

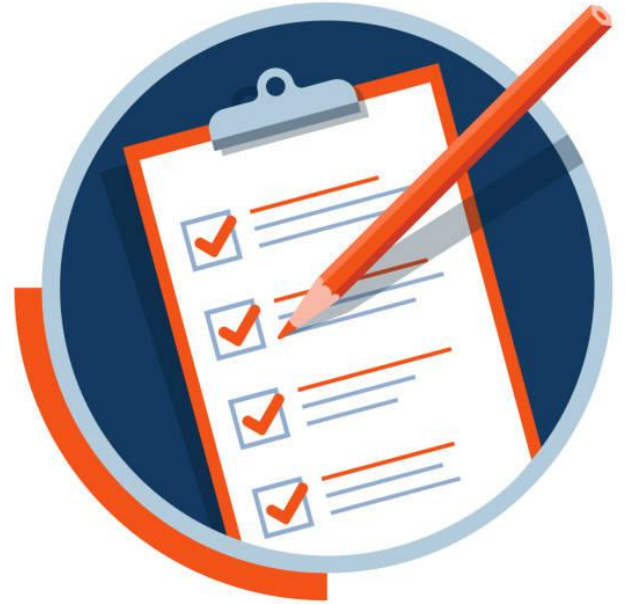
What Composition make good quality wine

Karan Vibhute
CODE Analytics
08.11.2024



Agenda

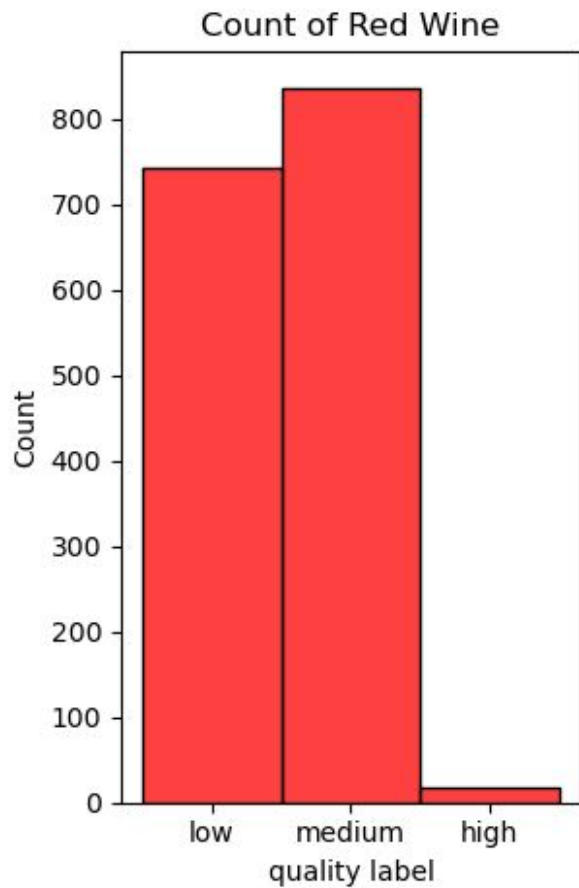
- ❑ Introduction to Data
- ❑ Red Wine data Analysis and Finding
- ❑ White Wine data Analysis and Finding
- ❑ Machine learning model for Red wine
- ❑ Machine learning model for White wine
- ❑ Final Summary



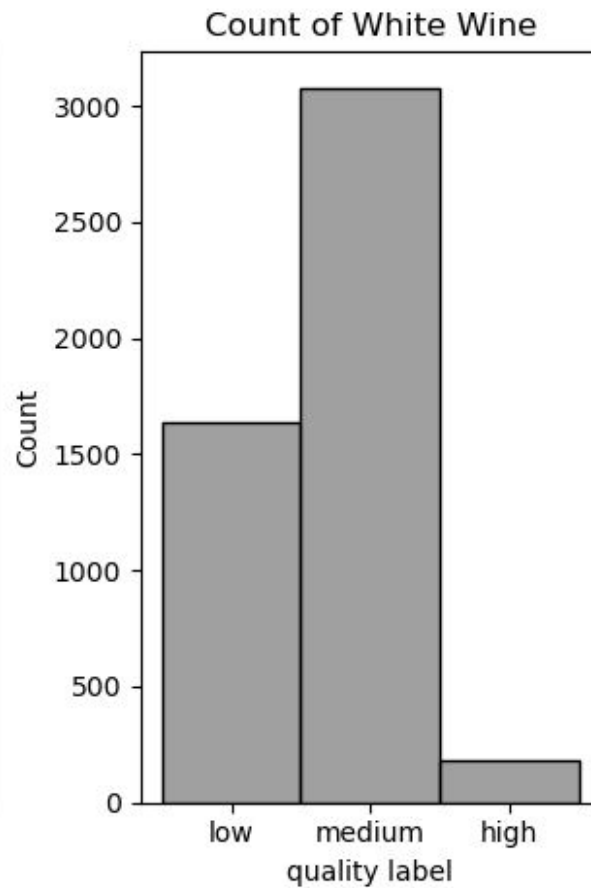


Variables in Data

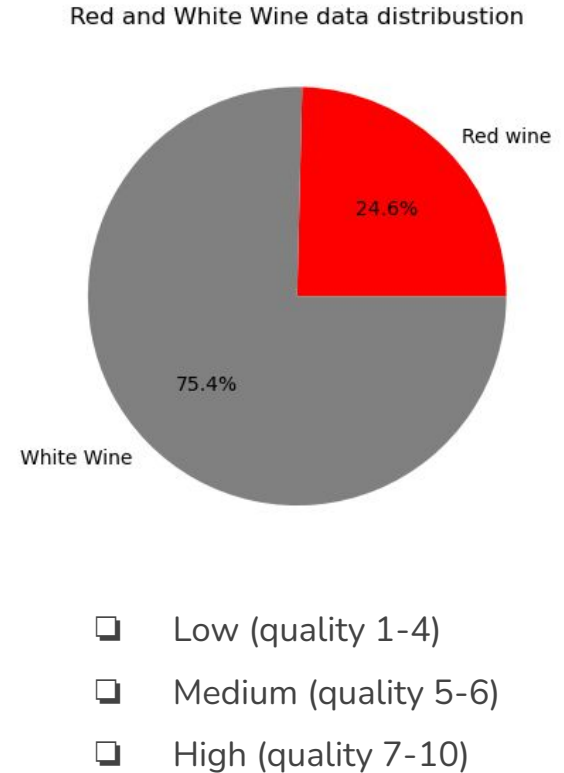
- ❑ Fixed acidity
 - ❑ Volatile acidity
 - ❑ Citric acid
- ❑ Residual Sugar
- ❑ Chlorides
- ❑ Free Sulfur dioxide
- ❑ Total sulfur dioxide
- ❑ Density
- ❑ pH
- ❑ Sulphates
- ❑ Quality
 - ❑ Low (1-4)
 - ❑ Medium (5-6)
 - ❑ High (7-10)



Total Count = **1599**

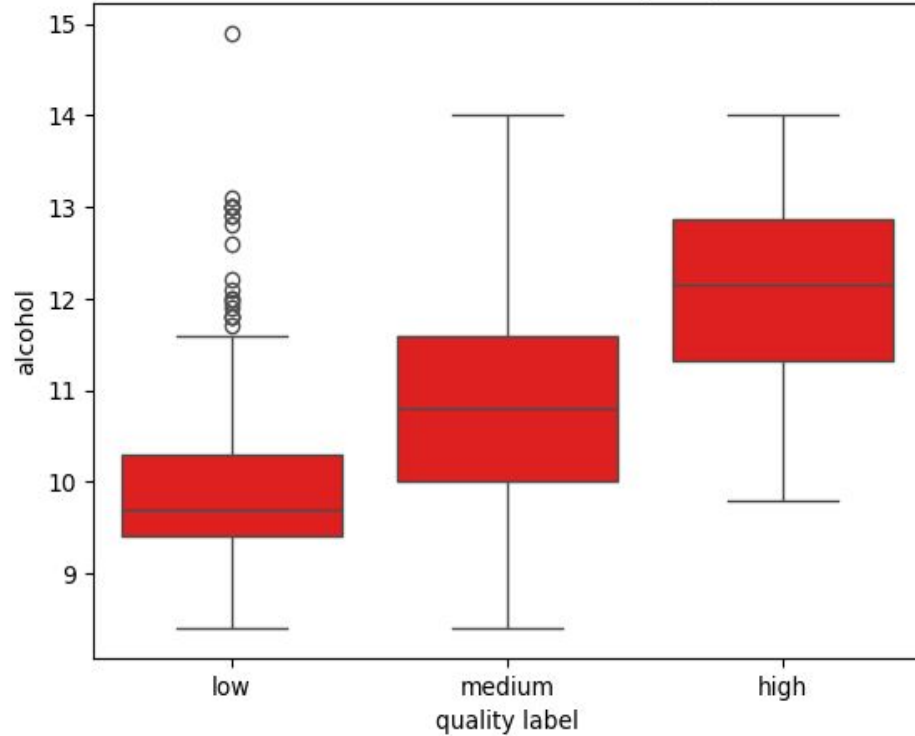


Total Count = 4898

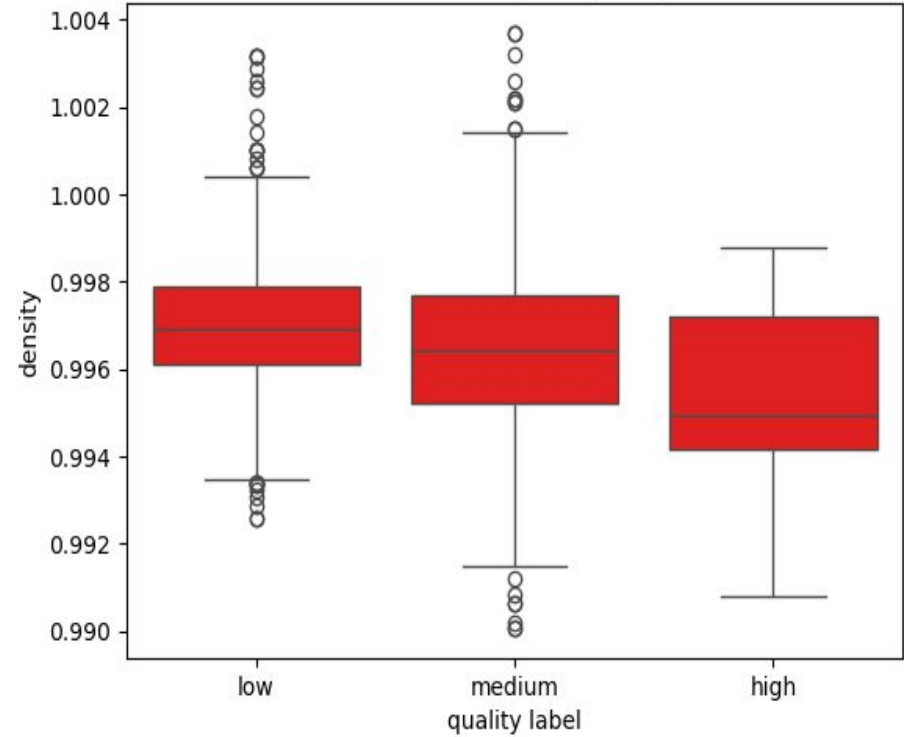


Red Wine data Analysis and Finding

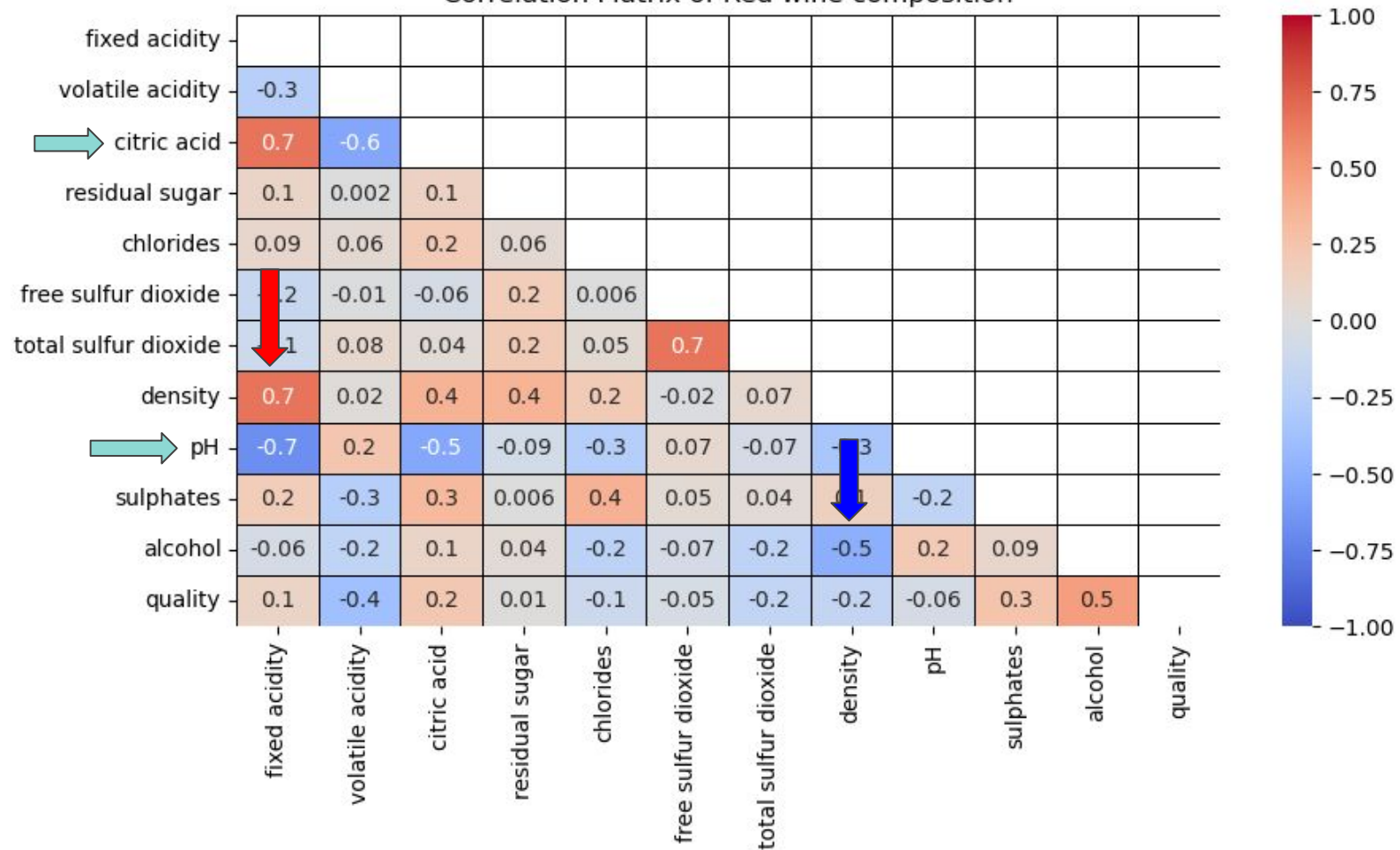
"Red Wine" Alcohol contain Vs quality Label



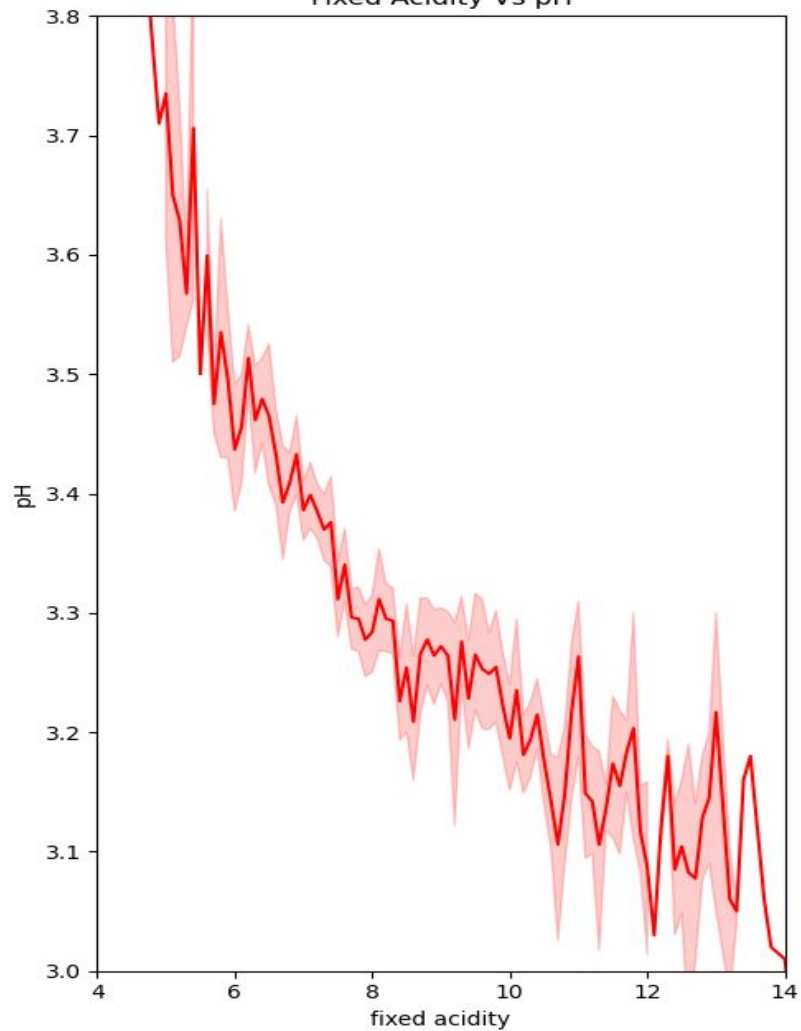
"Red Wine" density Vs quality Label



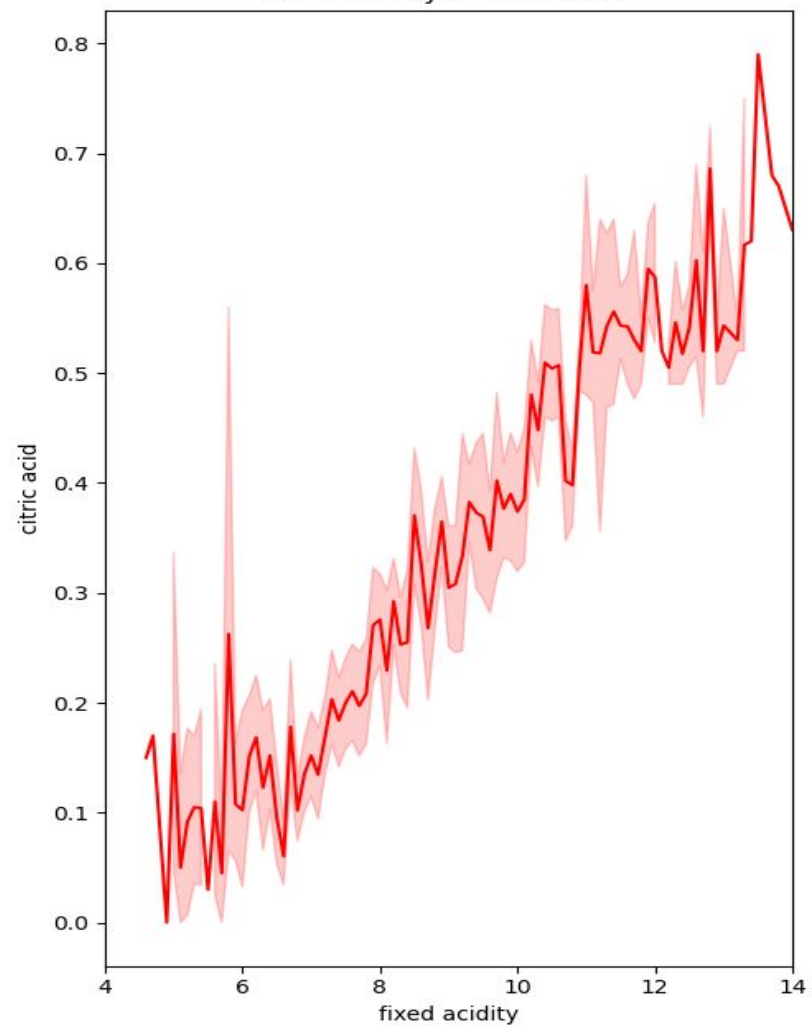
Correlation Matrix of Red wine composition



Fixed Acidity Vs pH



Fixed Acidity Vs Citric Acid



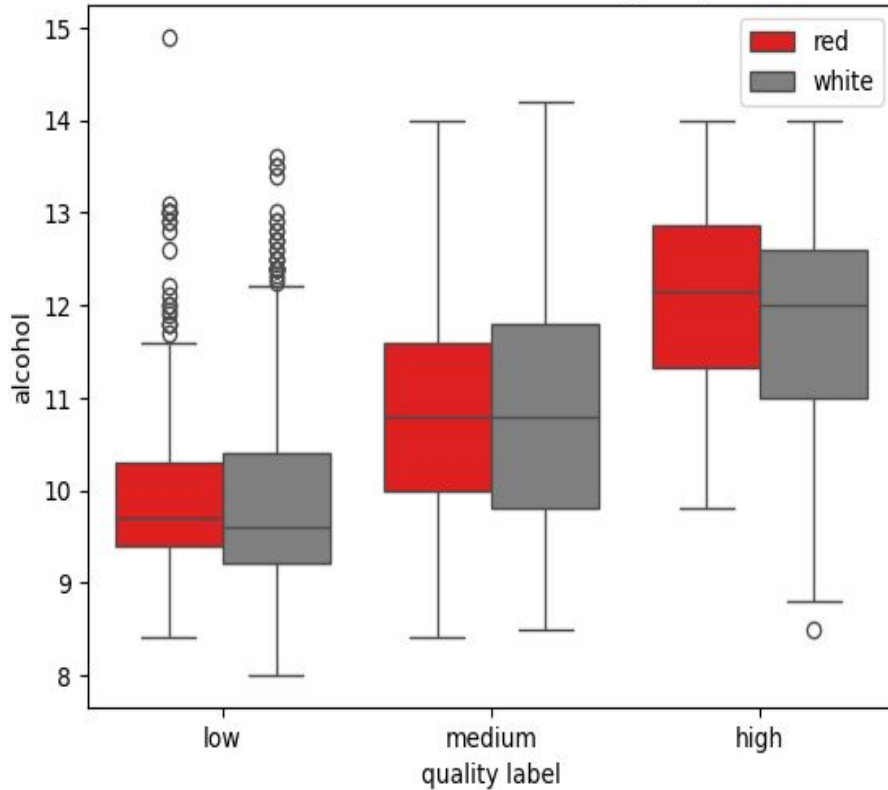


Findings for Red Wine

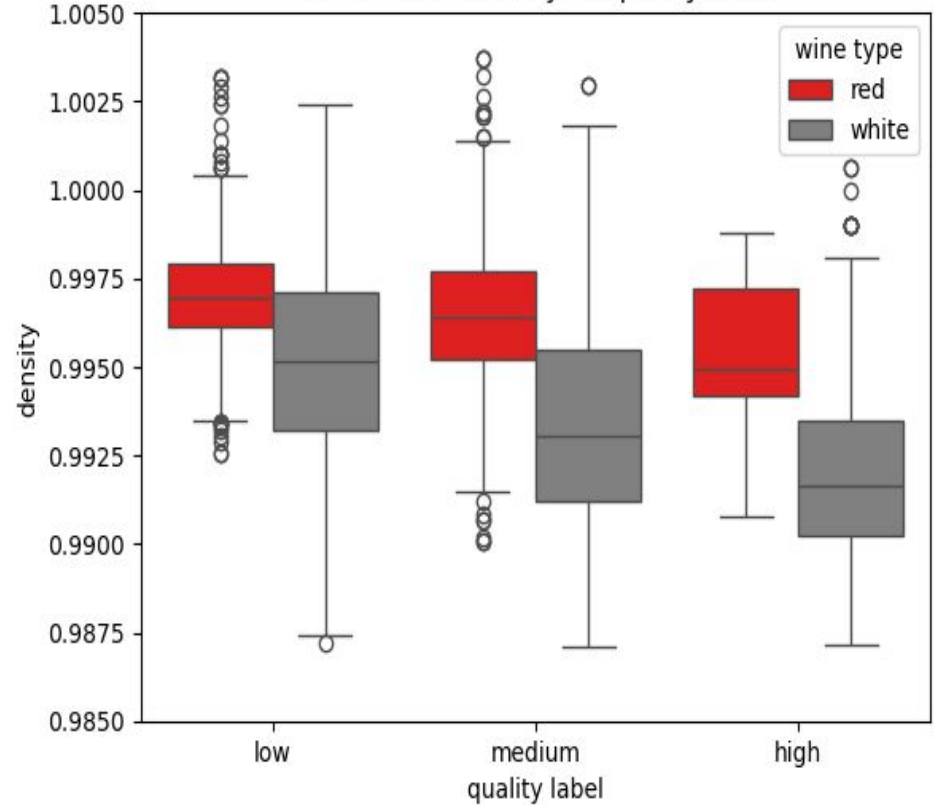
- ❑ As we Lower citric acid or higher the pH, it lower the fixed acidity and density of red wine
- ❑ If density is Low it increases Percentage of alcohol in Red wine
- ❑ Higher percentage of alcohol increases Quality of wine
- ❑ So if comparatively **Lower citric acid** or **higher the pH**, the Quality of Red wine will increases

White Wine data Analysis and Finding

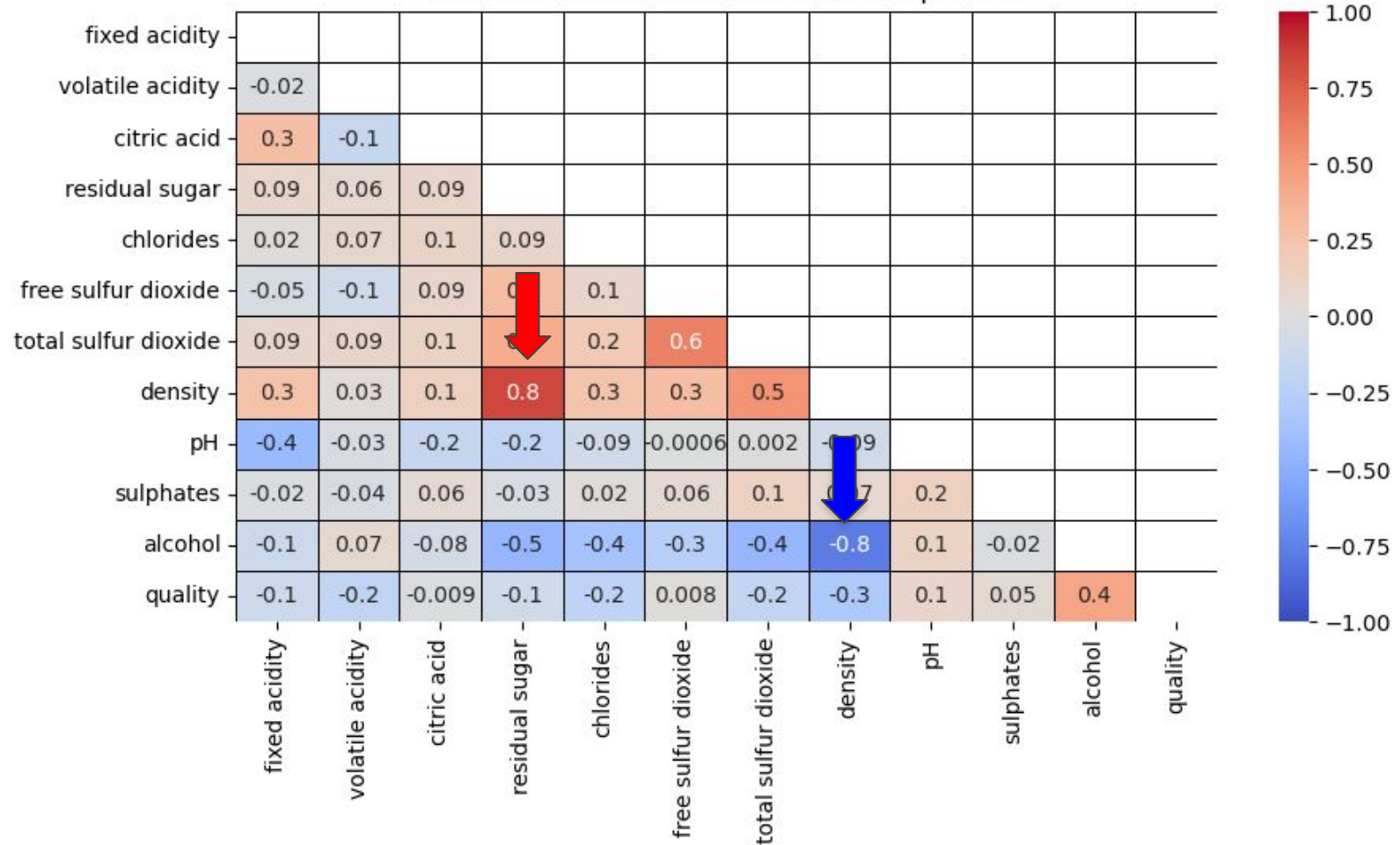
"Both Wine" alcohol Content Vs quality Label

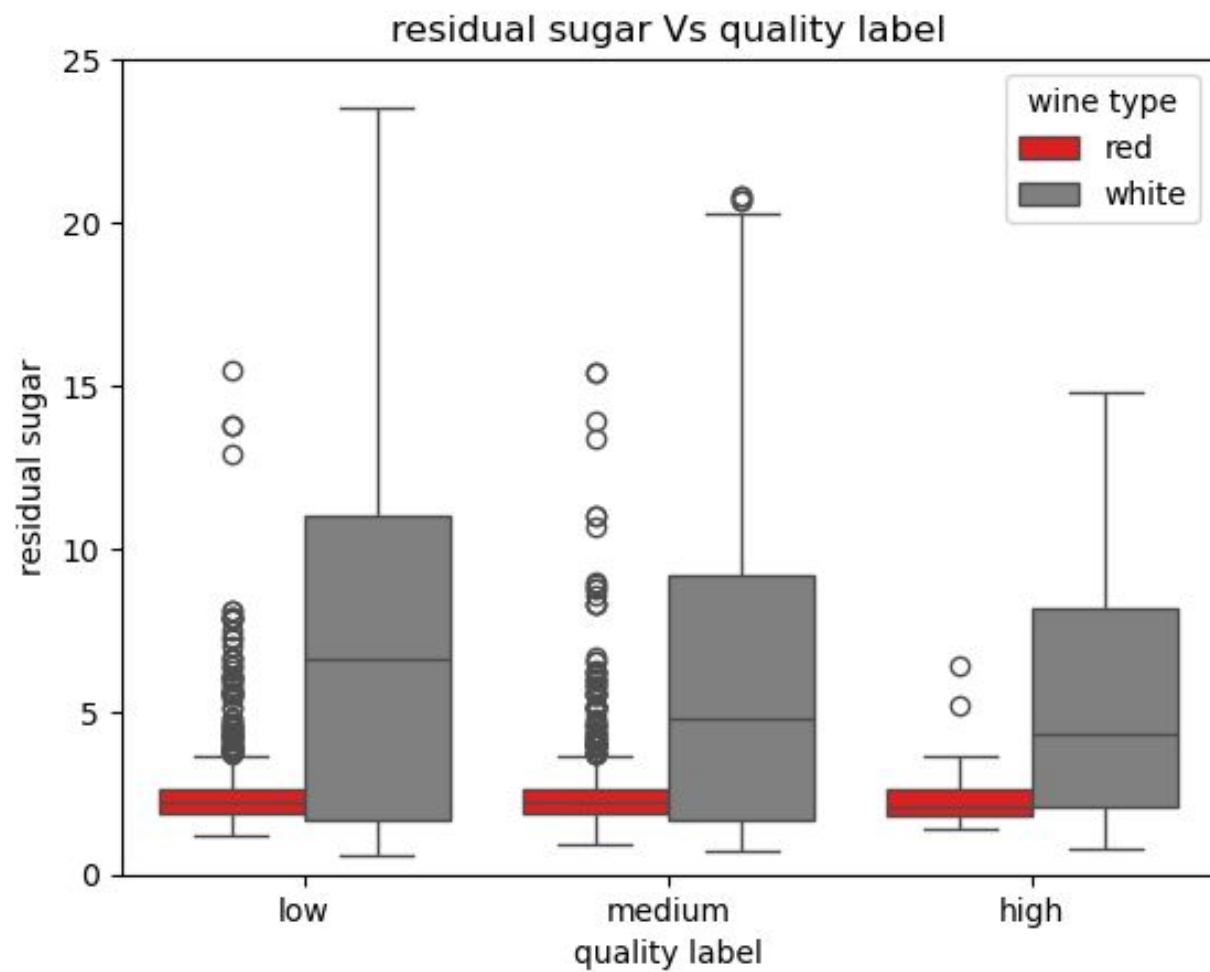


"Both wine" density Vs quality label



Correlation Matrix of White wine composition







Findings for White Wine

- ❑ As we Lower the residual Sugar, it lower the density of White wine
- ❑ If density is Low it increases Percentage of alcohol in White wine
- ❑ Higher percentage of alcohol increases Quality of white wine
- ❑ So if comparatively **lower the Residual Sugar**, the Quality of White wine will increases



Summary

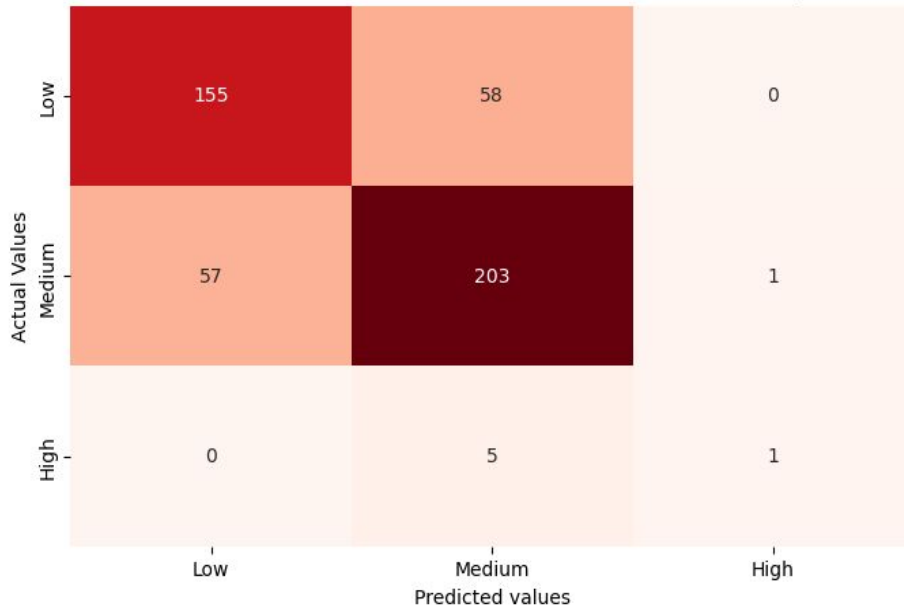
- ❑ To get good quality of **Red Wine** we can comparatively **decrease citric acid** or **increase the pH**
- ❑ To get good quality of **White wine** we can comparatively **lower the Residual Sugar**

Machine learning model for **Red Wine**

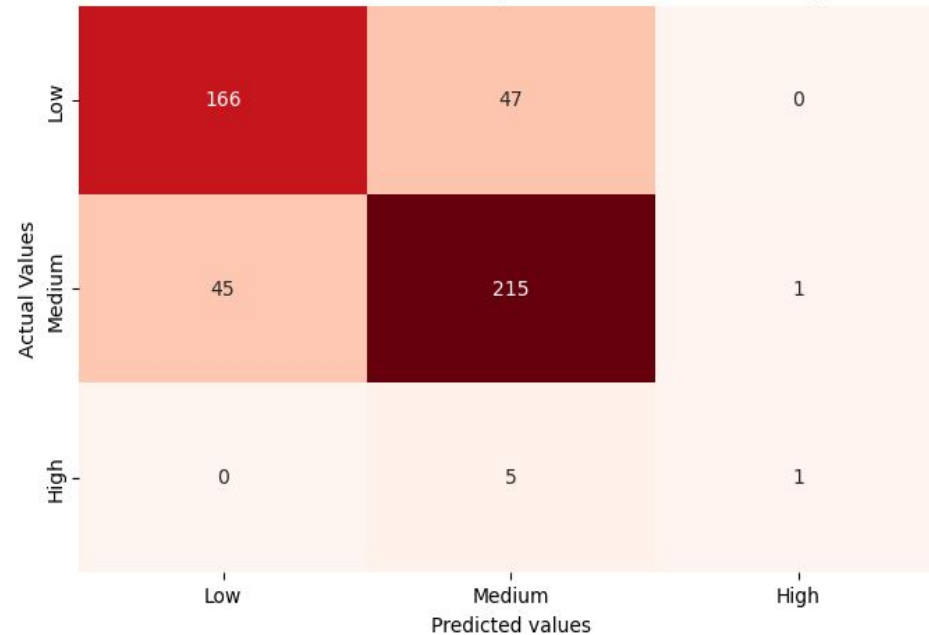


- ❑ Applied Models : Logistic Regression and Random Forest
- ❑ Chosen Model : Random Forest (It shows Better Result)

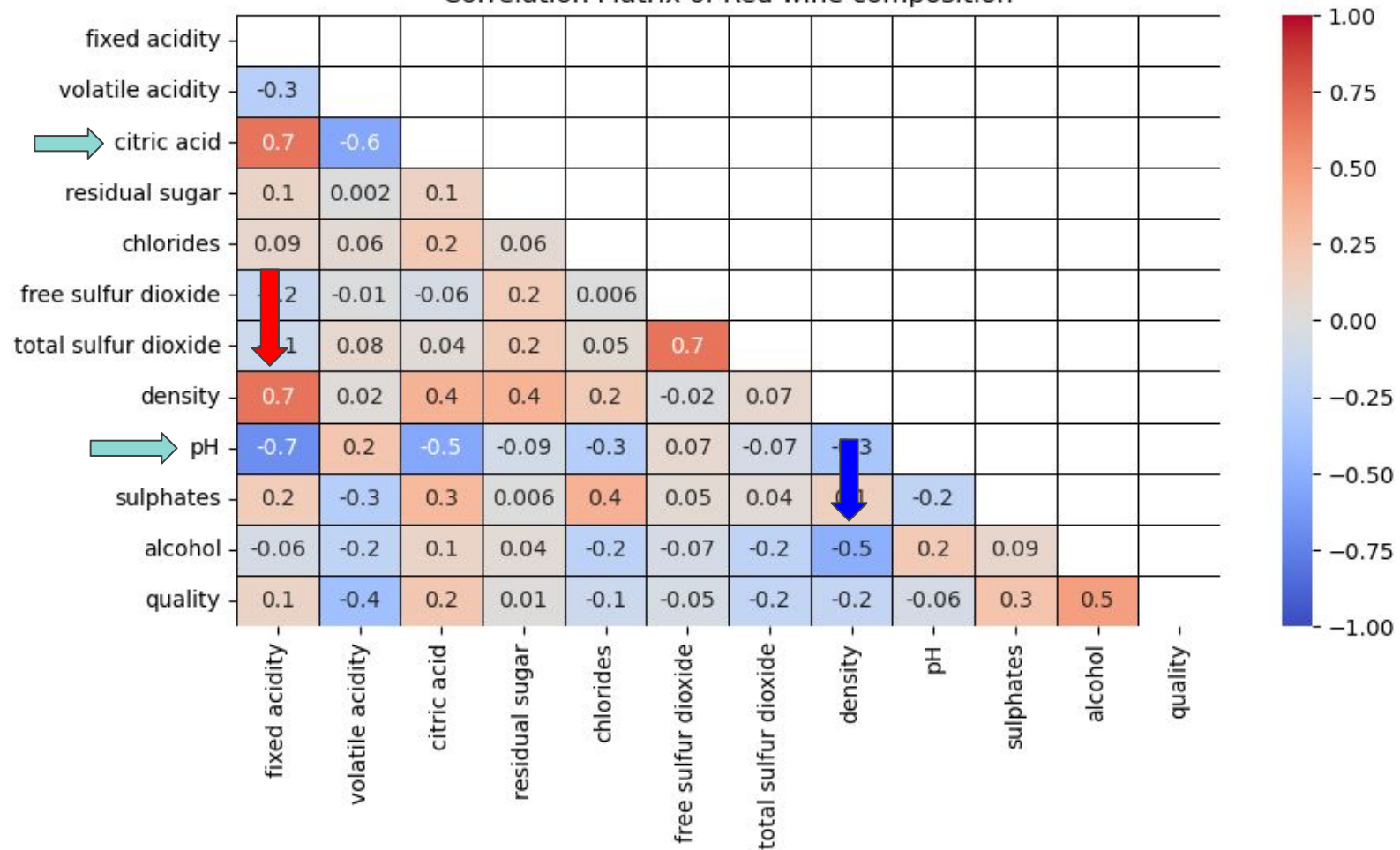
Red wine confusion matrix With Resulted Features - Accuracy 75 %



Red wine confusion matrix With Dependent Features - Accuracy 80.20 %



Correlation Matrix of Red wine composition

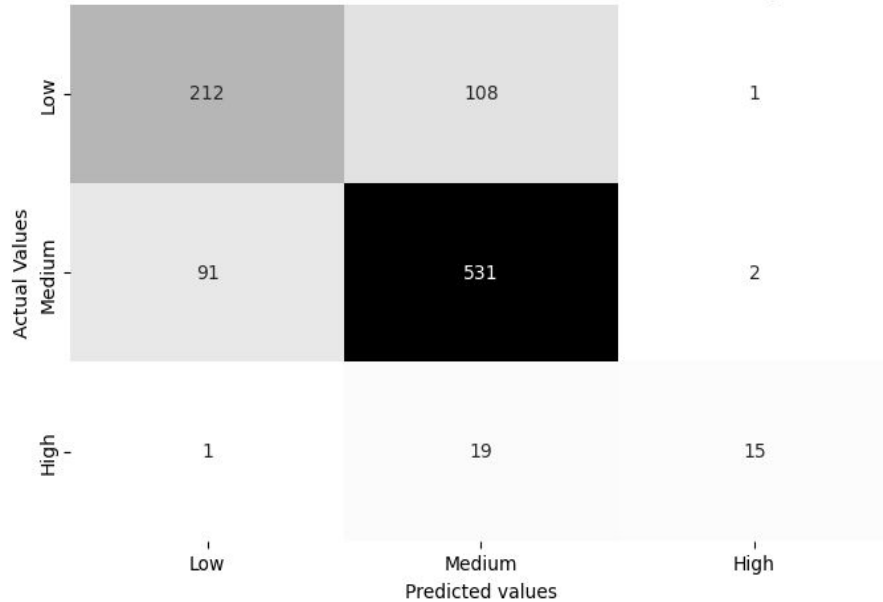


Machine learning model for White Wine

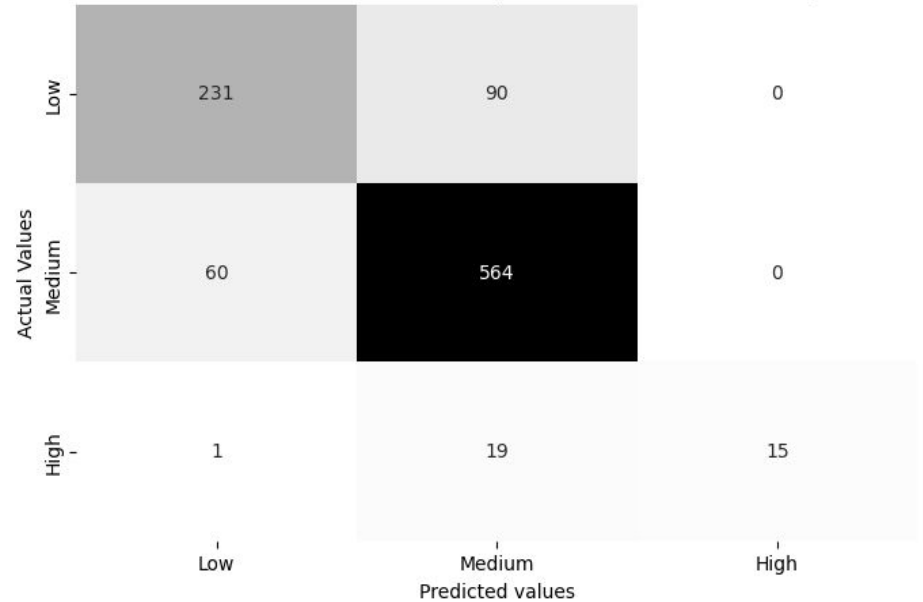


- ❑ Applied Models : Logistic Regression and Random Forest
- ❑ Chosen Model : Random Forest (It shows Better Result)

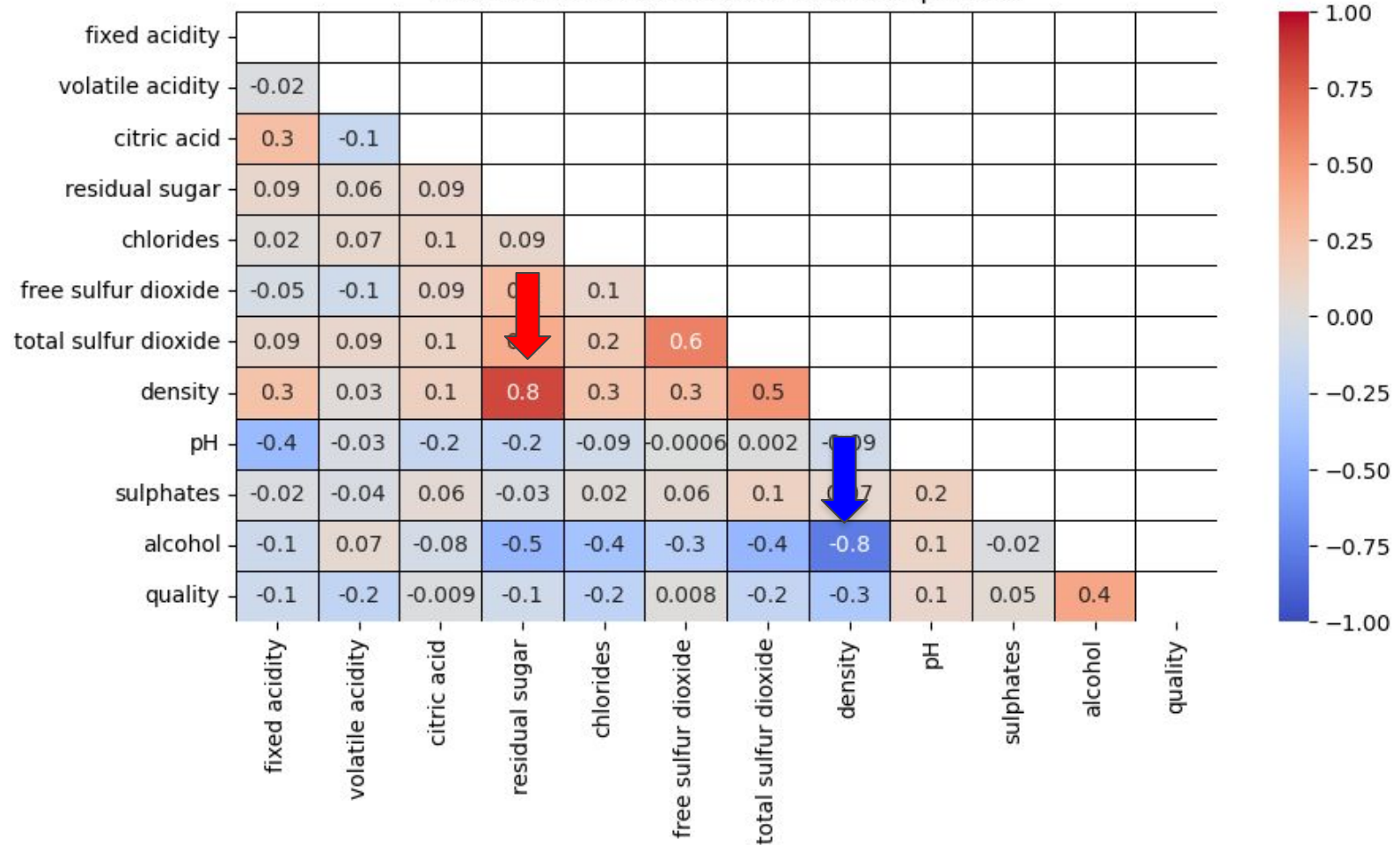
White wine confusion matrix With Resulted Features - Accuracy 76.93 %



White wine confusion matrix With Dependent Features - Accuracy 82.44 %



Correlation Matrix of White wine composition





Hyperparameter and Normalization

Data Scaling	Red Wine		White Wine	
	Before	After	Before	After
Hyperparameter	80.20	78.12	82.44	82.55
Normalization	80.20	79.58	82.44	82.65

Table ; Accuracy change in % after Data Scaling



Final Summary

- ❑ For Red wine all composition important
- ❑ For White wine all composition important only Sulphates are low correlated directly or indirectly
- ❑ Random forest is best machine learning model for this data set
- ❑ Due to less data for High type of both wines Accuracy affected
- ❑ Further scope is to check data in details where Wine quality is Higher but predicted lower



THANK YOU

Any Question ?