Self-Assessment:

- My Primary responsibility in our project is to focus on the machine learning model which includes identity the model(s), continue to refine, train, and test the model
- Other Roles I played is mostly consultant work with Teammates, it doesn't make sense to rotate role every week, So we with consultant model. Where Primary role is fixed, and other TM's will act as consultant and gives continuous feedback on Implementation.
- Initially stage of project we struggled finding right data set, as a team we figured out with some extra effort outside regular class hours
- We did peer review and pair programming to solve dependency on other module (e.g., DB & ML dependency)
- Pre-defining Data schemas and communication protocol helped us resolving dependencies

Team Assessment:

- Strength of our team is flexibility.
- Everyone is flexible to adjust their schedules if needed and accept adhoc requests
- Our Primary communication channel is Slack for messages and Zoom for Group work
- Everyone listens to others and experiment new learnings

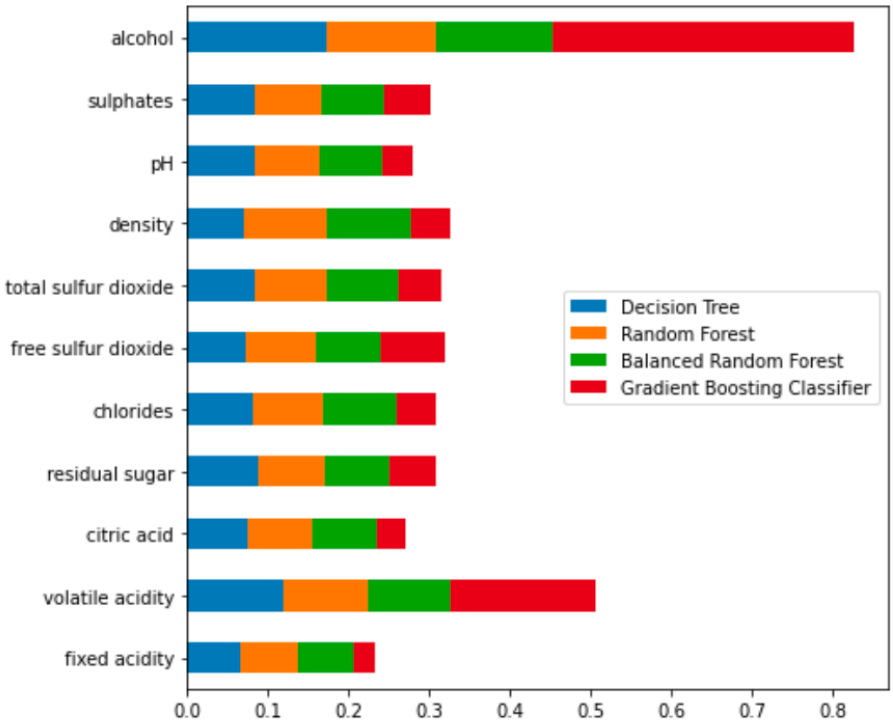
Summary of Project:

- Our project focuses on wine quality and how the 11 attributes (Chemical components) contribute to final quality score (1-10) and quality grade (Low, Medium, High)
- We used Classification Algorithm to predict the quality of the wine based on their features
- Models explored:
    - Decision Tree Classifier
    - Random Forest Classifier
    - Balanced Random Forest
    - Gradient Boosting Classifier
- Results based on our Analysis:
    - We can predict Quality of Wine based on its chemical composition
    - Alcohol content and Sulfur Dioxide plays key role in Wine Quality
    - Random Forest Model fits best to predict the quality of the wine
    - Bonus: We extend our model to predict the color of wine

Feature Comparison:

	low	medium	high
fixed acidity	7.329992	7.177257	7.085709
volatile acidity	0.397410	0.313863	0.289170
citric acid	0.304178	0.323583	0.334628
residual sugar	5.646225	5.549753	4.827721
chlorides	0.064404	0.054157	0.044576
free sulfur dioxide	29.480495	31.165021	31.055208
total sulfur dioxide	119.277055	115.410790	109.891151
density	0.995756	0.994558	0.993027
pH	3.214522	3.217726	3.227651
sulphates	0.524270	0.532549	0.541488
alcohol	9.873544	10.587553	11.433359
quality	4.884228	6.000000	7.158966

Feature Importance by Model:



Model Score:

	Accuracy	Precision	Recall	F1-Score	Support
Decision Tree	0.643077	0.643077	0.643077	0.643077	0.643077
Random Forest	0.732923	0.732923	0.732923	0.732923	0.732923
Balanced Random Forest	0.676923	0.676923	0.676923	0.676923	0.676923
Gradient Boosting Classifier	0.622154	0.622154	0.622154	0.622154	0.622154

Random Forest Classification:

	precision	recall	f1-score	support
high	0.78	0.59	0.67	320
low	0.76	0.75	0.76	596
medium	0.68	0.76	0.72	709
accuracy			0.72	1625
macro avg	0.74	0.70	0.71	1625
weighted avg	0.73	0.72	0.72	1625