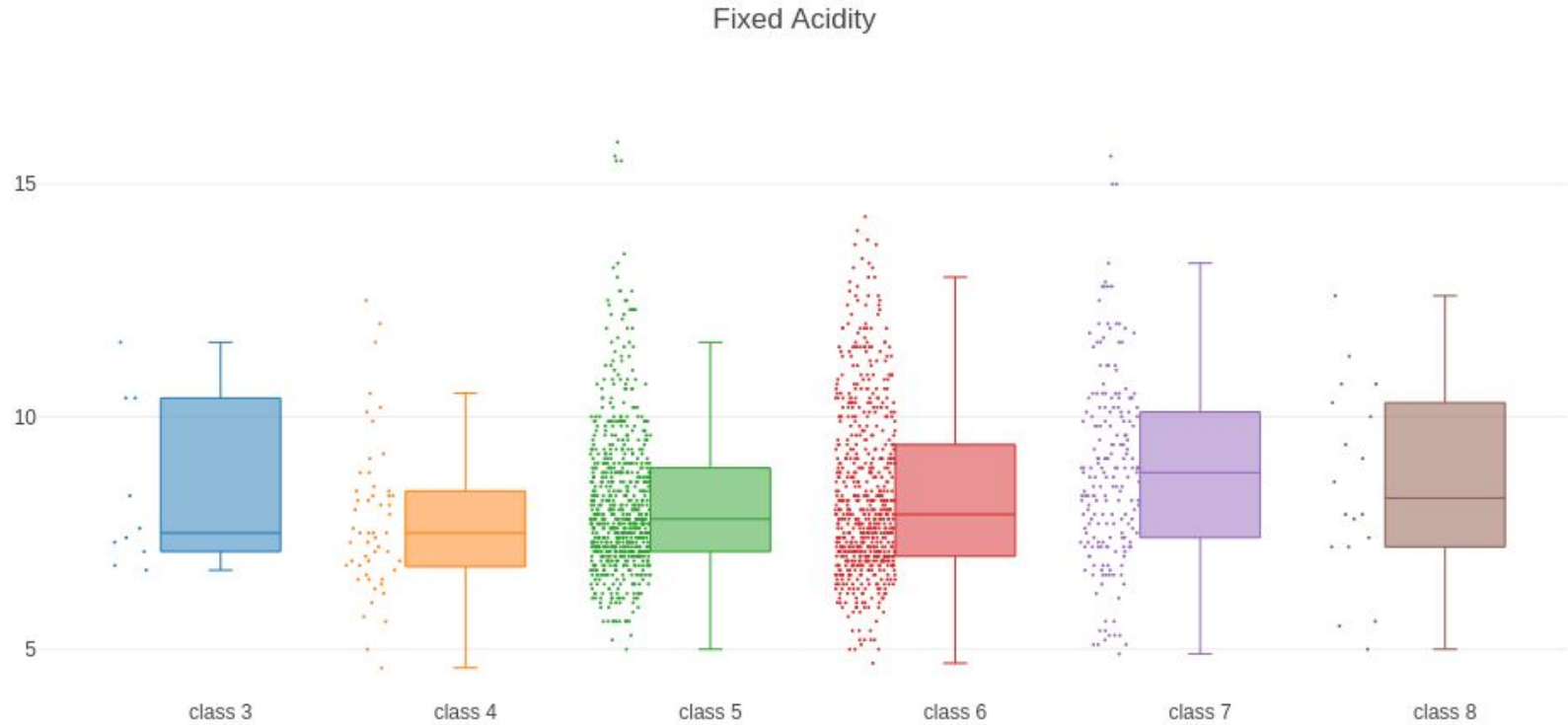
Questions ???

Backup

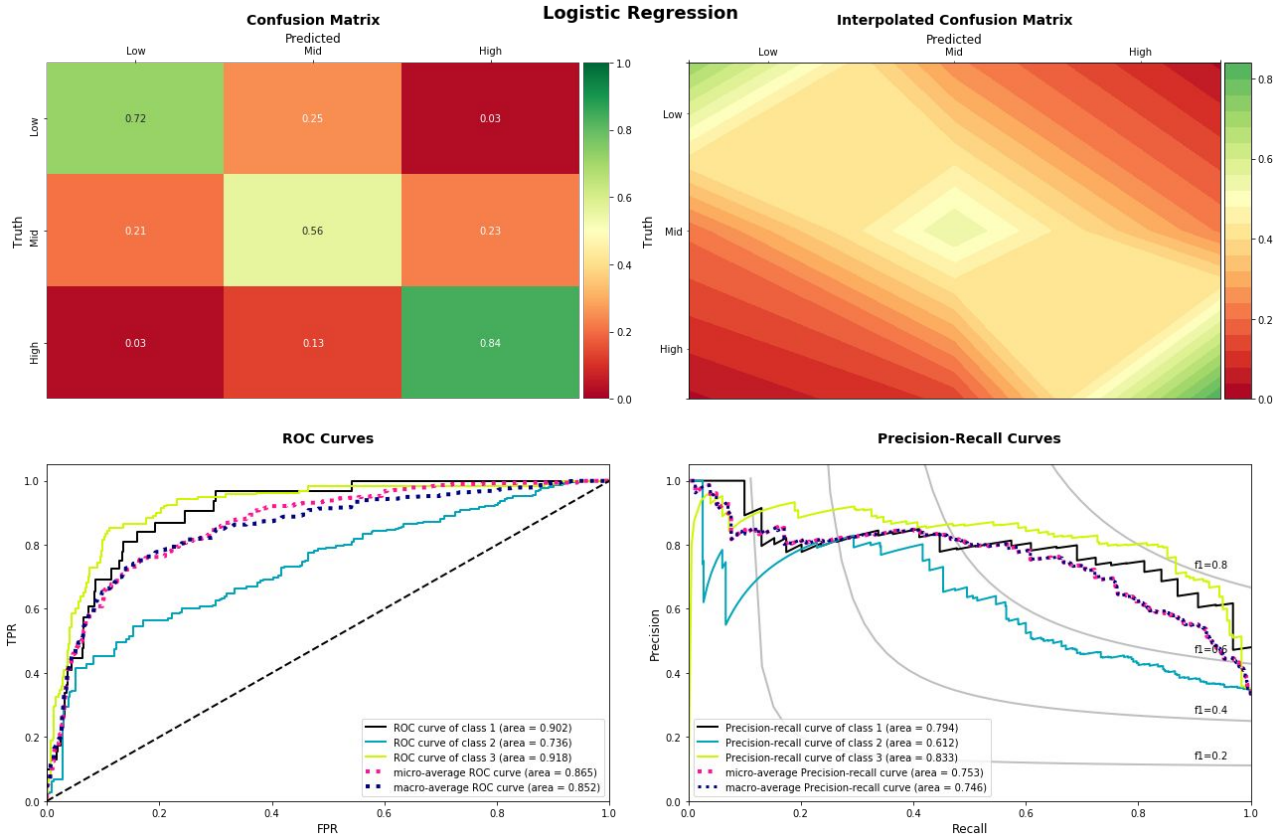
Challenges & Lessons Learned

- Traditional performance metrics are not as easily interpretable for multinomial classification (One-vs-Rest creates artificial class imbalance which skews performance metrics)
- Ordinal classes can bleed into each other (especially when target is qualitative)

Class Overlap

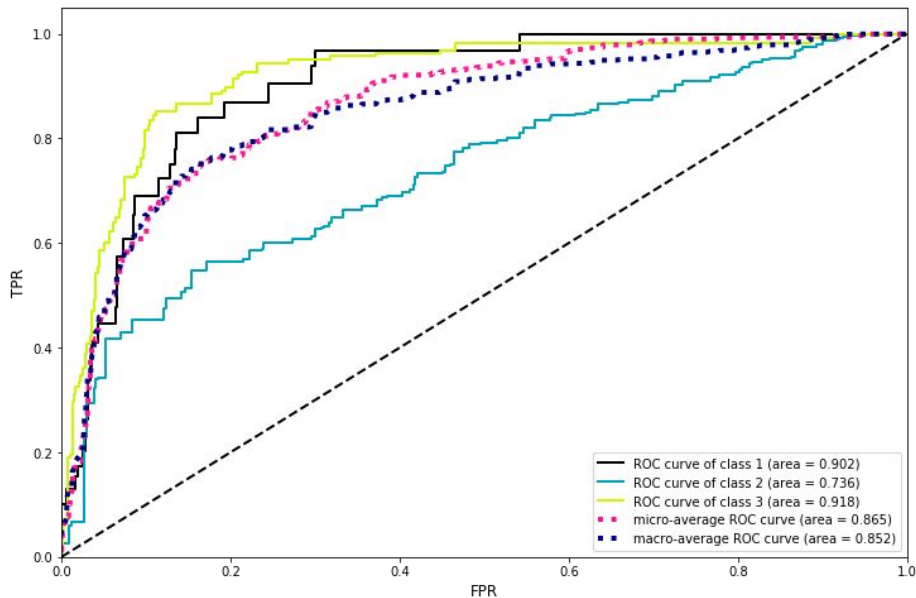


Logistic Regression

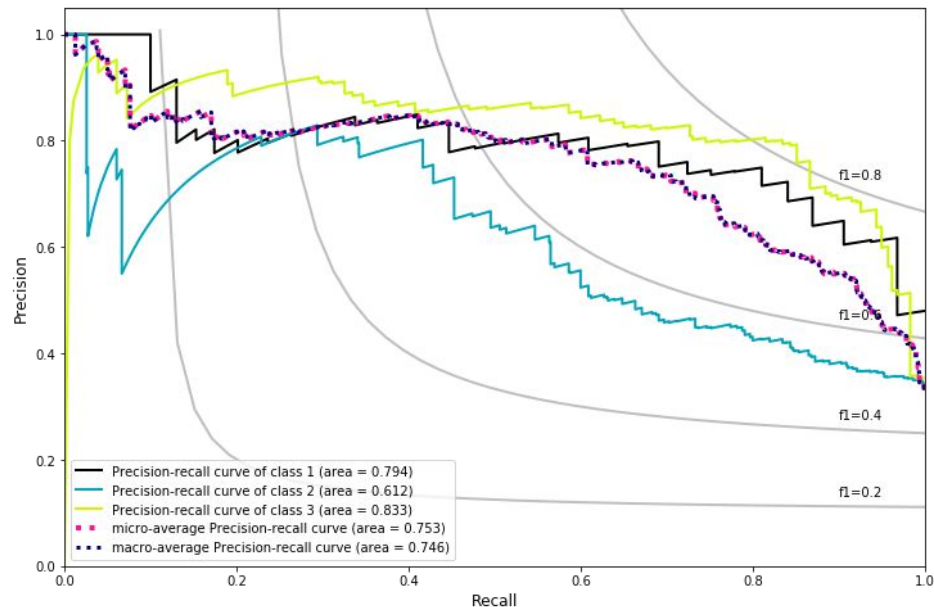


ROC & P-R Curves Complementary

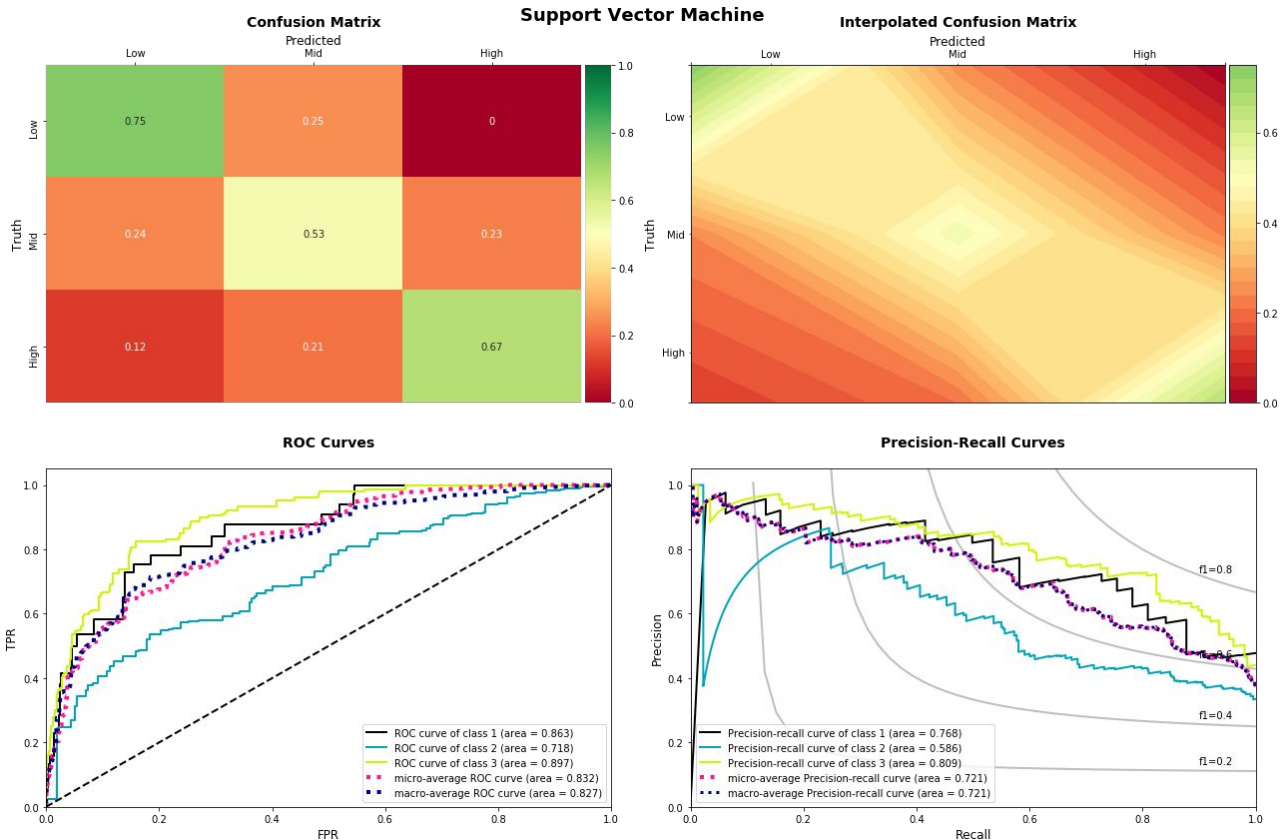
ROC Curves



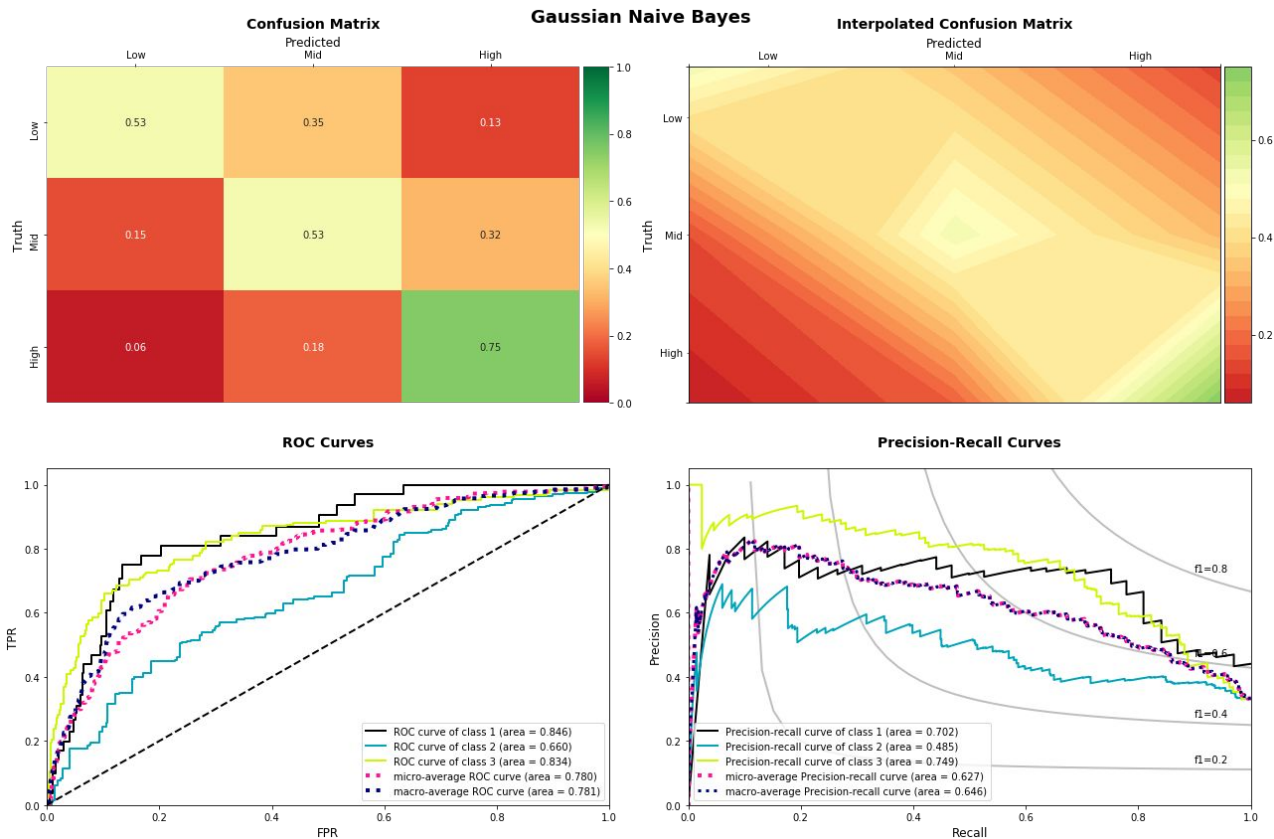
Precision-Recall Curves



Support Vector Machine



Gaussian Naive Bayes



K-Nearest Neighbors

