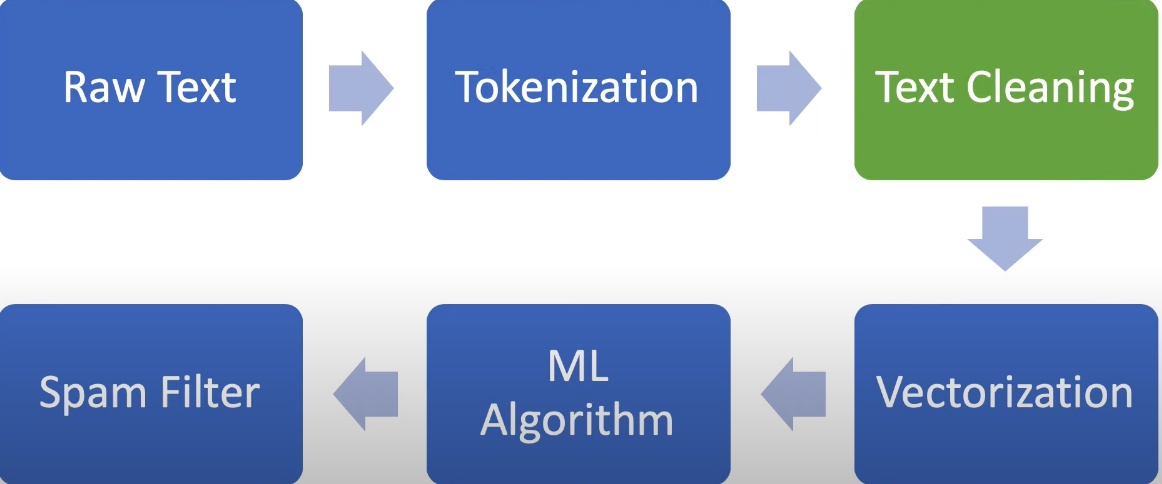
# Sentiment analysis for product Rating using SVM and Naïve Bayes

1. Imported the necessary Libraries to implement Sentiment Analysis
2. Importing pandas to read CSV file (Dataset) from the computer.
3. Also the display size of the text in the table is increased using 'set\_option('display.max\_colwidth',200)' in pandas library



Product Rating

Fig: Processing of the model

Preprocessing

1. The columns in Data frame are converted into integer and string datatype. also the text in the review column are converted into lower cases because our model can predict the word 'THE' an 'the' as different entities.
2. Using Regular Expression for correcting the word so that word redundancy can reduced.
3. Removing punctuation using String library
4. Removing Number Because the Number will not that helpful to reflect the sentiment / emotion
5. Tokenization : we have to iterate through each and every word in the sentence of the dataset
6. Removing Stop word : from the library nltk.corpus.stopwords.words('english')
7. Lemmatization: Finds root word of the similar words which helps to reduce the computation of our model.
8. Finding NA values in our huge dataset. we find those ones and fill them with the average value in the dataset i.e. rating
9. Splitting Dataset for Training and Testing
10. Feature Extraction: generating matrix value of text Dataset using Countvectorizer then converting them into TFIDF (term frequency–inverse document frequency)
11. Implementing the Naive Bayes model and fitting the training dataset
12. Calculating precision and then generating confusion matrix for Accuracy of our model for both naïve and Svm.