**IMPLEMENTATION:**

**MODULES:**

* User
* Admin
* Data Preprocessing
* Machine Learning

**MODULES DESCRIPTION:**

**User:**

The User can register the first. While registering he required a valid user email and mobile for further communications. Once the user register then admin can activate the user. Once admin activated the user then user can login into our system. User can upload the dataset based on our dataset column matched. For algorithm execution data must be in float format. Here we took Email dataset for testing purpose. User can also add the new data for existing dataset based on our Django application. User can click the Classification in the web page so that the data calculated Accuracy and Precision, Recall based on the algorithms. User can click

Prediction in the web page so that user can write the review after predict the review

That will display results depends upon review like postive,negative or neutral

**Admin:**

Admin can login with his login details. Admin can activate the registered users. Once he activate then only the user can login into our system. Admin can view the overall data in the browser. Admin can click the Results in the web page so calculated Accuracy and Precision, Recall based on the algorithms is displayed. All algorithms execution complete then admin can see the overall accuracy in web page.

**Data Preprocessing:**

A dataset can be viewed as a collection of data objects, which are often also called as a records, points, vectors, patterns, events, cases, samples, observations, or entities. Data objects are described by a number of features that capture the basic characteristics of an object, such as the mass of a physical object or the time at which an event occurred, etc. Features are often called as variables, characteristics, fields, attributes, or dimensions. The data preprocessing in this forecast uses techniques like removal of noise in the data, the expulsion of missing information, modifying default values if relevant and grouping of attributes for prediction at various levels.

**Machine learning**:

Based on the split criterion, the cleansed data is split into 60% training and 40% test, then the dataset is subjected to machine learning classifiers such as Iterative Dichotomiser3 (ID3). The accuracy and Precision, Recall of the classifiers was calculated and displayed in my results. The classifier which bags up the highest accuracy could be determined as the best classifier.