
# Machine Learning Project Synopsis

# Wine Quality Prediction

Submitted by,

Reevan Mario Miranda 4NM16CS118

Sriharivishnu 4NM16CS151

Submitted to,

Mrs. Divya Jennifer D'Souza

Assistant Prof. Gd I

Submission date:

04 September 2019

## Abstract

The aim of the project is to predict the quality of Red wine on a scale of 0–10 given a set of features as inputs. Linear Regression method of Supervised learning is used to predict the quality.

Wine quality dataset from UCI Machine learning repository is used:
https://archive.ics.uci.edu/ml/datasets/Wine+Quality

Input variables are: fixed acidity, volatile acidity, citric acid, residual sugar, chlorides, free sulfur dioxide, total sulfur dioxide, density, pH, sulphates, alcohol. And the output variable is Quality (score between 0 and 10). Below is a screenshot of the top 5 rows of the dataset.

| | fixed acidity | volatile acidity | citric acid | residual sugar | chlorides | free sulfur dioxide | total sulfur dioxide | density | pH | sulphates | alcohol | quality |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 7.4 | 0.70 | 0.00 | 1.9 | 0.076 | 11.0 | 34.0 | 0.9978 | 3.51 | 0.56 | 9.4 | 5 |
| **1** | 7.8 | 0.88 | 0.00 | 2.6 | 0.098 | 25.0 | 67.0 | 0.9968 | 3.20 | 0.68 | 9.8 | 5 |
| **2** | 7.8 | 0.76 | 0.04 | 2.3 | 0.092 | 15.0 | 54.0 | 0.9970 | 3.26 | 0.65 | 9.8 | 5 |
| **3** | 11.2 | 0.28 | 0.56 | 1.9 | 0.075 | 17.0 | 60.0 | 0.9980 | 3.16 | 0.58 | 9.8 | 6 |
| **4** | 7.4 | 0.70 | 0.00 | 1.9 | 0.076 | 11.0 | 34.0 | 0.9978 | 3.51 | 0.56 | 9.4 | 5 |

Plotting the heatmap of correlation between features: