**Experiment No 8**

**Title:** To Implement Naïve Bayes for given problem

**Tools Required: Anaconda Navigator**

**Concept :**

Naive Bayes methods are a set of supervised learning algorithms based on applying Bayes’ theorem with the “naive” assumption of conditional independence between every pair of features given the value of the class variable. Naive Bayes is the most straightforward and fast classification algorithm, which is suitable for a large chunk of data. Naive Bayes classifier is successfully used in various applications such as spam filtering, text classification, sentiment analysis, and recommender systems. It uses Bayes theorem of probability for prediction of unknown class.

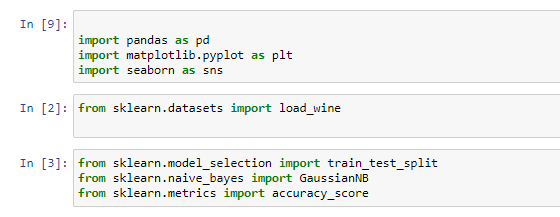
**Example Problem**

You will use wine dataset which is a very famous multi-class classification problem. "This dataset is the result of a chemical analysis of wines grown in the same region in Italy but derived from three different cultivars." (UC Irvine)

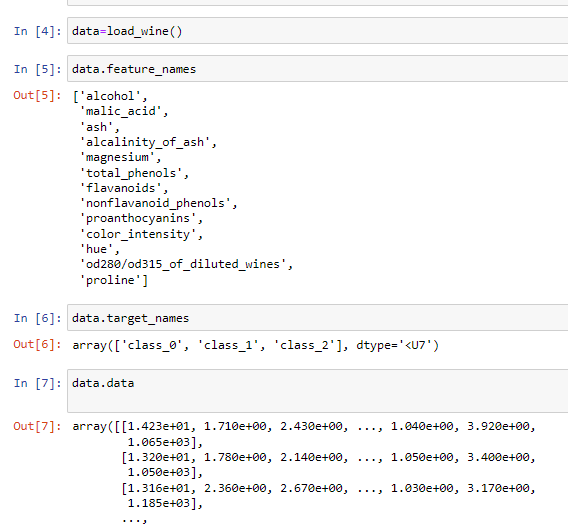
Dataset comprises of 13 features (alcohol, malic\_acid, ash, alcalinity\_of\_ash, magnesium, total\_phenols, flavanoids, nonflavanoid\_phenols, proanthocyanins, color\_intensity, hue, od280/od315\_of\_diluted\_wines, proline) and type of wine cultivar. This data has three type of wine Class\_0, Class\_1, and Class\_2. Here you can build a model to classify the type of wine. The dataset is available in the scikit-learn library.

**Steps:**

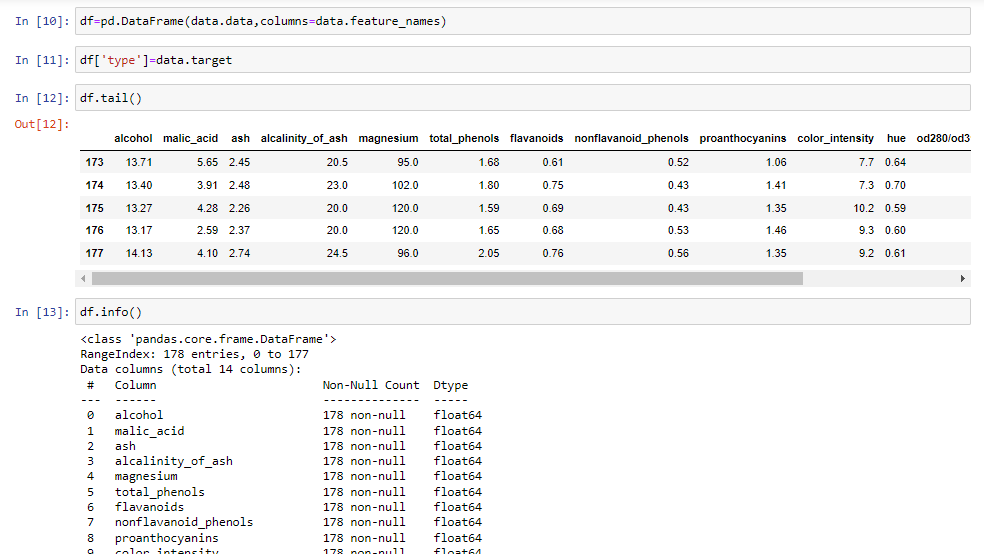
1. Import the required libraries and the dataset from scikit-learn datasets.

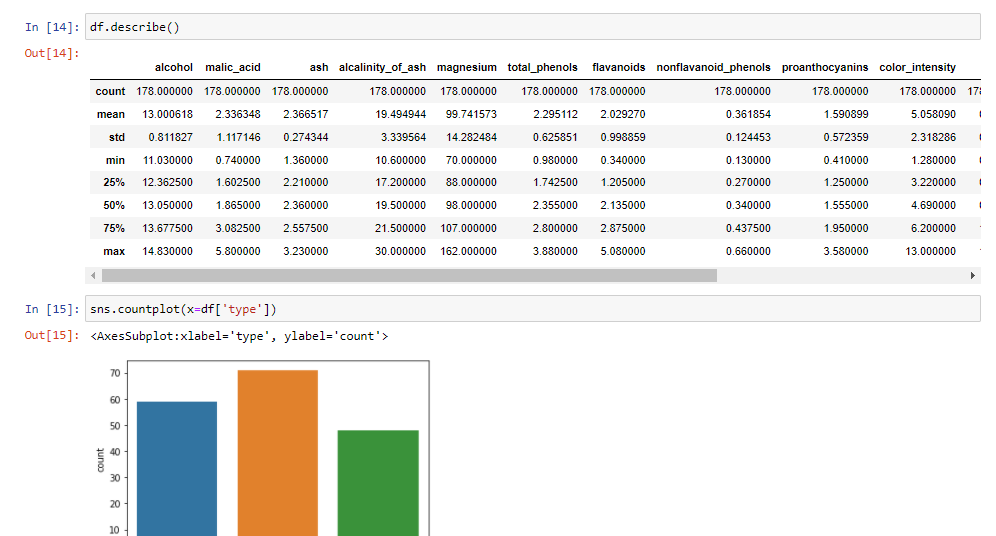


1. Exploring Data: You can print the target and feature names, to make sure you have the right dataset.



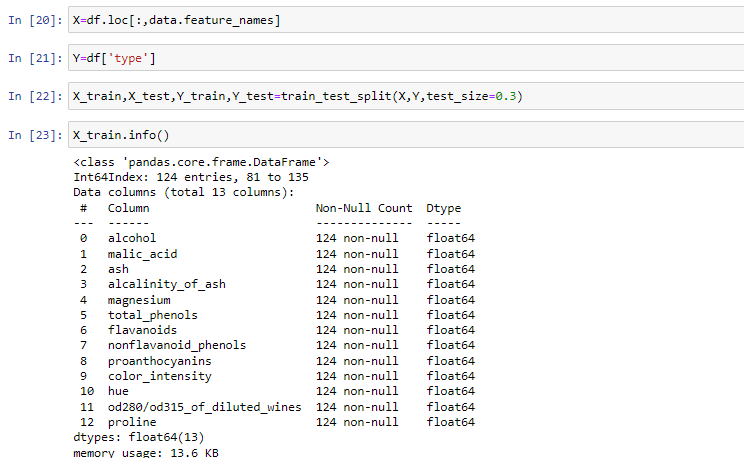
1. Data Exploration- It's a good idea to always explore your data a bit, so you know what you're working with.eg printing first five rows of whole dataset.

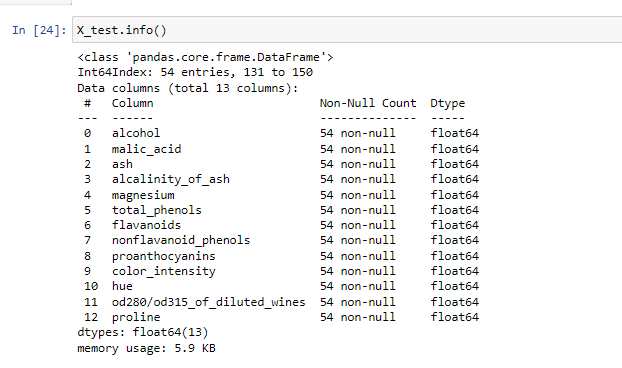




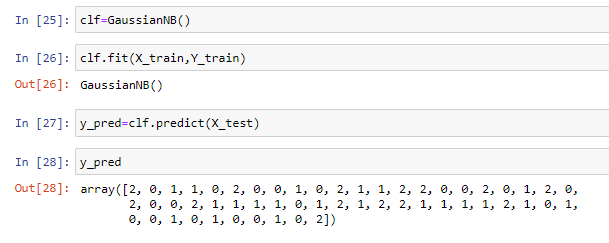


1. Splitting Data:First, you separate the columns into dependent and independent variables(or features and label). Then you split those variables into train and test set.





1. Model Generation: Performing naïve bayes analysis using scikit learn
   1. Create Gaussian classifier object
   2. Train gaussian classifier
   3. Predict the response for test dataset



1. Evaluating Model: After model generation, check the accuracy using actual and predicted values.We should estimate how accurately the classifier predicts the outcome. The accuracy is computed by comparing actual test set values and predicted values.

