Value prediction system for sentimental analysis of customer reviews

Team:

Puneet Desai (44876946)
Shashank Gururaj (44888163)
Shreya Nair (44775900)
Shivam Gaur (44922647)
Sangay Pelzang (44961093)

Set up Guide for Project

1. Contents of Data_Mining_Project.zip folder

The zip folder consists of the following four .ipynb (ipython notebook) files in the System_Code folder. The purpose of these files is mentioned below:

1. *CountVector_Step_By_Step_Implementation.ipynb*: This file demonstrates step by step implementation of the various steps involved in implementing the project. The text transformation uses Count Vectorization Technique.

2. *TFIDF_Step_By_Step_Implementation.ipynb:* This file demonstrates step by step implementation of the various steps involved in implementing the project. The text transformation uses TFIDF Technique.

The objective of the above two files is to explain the process flow of various stages in the software execution and the code used to implement those steps.

3. *Sentimental_Analysis_OfCustomerReviews_System_Using_Count_Vectorization_Technique.ipynb:* This file consists of the code to run the software and uses Count Vectorization Technique for text transformation. The software consists of GUI for easy usability and allow users to predict the rating of new review, analyze and compare different classification models based on their accuracy, precision, recall and f-measure.

4. *Sentimental Analysis OfCustomerReviews System Using TFIDF Technique.ipynb:* This file consists of the code to run the software and uses TFIDF Technique for text transformation. The software consists of GUI for easy usability and allow users to predict the rating of new review, analyze and compare different classification models based on their accuracy, precision, recall and f-measure.

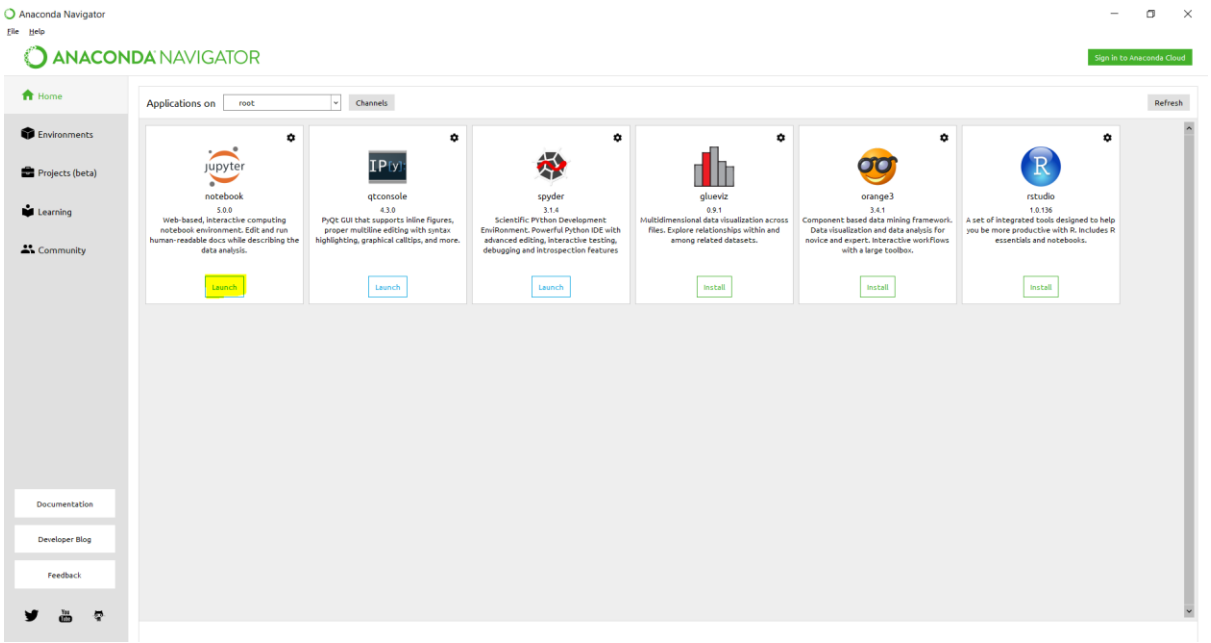
2. Instructions to run the .ipynb files.

To run the above four files, follow the following steps:

1. Download Anaconda 5.0.0 for Windows or Mac OS from the link: https://www.anaconda.com/download/. We recommend downloading Anaconda for the latest Python 3.6 version.
2. Install and run the downloaded version of Anaconda on the computer.
3. 'Jupyter Notebook' has been installed. To run the 'Jupyter Notebook' execute the below command:

```
jupyter notebook
```

4. Search Anaconda Navigator in search bar of your OS and open Anaconda Navigator. The below screen will be displayed.



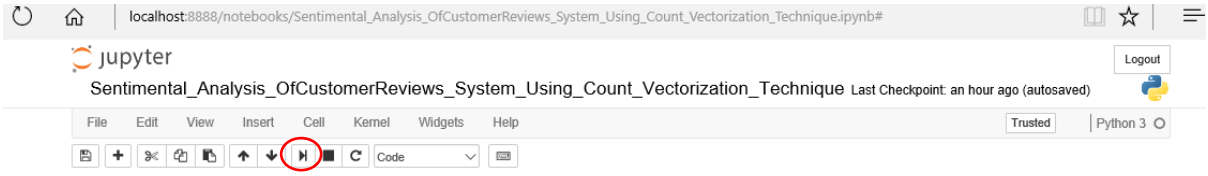
5. Click on “Launch” button in the Jupyter Notebook Tab. Home screen of ‘Jupyter Notebook’will be open in your default web browser.

6. To run the .ipynb files, click on the upload button on the top left corner of the ‘Files’ tab of the home screen.

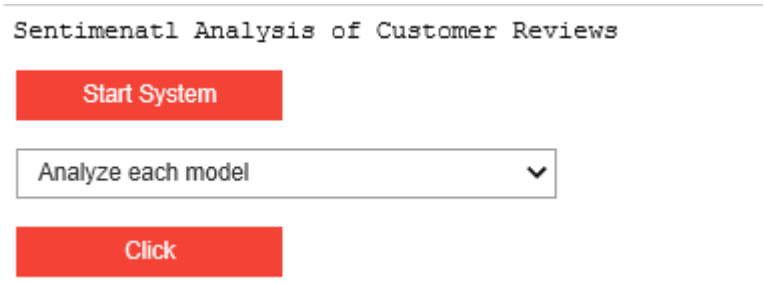
A dialog box will appear. Browse the file to run and click open. Then click on the blue upload button.

7. Before running the code files, upload the dataset file, "Amazon_Unlocked_Mobile.csv" in the same way as the above code files are uploaded in the ‘Jupyter Notebook’.

8. To run the file, search the uploaded file in the list of files under the ‘Files’ tab on the home screen. The file will open in new tab of the browser. Then click on the ‘Run Cell’ button to run the code (shown below) **Please Note:** Execution might take a few minutes.



9. Go to the bottom of the code to view the output of the code. At first click on START SYSTEM, and then select an option from the drop box.



10. Three options are available in the drop-down box:
    1. *Analyze each model:* Displays ROC curve, Accuracy and Precision, Recall and F-Measure for each model. The model for analysis can be chosen further from the new dropdown box.
    2. *Predict rating for New Review:* Predicts the rating of the new customer review provided in the text box.
    3. *Compare Models:* Displays a bar graph of accuracy of all the models.