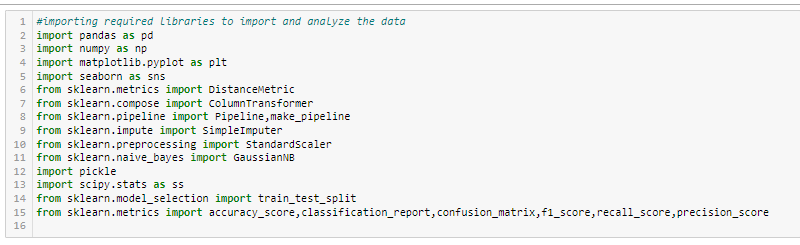
**Documentation on Apple Quality Prediction using Naïve bayes**

**Dataset** : <https://www.kaggle.com/datasets/nelgiriyewithana/apple-quality>

**Overview :** The dataset contains information about various attributes of a set of fruits, providing insights into their characteristics. The dataset includes details such as fruit ID, size, weight, sweetness, crunchiness, juiciness, ripeness, acidity, and quality.

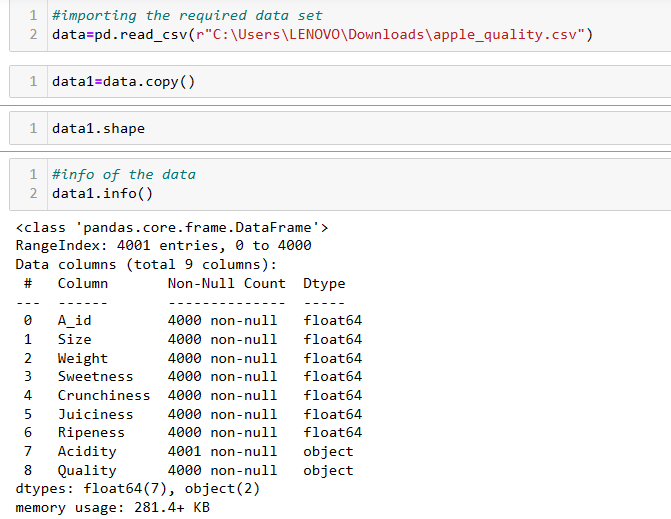
**Problem Statement :** Fit a model to predict the quality of the fruit using Naïve Bayes .

1. **Import Libraries**

****

**Data Collection**

1. Importing the dataset and preview of the data :



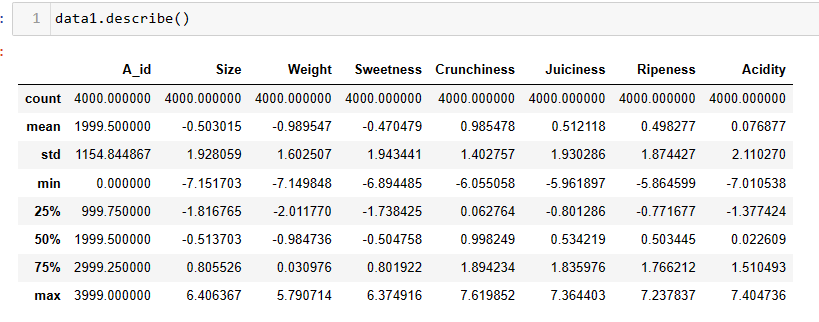
**Data Cleaning / Data manipulation**

1. Data Cleaning :

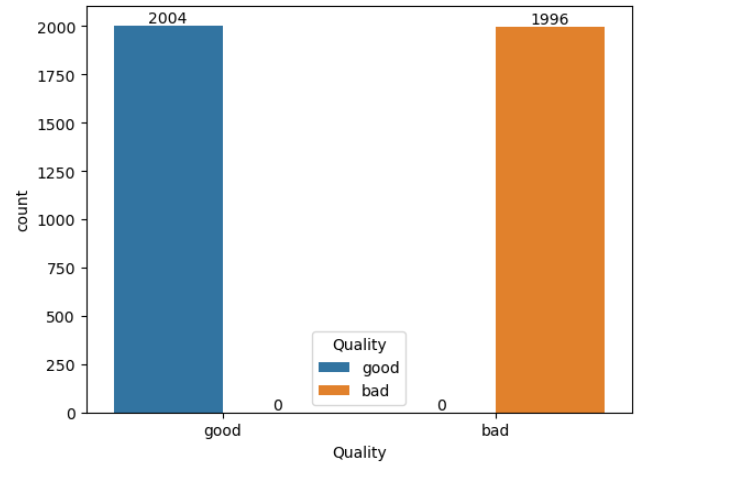
* Found 1 null value in each column
* Dropped the null values present in the dataset
* Changed the data type of column “Acidity” to float
* After cleaning dataset is of shape (4000,9)

**Exploratory Data Analysis**

1. Statistical sights :



1. Exploratory Data Analysis :



Based on the count plot of class variable, the dataset is balanced.

**Feature Engineering**

1. Feature engineering :

* Extracted required feature variable and class variable from the dataset
* Checked if all the feature variables are following Gaussian Distribution
* All the feature variable are following Gaussian Distribution.

**Trainig the model**

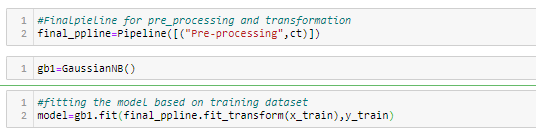
1. Dividing the dataset to train and test :



**Fitting pipelines**

****

1. Fitting final Gaussian Naïve Bayes model using training data

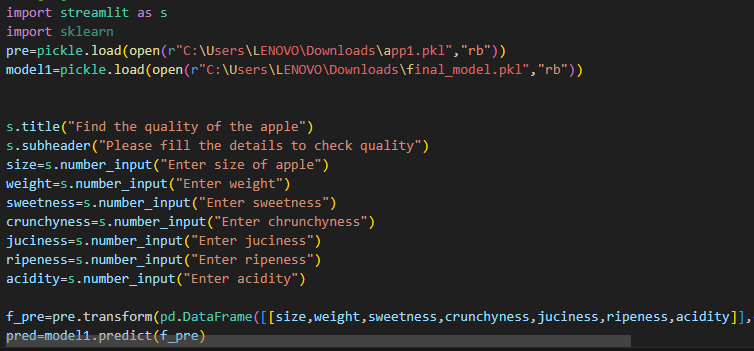


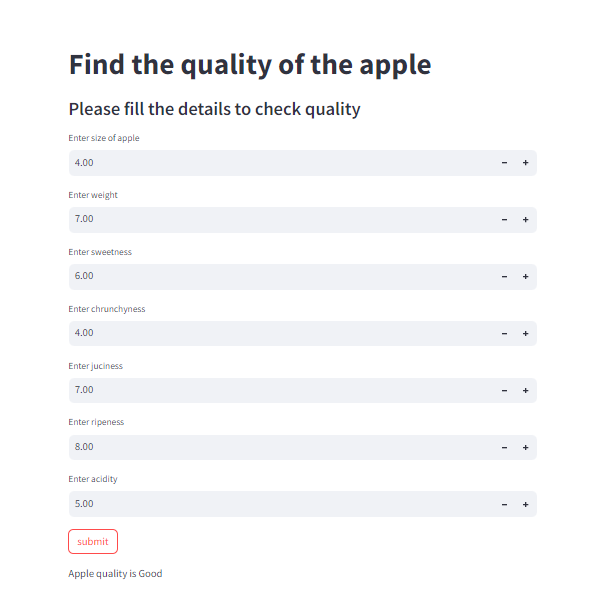
1. Deployment :

Dumped the model and loaded in VSCode to run :



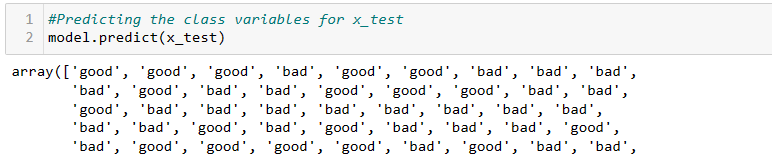
Execution of program in streamlit





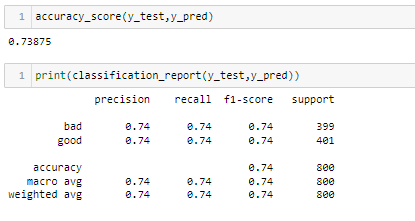
**Testing the model**

1. Predicting the Class variable for test data using the model



**Performance of the model**

1. Performance of the model :



The accuracy\_score of the model is 75.75% which concludes that the performance is good.