**Email/SMS Spam Classifier**

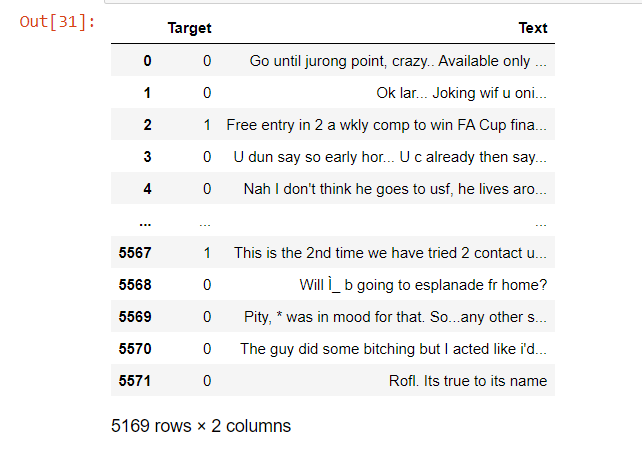
The following code helps in understanding if the provided message is spam or not, using keywords by Bayes' Classification Algorithm.

The classifier first takes a body of known spam and ham (non-spam) emails to evaluate. Then, it evaluates each email/sms in a test body as spam or ham, with the difference between ham and spam only known to the classifier for the purpose of calculating the success rate.

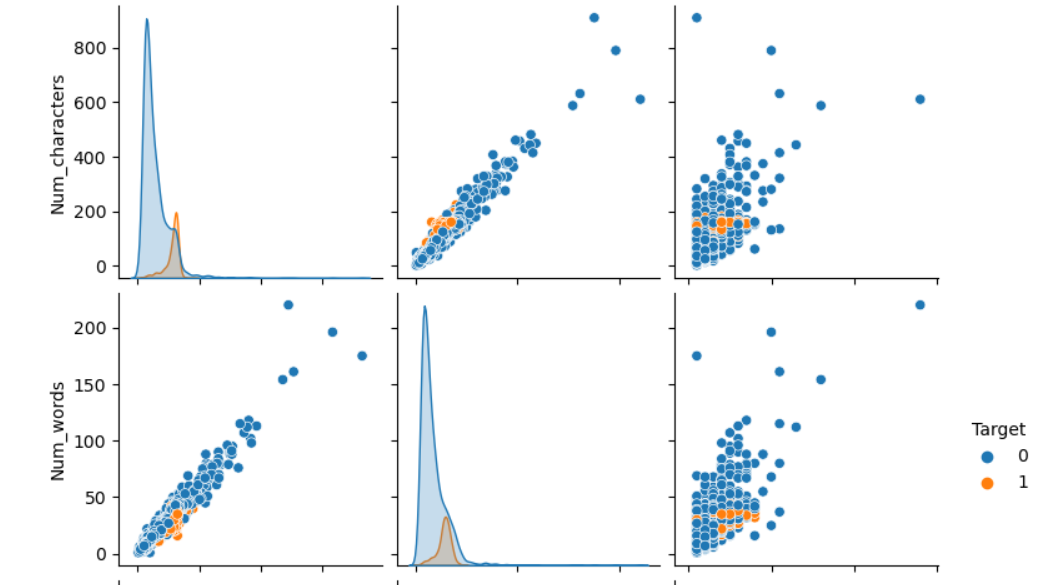
Several parameters can be changed to optimize the effectiveness of the classifier. By tweaking these parameters, rates in the upper 90% range for both spam and ham classification can be reached.

* Firstly we take the data from a random csv file which contains different sentences and keywords.

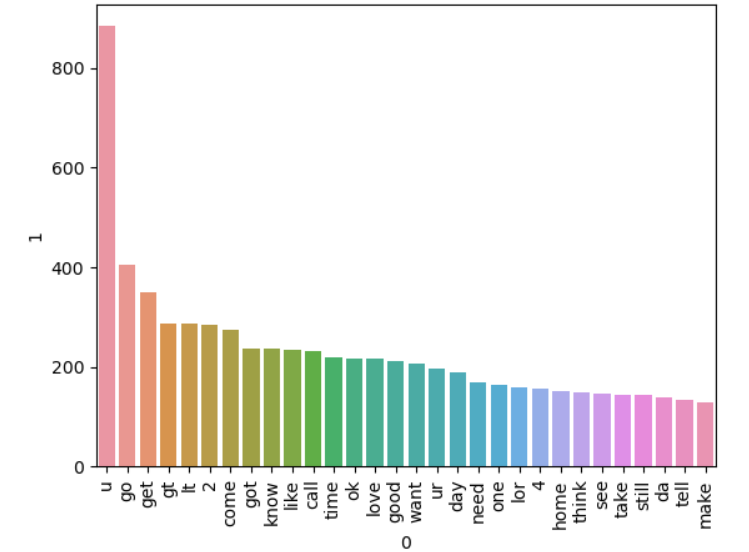
We clean the data so that the columns containing the duplicate or empty values could be reduced which ultimately will reduce the inconsistency in data.



* After this EDA is performed by importing nltk in order to analyse and visualise the data which would help in getting a clearer picture of the number of ham and spam from the given set of keywords.

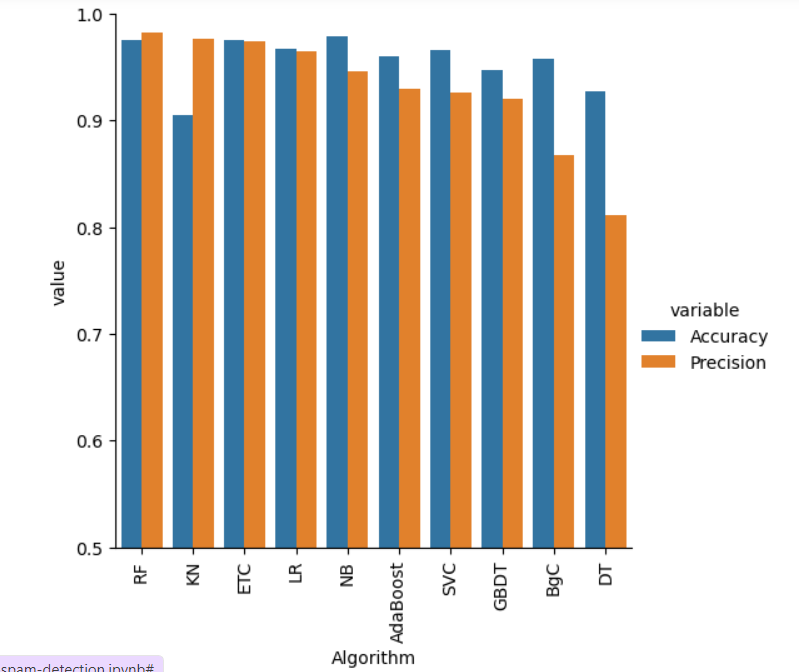


* Following this the preprocessing of the data is done which includes the usage of the function stopwords and counter which creates a dictionary for each word and we can count the frequency of each one of them.



* Lastly we build the model using various operations such as fitting and finding the best approach of classifier which gives us the best precision and accuracy.

We apply the stacking and the estimators for the data to be executed. Also the voting classifier has been used which is a machine learning estimator that trains various base models or estimators and predicts on the basis of aggregating the findings of each base estimator.



Atlast the code is executed from which we are able to predict whether the provided statement is ham or spam which can be seen as below.

