DATA SCIENCE AND ANALYTICAL LAB PROJECT.

BY

18SW142 & 18SW68

PREDICTION OF WINE QUALITY.

ABSTRACT

The better quality of wine is quite important for the good health of the addicted people. Today, the prediction of its quality is the major focus of many researchers and health scientists in order to provide an assumption of how the wine affects the human body organs, and averting the chances of lung cancer is the major objective of this research study. The whole process of predicting the wine quality was carried out by importing the python libraries i.e. NUMPY, PANDAS, MATPLOTLIB and SEABORN, whereas SKLEARN is used for training and splitting the data values. The Random Forest Classifier was used for finding and computing the results of accuracy score; the whole data values of the dataset taken from KAGGLE are fetched from a CSV/EXCEL file. Besides it, the process of data analysis was performed to analyze the shape and missing values in the dataset along with the data visualization for comparing different columns and its results. The 12 columns were taken as labels aimed at quality prediction of wine which was involved in data visualization where the assumption of bad and good quality of wine was observed with the help of sketching the correlation method for columns. The experimental model training was performed by dint of Random Forest algorithm. The final concluding results and observations led to evaluation of accuracy through testing the dataset in which it was found that the model is predicting 0.94% correct results, and the complete model of the predictive system was built accordingly based on the quality analysis of wine as bad or good.

GITHUB LINK OF PROJECT:

https://github.com/LaibaMemon/Laiba-EDA-on-Wine-Quality-Prediction-data/blob/main/WINE%20QUALITY%20PREDICTION.ipynb