

SENTIMENT ANALYSIS



BY

V BHUVIDHA	(16BF1A0555)
C RAMYA SREE	(16BF1A0522)
G HARSHAVARDHAN	(16BF1A0537)
K POORNACHANDRA	(16BF1A0549)

**UNDER THE GUIDANCE OF
DR.P.DILEEP KUMAR REDDY
ASSOCIATE PROFESSOR.**

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
SRI VENKATESWARA COLLEGE OF ENGINEERING
KARAKAMBADI ROAD
TIRUPATI-517507**

OUTLINE



Abstract
Introduction
Existing System
Proposed System
Implementation Details
Modules
 UML Diagrams
Types of Testing
Test Cases
Screen Layouts
Conclusion
References

ABSTRACT



- Sentiment Analysis is a major subject in machine learning which aims to extract subjective information from the textual reviews.
- It is the task of identifying whether the opinion expressed in a text is positive or negative or neutral. It is used to determine the attitude of the reviewer with respect to the various topics or the overall polarity of reviews.
- Social media websites have emerged as one of the platforms to raise user's opinions and influence the way any business is commercialized.
- Opinion of people matters a lot to analyze how the propagation of information impacts the lives in a large-scale network.
- It is simple and efficient Neural Language Model approach for text classification that relies only on unsupervised word representation inputs.
- For the classification of sentiments whether being positive, negative or neutral we use the best algorithm of Neural networks. By using all these techniques we try to find the maximum accuracy than others.

INTRODUCTION



- ***Sentiment analysis*** is contextual mining of text which identifies and extracts subjective information in source material, and helping a business to understand the social sentiment of their brand, product or service while monitoring online conversations.
- To the use of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information.
- Since publicly and privately available information over Internet is constantly growing, a large number of texts expressing opinions are available in review sites, forums, blogs, and social media.

EXISTING SYSTEM



- The proposed Global Belief-Recursive Neural Network represents the state-of-the-art for granular sentiment analysis.
- In order to correctly capture contextual sentiment, a backward step from upper tree nodes is introduced here.
- A different approach is obtained by considering aspect specific sentiment analysis using hierarchical deep learning according to [LSM].
- Here, separate aspect sentiment models (SAS) or Joint Multi-Aspect Sentiment models (JMAS) train root-node-level soft max classifiers with aspect and sentiment as classification outputs.
- However, a prelabeling of the aspects/product features is required for these models.
- This work should explore different recurrent neural networks (RNN) including bi-directional recurrent NN and Long Short-Term Memory (LSTM) RNN's that try to capture aspect specific sentiment through context.

PROPOSED METHOD



❖ Sentiment Analysis using Recurrent Neural networks in Keras with Tensorflow as backend

Algorithm for Sentimental Analysis using RNN

- 1) First we need to convert the raw text-words into so-called tokens which are integer values.
- 2) Then we convert these integer-tokens into so-called embeddings which are real-valued vectors, whose mapping will be trained along with the neural network, so as to map words with similar meanings to similar embedding-vectors.

PROPOSED METHOD



- 3) Then we input these embedding-vectors to a Recurrent Neural Network which can take sequences of arbitrary length as input and output a kind of summary of what it has seen in the input.
- 4) Output from the RNN is squashed by an activation function (Sigmoid in this case)
- 5) output is between 0 and 1
{ 0: highly negative, 1 : highly positive }

IMPLEMENTATION DETAILS



Hardware Requirements:

- RAM Capacity : 12GB
- Processor : Any 8th Gen Processor
- Graphics Card : 4GB
- Speed : 2.5 GHZ or more

IMPLEMENTATION DETAILS



Software Requirements:

•Operating System	:	Windows/Linux
•Programming Language	:	Python Environment
•Browser	:	Mozilla/IE/Chrome
•Framework	:	Tensor Flow, NumPy
•Deep Learning Module	:	Keras
•Web Application	:	Flask
•IDE	:	Spyder
•Python Version	:	3.8.2

MODULES



There are five modules in this application. They are:

- Download
- imdb
- Tokenizer
- Sequential model
- Flask

MODULES



- **Download:** This module is used to extract text reviews or text from the given url.
- **Imdb:** This module provides functionality to load data. This also has methods to load training data and test data.
- **Tokenizer:** Since computer cannot understand text, first we need to convert the raw text-words into so called tokens which are integer values. Tokenizer provides the required functionality.

MODULES



- **Sequential model:** This model is used to create Recurrent Neural Network. It provides methods to add different types of network layers sequentially, each performing unique operation on the tokenized text data to learn the mapping and understanding the context. It provides method to predict output of unseen input.
- **Flask:** This is a python framework used to create web pages. This helps us to take the corresponding input from the user and to render the output of our model through web page in browser.

UML DIAGRAMS

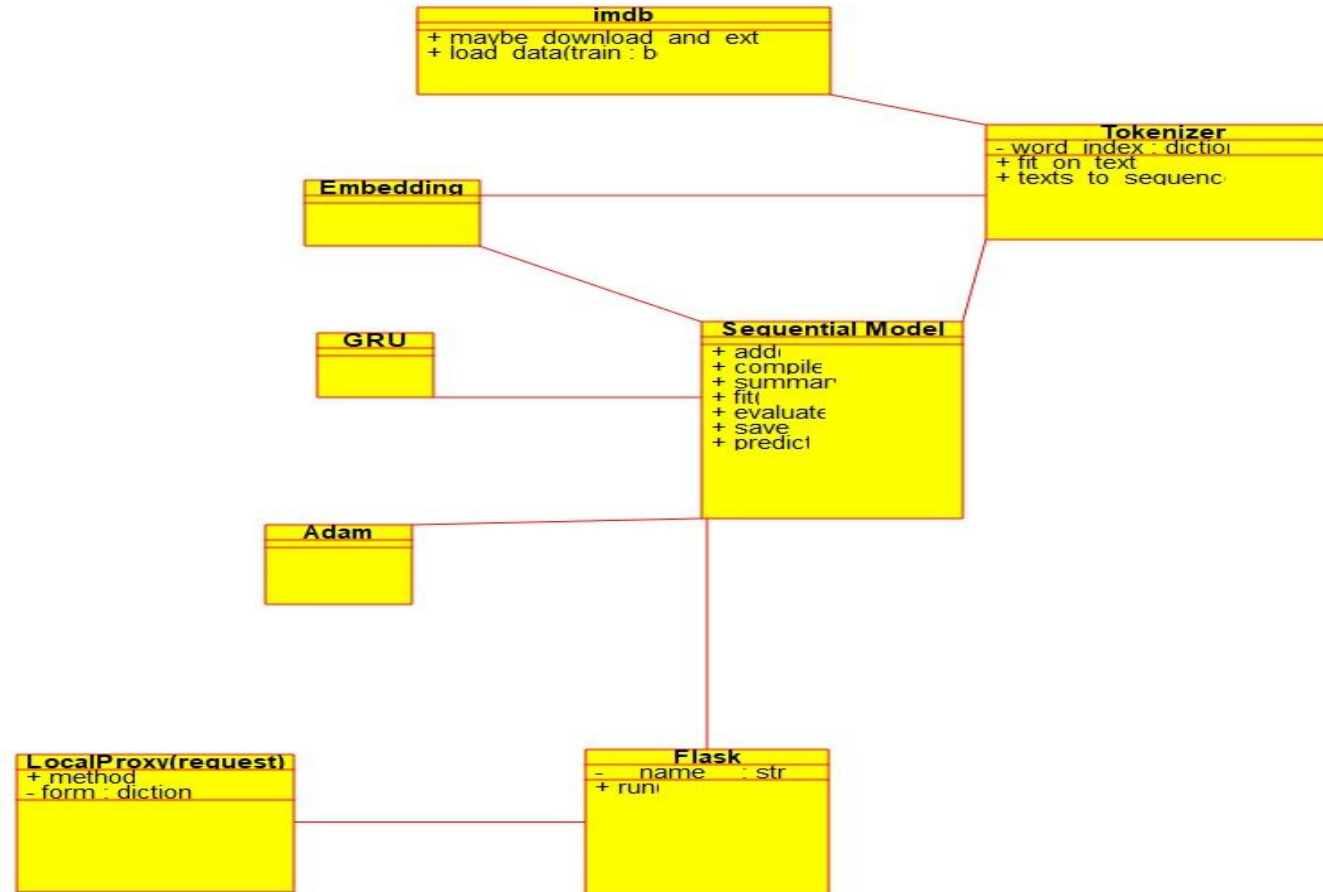


Goals:

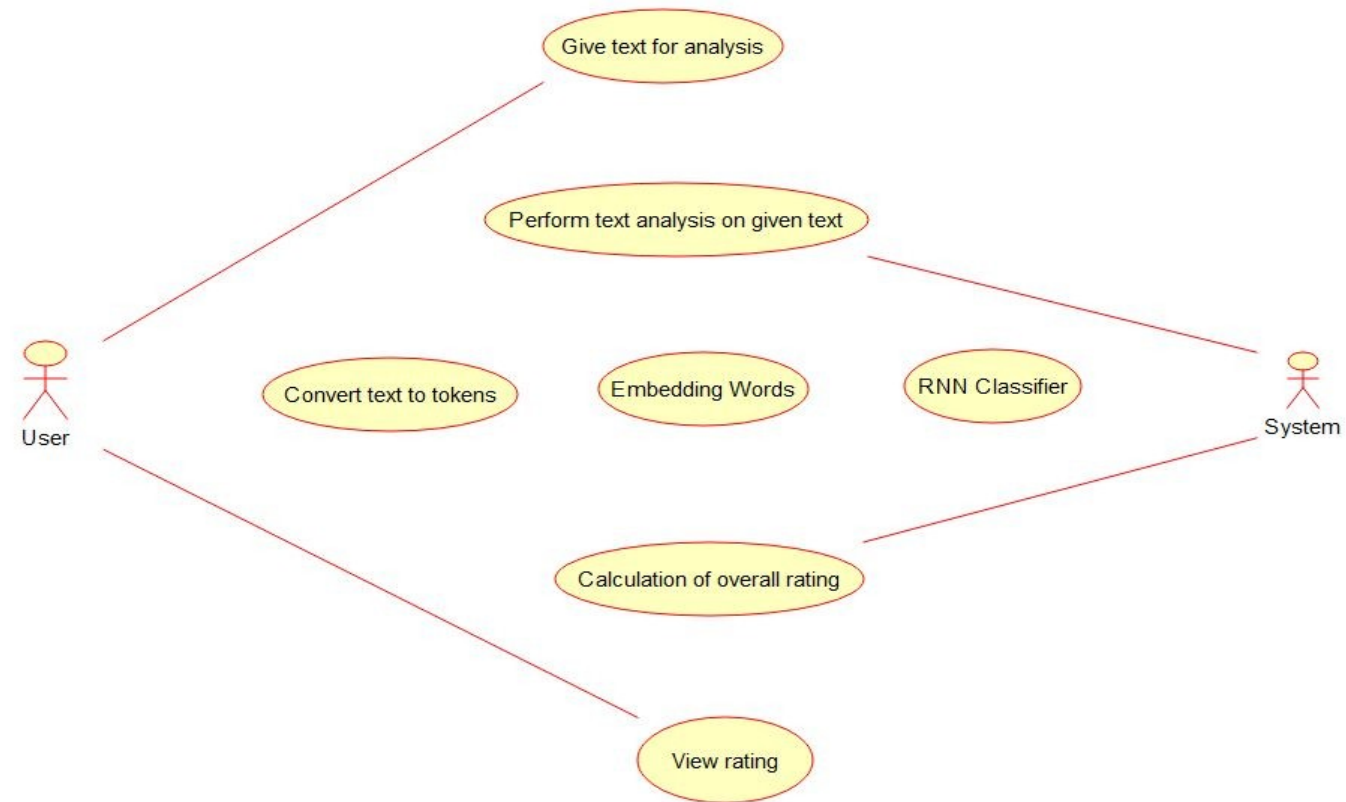
The primary goals in the design of the UML are as follows:

- Be independent of particular programming languages and development processes.
- Provide a formal basis for understanding the modeling language.
- Encourage the growth of the OO tools market.
- Support higher-level development concepts such as collaborations, frameworks , patterns and components.
- Integrate best practices.

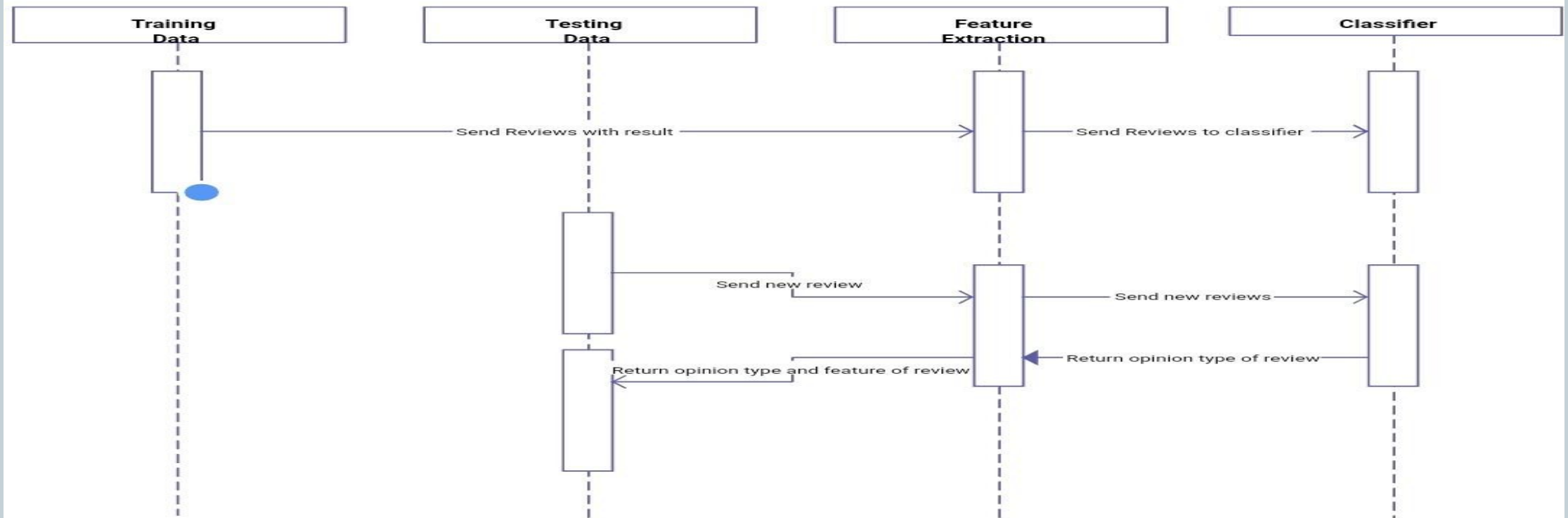
CLASS DIAGRAM



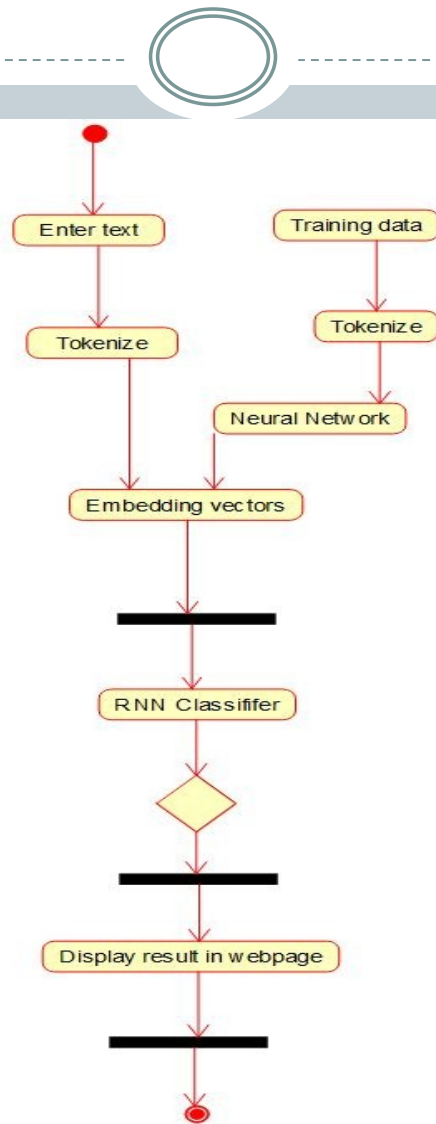
USE CASE DIAGRAM



SEQUENCE DIAGRAM



ACTIVITY DIAGRAM



TYPES OF TESTING



UNIT TESTING:

Each module can be tested in the following two Strategies:

Black box Testing:

It has been used to find errors in the following categories:

- Incorrect or missing functions Interface errors.
- Errors in data structure or external database access
- In this testing only, the output is checked for correctness.
- The logical flow of the data is not checked.
- Performance errors Initialization and termination errors.

TYPES OF TESTING



White box Testing:

It has been used to generate the test cases in the following cases:

- Execute all logical decisions on their true and false sides.
- Guarantee that all independent paths have been executed.
- Execute all loops at their boundaries and within their operational bounds.

INTEGRATION TESTING:

- Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields.
- It is specifically aimed at exposing the problems that arise from the combination of components.

TYPES OF TESTING



FUNCTIONAL TESTING:

- Functional tests provide systematic demonstration that functions tested are available specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

- Valid Input : identified classes of valid input must be accepted.
- Invalid Input : identified classes of invalid input must be rejected.
- Functions : identified functions must be exercised.
- Output : identified classes of application outputs must be exercised.
- Systems/Procedures : interfacing systems or procedures must be invoked.

SYSTEM TESTING:

- Involves in-house testing of the entire system before delivery to the user. Its aim is to satisfy the user the system meets all requirements of the client's specifications.

TYPES OF TESTING



Testing Approach:

Testing can be done in two ways:

- Bottom up approach
- Top down approach

Bottom up Approach:

- Testing can be performed starting from smallest and lowest level modules and proceeding one at a time. For each module in bottom up testing a short program executes the module and embedded within it provides the needed data so that the module is asked to perform the way it will when used within the larger system. When bottom level modules are tested attention turns to those on the next use the lower level ones they are tested individually and then taken with the previously examined lower level modules.

TYPES OF TESTING



Top down Approach:

- This type of testing starts from upper level modules.
- Since the detailed activities usually performed in the lower level routines are not provided stubs are written.
- A stub is a module shell called by upper level module and that when reached properly will return a message to the calling module indicating that proper interaction occurred.
- No attempt is made to verify the correctness of the lower level module.

TYPES OF TESTING



ACCEPTANCE TESTING

- User Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the system meets the functional requirements.

Test Results:

- All the test cases mentioned above passed successfully. No defects encountered.

TEST CASES



Input text	Expected output	Output	Result
The movie is bad	Negative with confidence range between 0.9 and 1	Negative with confidence 0.98	Pass
Udacity is best platform for Professional learning	Positive with confidence range between 0.9 and 1	positive with confidence 0.97	Pass
The sound quality is good but not upto expectations	Positive with confidence range between 0.4 and 0.6	Positive with confidence 0.55	Pass

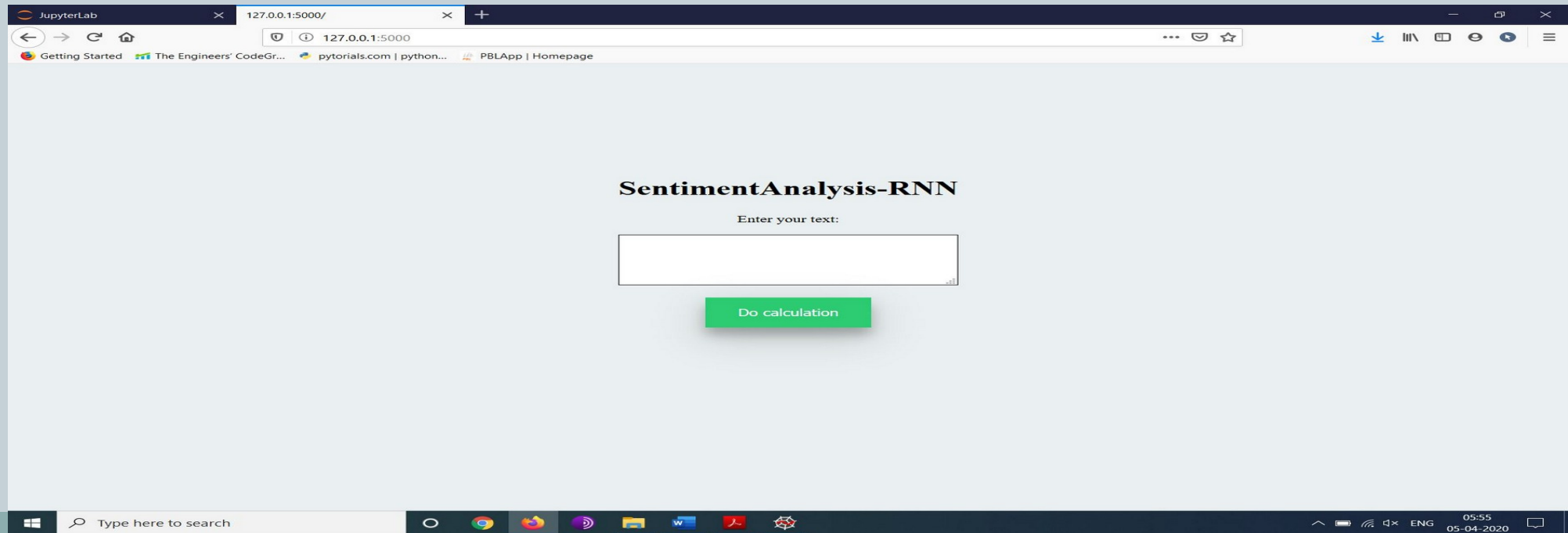
SCREEN LAYOUTS



Screen Layout 1: Home Page

Description:

- User has to enter the review in the text box available and click on submit button to process the review.



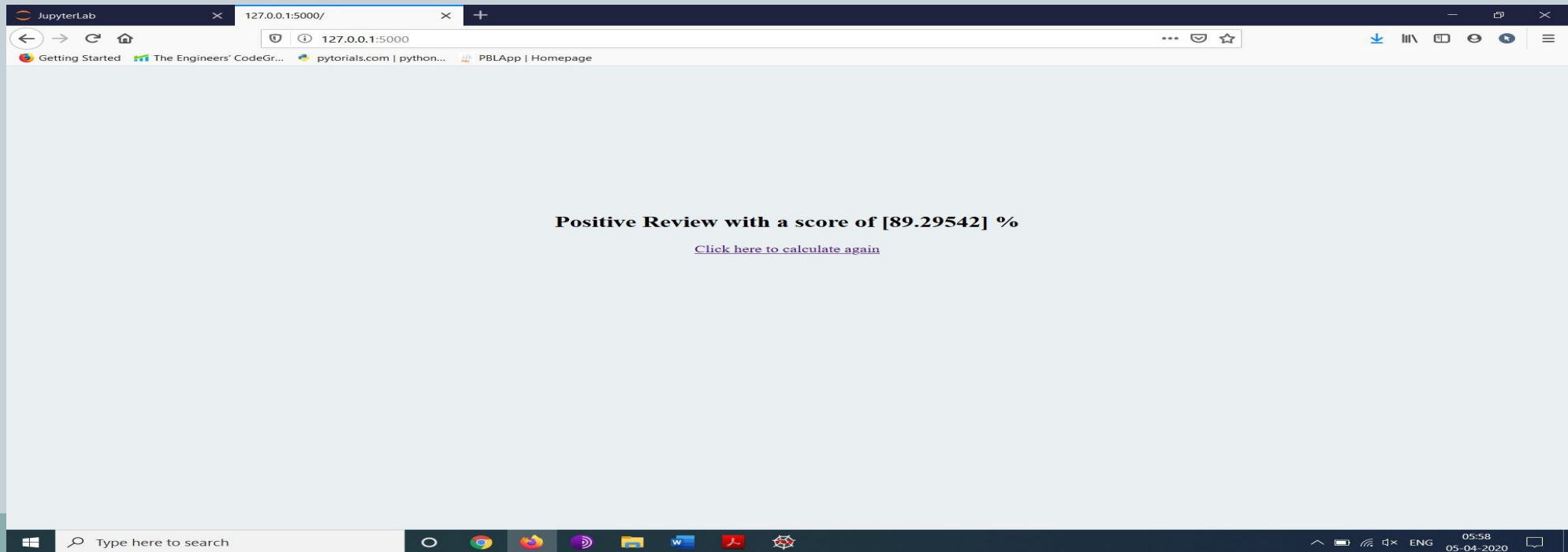
SCREEN LAYOUTS



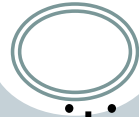
Screen Layout 2: Result Page

Description:

- This page displays the evaluated results of the sentiment of the review with score of it.



CONCLUSION



- The field of sentiment analysis is an exciting new research direction due to large number of real-world applications where discovering people's opinion is important in better decision-making.
- The development of techniques for the document-level sentiment analysis is one of the significant components of this area. Recently, people have started expressing their opinions on the Web that increased the need of analyzing the opinionated online content for various real-world applications.
- A lot of research is present in literature for detecting sentiment from the text. Still, there is a huge scope of improvement of these existing sentiment analysis models. Existing sentiment analysis models can be improved further with more semantic and commonsense knowledge.

REFERENCES



Web Resources:

- [1] <https://www.machinecurve.com/index.php/2020/02/08/how-to-use-padding-with-keras/>
- [2] <https://keras.io/getting-started/sequential-model-guide/>
- [3] "http://ai.stanford.edu/~amaas/data/sentiment/aclImdb_v1.tar.gz"
- [4] <https://blog.keras.io/keras-as-a-simplified-interface-to-tensorflow-tutorial.html>
- [5] <https://machinelearningmastery.com/data-preparation-variable-length-input-sequences-sequence-prediction/>



Book Resources:

- [1] Laura Auria, Rouslan A Moro, “Support Vector Machines”, 2008, web, 20 August 2014.
- [2] H Kim, P Howland, and H Park “Dimension reduction in text classification with support vector machines”, Journal of Machine Learning Research, 2005.
- [3] Pang- Ning Tan, Michael Steinbach, Vipin Kumar, “Introduction to Data Mining”, Pearson publications.



A hand holding a white card with the text "Any Questions?" written on it. The word "Any" is written in red, and "Questions?" is written in black. The hand is wearing a blue and white striped shirt. The card is held against a light gray background.

Any
Questions?

THANK YOU