

Day 53

DIY

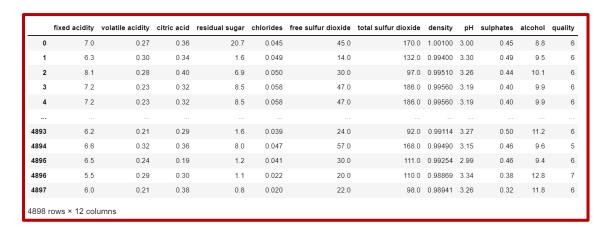
Q1. Problem Statement: Model Evaluation Metrics for Regression

Write a Python program that reads the *winequality-white.csv* (provided on LMS) file into a DataFrame, the following are the tasks that are to be taken into consideration while constructing diffract model and in the end evaluate them based on RMSE, MAPE, RMSLE.

- 1. Load the given dataset into a data frame
- 2. Find missing values and drop them if you find any
- 3. Check data types for all features
- 4. Extract dependent and independent variables into the y & x data frame ("alcohol" is our dependent feature)
- 5. Split your data into train and test, by 20% as test size
- 6. Create a new data frame for comparison of all models containing column as model name, RMSE, MAPE, RMSLE
- 7. Build linear regression, SVM, ridge, lasso, Decision Tree and measure their RMSE, MAPE, RMSLE and make the final data frame

Dataset:





Sample Output:

7. Build linear regression, SVM, ridge, lasso, Decision Tree and measure their RMSE, MAPE, RMSLE and make the final data frame

	Model Name	MAPE	RMSE	RMSLE
0	Linear regression	2.947415	0.398478	0.035545
1	SVM regression	7.310045	1.010017	0.085515
2	Ridge regression	2.952137	0.398849	0.035566
3	Lasso regression	3.381967	0.444636	0.039187
4	Decision Tree regression	4.154571	0.572211	0.049829