



ONLINE MARKET PLACE FOR ORGANIC FOODS USING BLOCKCHAIN

Project ID: 19_20-J 06

Software Requirement Specifications (SRS) for the component

‘Aspect-Based Sentiment Analysis for Customer Reviews of Organic Foods ’

A.W.M.J.S. Bandara (IT15145994)

BSc Special (Hons) - Information Technology

(Specialization in Information Technology)

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

August, 2019

Declaration

I declare that this is my own work and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where acknowledgement is made in the text.

Name	Student ID	Signature
A.W.M.J.S.Bandara	IT15145994	

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

.....

.....

Signature of the supervisor

Date

Table of Contents

Contents

Declaration	2
List of Tables	4
List of Figures.....	4
1. Introduction	5
1.1 Purpose.....	5
1.2 Scope	5
1.3 Definitions, Acronyms, Abbreviations	6
1.4 Overview	7
2. Overall Description.....	9
2.1 Product Perspective	10
2.1.1 User Interfaces	10
2.1.2 Hardware Interfaces	12
2.1.3 Software Interfaces	12
2.1.4 Communication Interfaces.....	12
2.1.5 Memory constraints	12
2.1.6 Operations	12
2.1.7 Site adaptation requirement	12
2.2 Product functions.....	13
2.2.1 Use case diagram.....	14
2.2.2 Use case scenario	15
2.3 User characteristics.....	16
2.4 Constraints.....	17
2.5 Assumptions and dependencies	17
2.6 Apportioning of requirements	17
3. Specific Requirements	18
3.1 External interface requirements.....	18
3.1.1 Hardware interfaces.....	18
3.1.2 Software interfaces	18
3.1.3 Communication interfaces	18
3.2 Classes/Objects.....	19

3.3 Performance requirements	20
3.4 Design constraints	20
3.5 Software system attributes	20
3.5.1 Reliability	20
3.5.2 Availability	20
3.5.3 Security	21
3.5.4 Maintainability	21
4. Supporting requirements	21
References	21

List of Tables

Table 1. 1 Definitions, Acronyms, Abbreviations	6
Table 1. 2 Use case scenario – Collecting the reviews to be analyzed.	15
Table 1. 3 Use case scenario – Data cleansing	15
Table 1. 4 Use case scenario – Sentiment analysis of the reviews.	16
Table 1. 5 Use case scenario – View analyzed results.	16

List of Figures

Figure 1. 1 Overview	7
Figure 1. 2 Product Overview Page.....	10
Figure 1. 3 Overall Aspect-based Sentiment Analysis Overview	11
Figure 1. 4 Showing the final output flow	13
Figure 1. 5 Use case diagram	14
Figure 1. 6 Class diagram	19

1. Introduction

Aspect-based Sentiment Analysis (ABSA) is a text analysis technique that breaks down text into aspects (attributes or components of a product or service) and allocates each one a sentiment level. This technique can help businesses become customer-centric and place their customers at the heart of everything they do. It's about listening to their customers, understanding their voices, analyzing their feedbacks and learning more about customer experiences, as well as their expectations for products or services.

1.1 Purpose

The purpose of this SRS document is to provide a detailed overview of Aspect-based Sentiment Analysis approach, its parameters and goals. This document describes the project's target audience and its user interface, hardware and software requirements. It defines how our client, team and audience see the product and its functionality. Nonetheless, it helps any designer and developer to assist in software delivery lifecycle (SDLC) processes.

Also, this document describes how the quality grader system will differ from other systems and how it has targeted to achieve its objectives while overcoming the existing problems.

In this document detailed overview of development of comprehensive software platform for analyzing user reviews for products in our system is explained.

1.2 Scope

In this document I have described how to build aspect-based sentiment analysis model to identify user reviews regarding organic products in the systems.

Instead of classifying the overall sentiment of a text review on product into positive or negative, aspect-based analysis allows us to associate specific sentiments with different aspects of the product. The results are more detailed, interesting and accurate because aspect-based analysis looks more closely at the information behind a text review.

Time to time people tend to use new words for comments which has explicit meanings which have different meaning rather than actual meaning. Those words are needed to be identified as they play a significant role by opinions posted in the system in order to make decisions on products.

Each time consumers interact with the product, by considering user reviews, products are receiving valuable insights.

.

1.3 Definitions, Acronyms, Abbreviations

ABSA	Aspect Based Sentiment Analysis
DB	Database
SRS	Software Requirement Specification
GUI	Graphical User Interface
UI	User Interface
API	Application Programming Interface
AI	Artificial Intelligence
NLP	Natural Language Processing

Table 1. 1 Definitions, Acronyms, Abbreviations

1.4 Overview

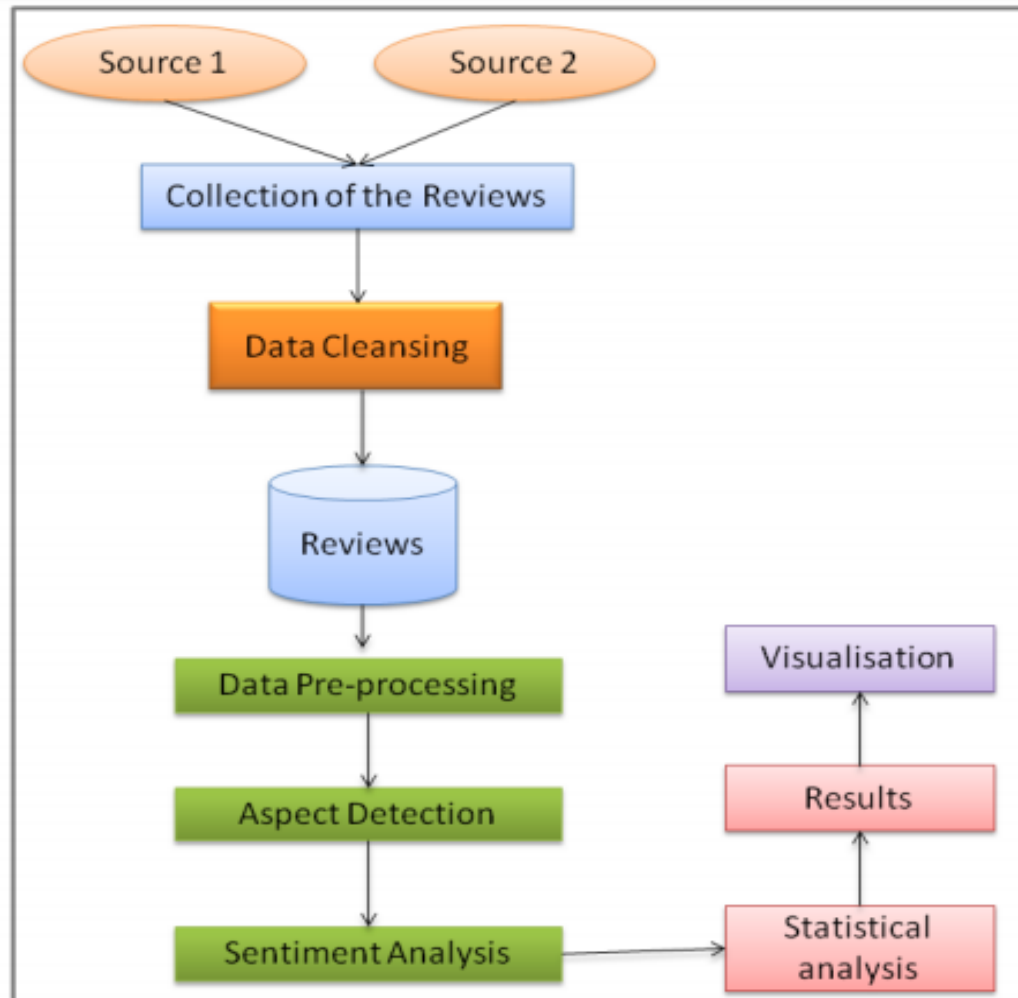


Figure 1. 1 Overview

This document presents the methodology and the detailed information about the process of designing the aspect-based sentiment analysis tool product reviews. System will contain following major modules and they will be having the following objectives as above diagram.

Data Collection

Before we can start any kind of text analysis, we need to gather information.

User reviews regarding organic foods, must be retrieved from the system in order to perform aspect-based sentiment analysis as the first step.

Data Pre-processing

After the data collection, the text data is cleaned and cleansed. This phase of the methodology is the most important phase; in order to build a robust, stable and efficient system, the data must be properly cleaned.

Removing duplicate rows and unnecessary columns, removal of punctuations, stop words, white spaces, special characters, converting text to lowercase, stemming and lemmatizing will be performed in this stage.

Each review converted to a sentence after the data pre-processing stage for further analysis.

Aspect and Sentiment Detection

Once the data is cleaned and the reviews were converted to cleaned sentences, the sentiment was extracted from each sentence and a polarity was assigned. Similarly, Aspects were detected for each sentence and a separate column was made for each aspect with its polarity score. Finally, the aspects and sentiment of the sentences were stored in a tabular form.

Data Mining

Statistical operations like correlation and linear regressions were performed on the aspects and polarity score of the reviews for products. Hence, these statistical operations provide a better insight into the data and help by understanding the relationships among different factors in the dataset.

Visualization

This will visualize the results to the user. Polarity of the reviews and overall rating for the product will be shown according to the aspect graphically using various diagrams.

2. Overall Description

Customers are more vocal than ever. They love leaving feedbacks regarding several aspects while making them a valuable resource for businesses in order to decision making purposes.

Also, it's not possible to make decisions on products by looking to user reviews one by one and give responses to customer reviews as soon as possible because business has to be dealt with tons of reviews from several consumers manually. By using aspect-based sentiment analysis approach it's possible to make decisions on products and provide responses as soon as possible without looking to whole review aspect-wise without consuming more time.

If customers are unhappy with the way you handle queries and complaints or finds the user experience of your software clunky and inefficient, they'll simply look elsewhere for an alternative. American Express found that customers are willing to spend more money with a company that delivers an excellent service, which goes to show that people are less price conscious these days, and more focused on a premium customer experience.

This has motivated the trend for continues development of sentiment analyzing tools for social media which can output valuable information for businesses to improve their business processes.

Sentiment analyzing is a widely studied Natural language processing task. We focus on aspect-based sentiment analysis as an automated method which is used to extract positive or negative (polarity) attitudes from user reviews.

Business is able interpret text entries from customers and gain meaningful insights. Not only does this help managers make key decisions based on insights from their customers, it also helps employees become more efficient and less frustrated with time-consuming, monotonous tasks.

This approach breaks down text into aspects (attributes or components of a product or service) and allocates each one a sentiment level. This technique can help business become customer-centric and place their customers at the heart of everything they do. It's about listening to their customers, understanding their voice, analyzing their feedback and learning more about customer experiences, as well as their expectations for products. Aspect-based sentiment analysis is particularly relevant at the moment because companies need to be more customer-centric than ever. This text analysis model lets our system to identifying on the specific aspects that make our customers happy or unhappy. By gaining a deeper understanding, system can be able to create a seamless customer experience and increase customer retention. Moreover, consumers can have an amazing experience throughout the organic food purchasing process considering existing user reviews which made by other consumers because of this aspect-based sentiment analysis approach.

2.1 Product Perspective

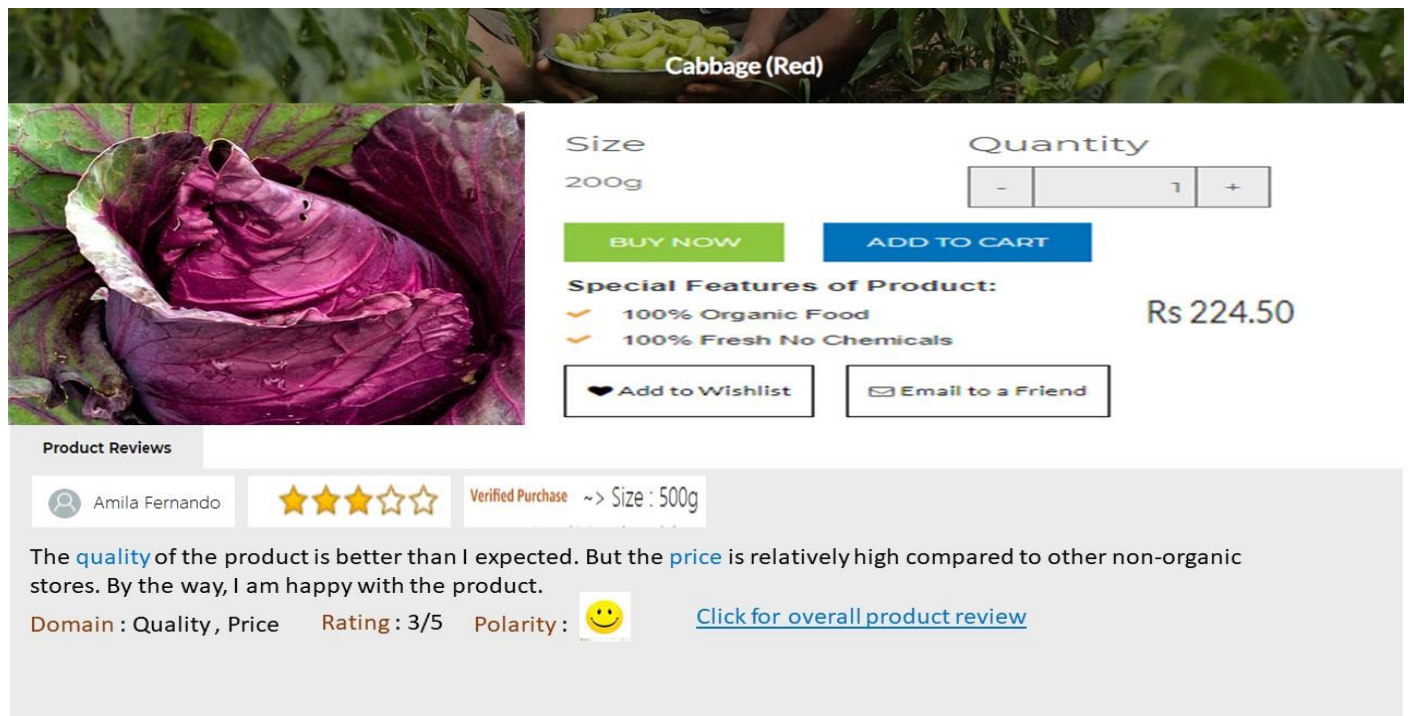
Currently there are so many customer reviews analyzing tools are available but most of those methods are not domain specific.

The typical sentiment analysis focus on predicting the positive or negative polarity of the given sentence(s). This task works in the setting that the given text has only one aspect and polarity. Proposed aspect-based sentiment analysis approach be able to predict the aspects mentioned in a sentence and the sentiments associated with each one of them.

Most sentiment analysis researches has focused on classifying the overall sentiment of a document into positive or negative. In here, however, often like to understand what are the specific sentiments towards different aspects of an entity, e.g. a restaurant review – “Food is decent but service is so bad.”, contains positive sentiment towards aspect food but strong negative sentiment towards aspect service. Classifying the overall sentiment as negative would neglect the fact that food was actually good.

Purpose of this model is to design a combined model with aspect prediction and sentiment prediction.

2.1.1 User Interfaces



The screenshot displays a product page for 'Cabbage (Red)'. It features a large image of the cabbage on the left. To the right, the product details include the size '200g', a quantity selector set to '1', and two buttons: 'BUY NOW' and 'ADD TO CART'. Below these, the 'Special Features of Product:' are listed as '100% Organic Food' and '100% Fresh No Chemicals'. The price 'Rs 224.50' is shown on the right. At the bottom, there are buttons for 'Add to Wishlist' and 'Email to a Friend'. A 'Product Reviews' section is visible at the bottom, showing a review by 'Amila Fernando' with a 3/5 star rating and a 'Verified Purchase' badge. The review text states: 'The quality of the product is better than I expected. But the price is relatively high compared to other non-organic stores. By the way, I am happy with the product.' Below the review, the domain is 'Quality, Price', the rating is '3/5', and the polarity is '😊'. A link 'Click for overall product review' is also present.

Cabbage (Red)

Size 200g

Quantity - 1 +

BUY NOW ADD TO CART

Special Features of Product:

- ✓ 100% Organic Food
- ✓ 100% Fresh No Chemicals

Rs 224.50

♥ Add to Wishlist ✉ Email to a Friend

Product Reviews

Amila Fernando ★★☆☆☆ Verified Purchase ~> Size : 500g

The **quality** of the product is better than I expected. But the **price** is relatively high compared to other non-organic stores. By the way, I am happy with the product.

Domain : Quality, Price Rating : 3/5 Polarity : 😊 [Click for overall product review](#)

Figure 1. 2 Product Overview Page

Description :

Customers who purchased organic products only can be able to make feedbacks/reviews and can rate through stars on them. To do so, customers should have a verified purchase towards that particular product. According to their feedbacks, corresponding aspects like quality, price, service will be identified. Individual review rating and polarity(positive, negative, neutral) for that individual feedback will also be identified and those data is shown through this interface.

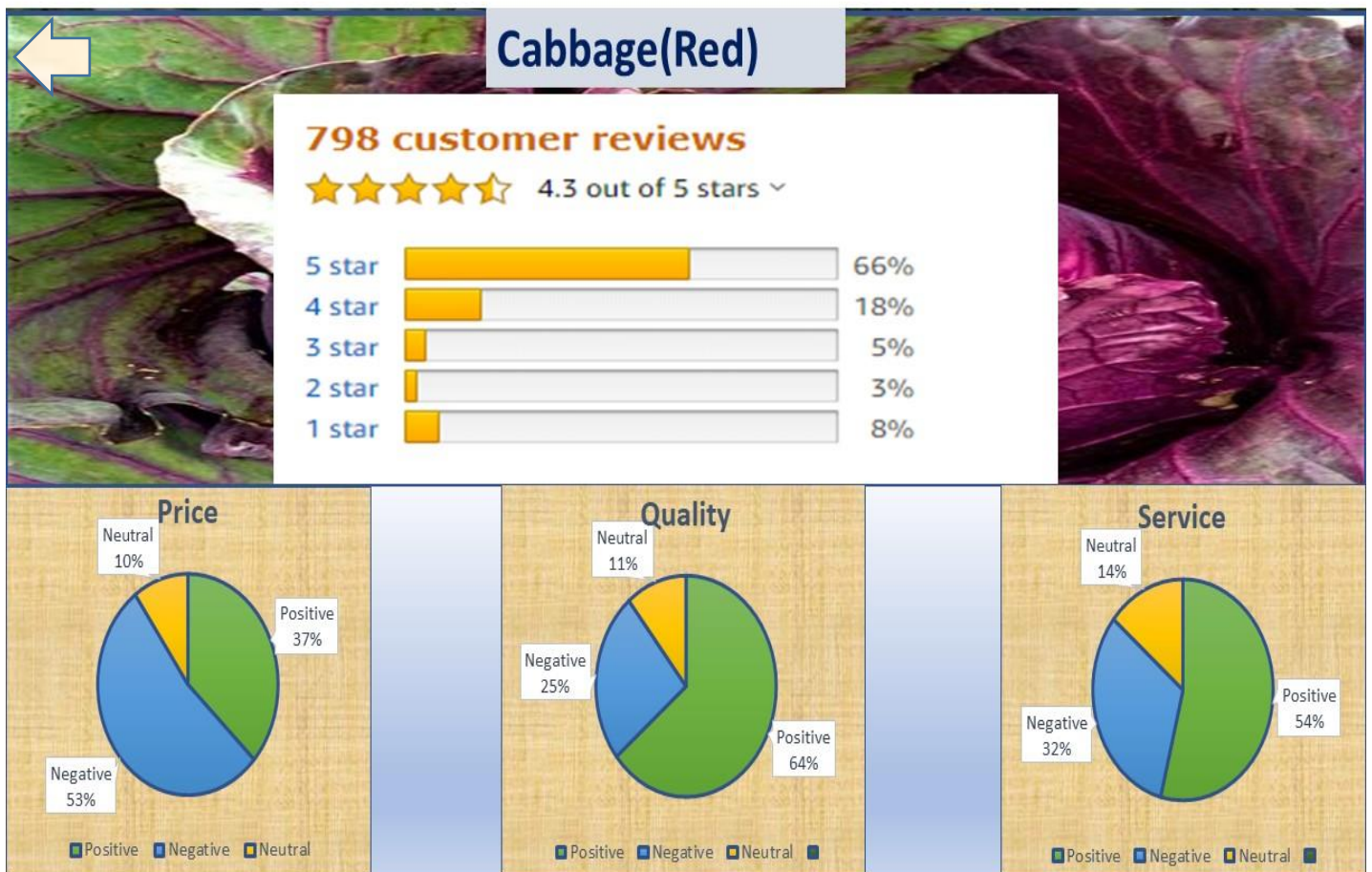


Figure 1.3 Overall Aspect-based Sentiment Analysis Overview

Description :

Aspect-based polarity for overall reviews/feedbacks will be graphically shown through this interface for consumers. Considering into these graphical data, consumers can be able to make decisions on particular organic products. Main aim of this research area is to provide more reliable service to consumers. That purpose is achieved through this interface.

2.1.2 Hardware Interfaces

No hardware interfaces will be used in this study.

2.1.3 Software Interfaces

- Python spaCy, NLTK and TextBlob libraries.
- Django framework for python.
- React.
- MongoDB.
- Visual Studio Code.

2.1.4 Communication Interfaces

This is a web application. We have designed a web site which is used to communicate with java script and the PHP as frontend and back end.

2.1.5 Memory constraints

A VPS of 1 core 1.5GHz processor, 4 GB of RAM, 10GB Hard disk will be used for the operations

2.1.6 Operations

Retrieve filtered (spams/irrelevant) product reviews from ontology model.

Aspect term extraction.

Aspect term polarity detection.

Aspect term categorization.

Aspect category polarity detection.

Visualize the results to the user.

2.1.7 Site adaptation requirement

Elements in the Graphical user interface must contain self-descriptive and easily understandable names.

Users will be able to see the aspect-based polarity of the reviews after the final production.

2.2 Product functions

This system has mainly four parts;

1. Aspect term extraction.
2. Aspect term polarity detection.
3. Aspect term categorization.
4. Aspect category polarity detection.

In this SRS document mainly discussed on the absolute positioning and generated the final output step by step.

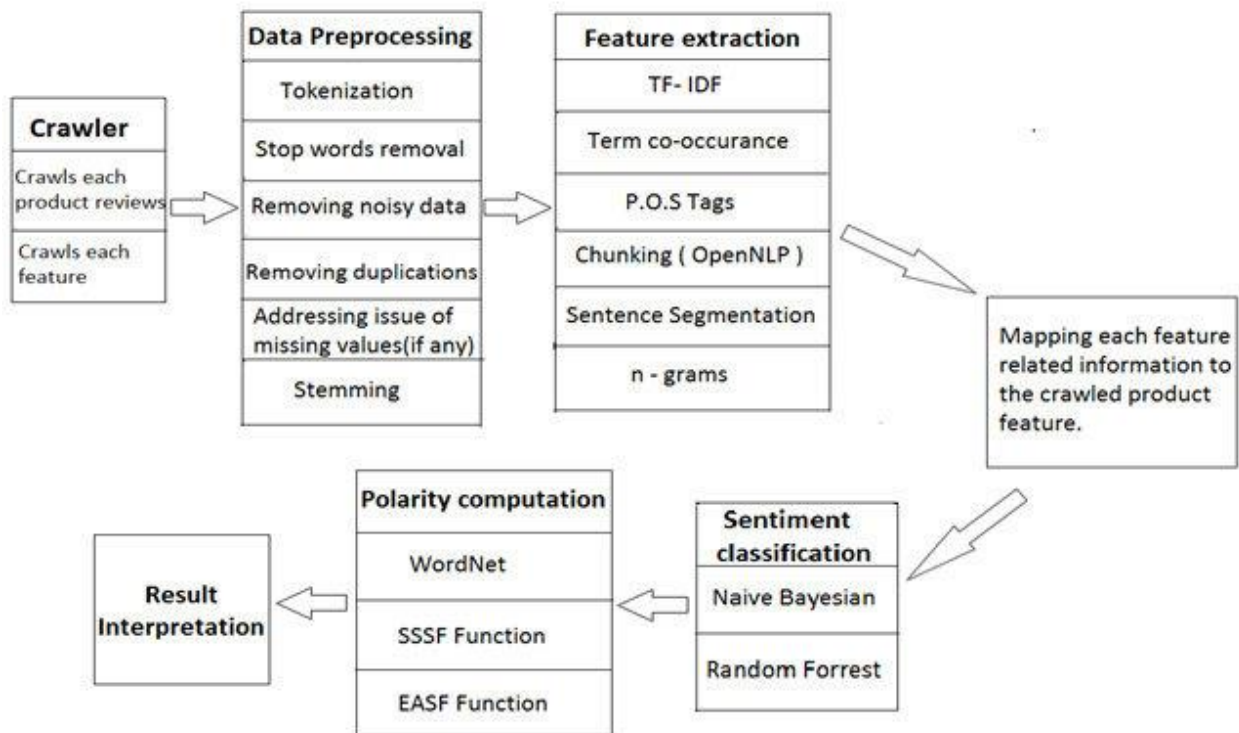


Figure 1. 4 Showing the final output flow

2.2.1 Use case diagram

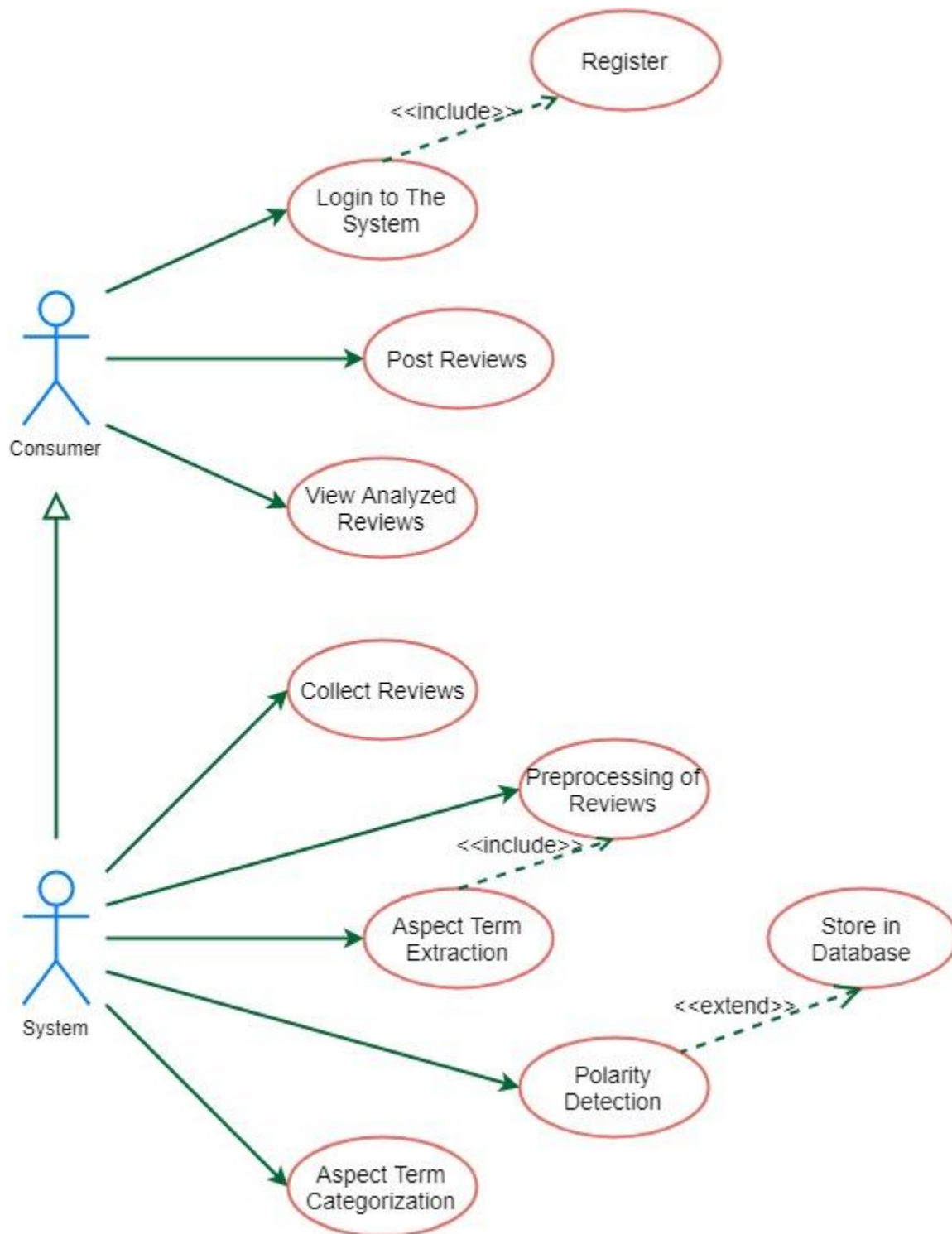


Figure 1. 5 Use case diagram

2.2.2 Use case scenario

Use case ID	UC_001
Use case	Collecting the reviews to be processed
Primary actor	System
Pre-condition	Reviews should be filtered using ontology model.
Post-condition	Provide filtered reviews for further analysis.
Main success scenario	<ol style="list-style-type: none">1. User enter reviews to the system.2. Retrieve those into ontology model by system.3. Filter those reviews accordingly.4. Provide those filtered reviews for cleansing..
Extension	1a. If the user review is not belong to the defined aspects, that review will be categorized as miscellaneous

Table 1. 2 Use case scenario – Collecting the reviews to be analyzed.

Use case ID	UC_002
Use case	Data cleansing
Primary actor	System
Pre-condition	Gather user reviews from ontology model
Post-condition	Build corpus.
Main success scenario	<ol style="list-style-type: none">1. Removing duplicate rows and unnecessary columns.2. Removal of punctuations, stop words, white spaces, special characters, converting text to lowercase etc.3. stemming and lemmatizing.4. Store the cleaned data in the database

Table 1. 3 Use case scenario – Data cleansing

Use case ID	UC_003
Use case	Sentiment analysis of the reviews.
Primary actor	System
Pre-condition	Corpus of the collected data.
Post-condition	Give aspect-based sentiment score.
Main success scenario	<ol style="list-style-type: none"> 1. Get corpus that stored in the database. 2. Use rules-based sentiment scores using nltk libraries. 3. Identify the aspects of the review. 4. Give sentiment score as positive / negative.

Table 1. 4 Use case scenario – Sentiment analysis of the reviews.

Use case ID	UC_004
Use case	View analyzed results.
Primary actor	System
Pre-condition	User should visit to the product.
Post-condition	Analyzed results will be shown in graphically.
Main success scenario	<ol style="list-style-type: none"> 1. Analyze the polarity of reviews. 1. Generate aspect-based graphs. 2. Result will be shown in graphically.

Table 1. 5 Use case scenario – View analyzed results.

2.3 User characteristics

Administrator: Administrators of the system can login to the main system as well as the lexicon library building system. They can view analytics data, feed data to the Machine Learning platform. Also, he/she can add/edit domain specific lexicons and their semantic scores manually to the library.

2.4 Constraints

Computer needs minimum of 1GHz 1 core processor, 2GB of RAM and internet connection for access Web GUI to use with the performance, users should have at least Minimum external storage of 40 MB

Hardware requirements are the mouse and keyboard.

I decided to provide this system in English language as a beginning stage. If I go further forward in time, I can release the Sinhala version too.

2.5 Assumptions and dependencies

This domain-based sentiment analysis system consists with English language. So, every user should have the English language literacy. And also, they should have post reviews in English.

AI should train around 3 months to get a good accuracy.

People have various kind of typing methods and they use different letters for same words.

No data will be lost/corrupted during the communication between server and web application.

2.6 Apportioning of requirements

Primary requirements of this project are described in section 1 and 2. Requirements that are specified in the section 3 are referred to as functional requirements. All these requirements are intended to be consistent. If there are any inconsistencies present, they will be logged as defects. Requirements that are mentioned in both section 2 and section 3 will be implemented by referring to the section 3 because it is more detailed than section 2.

3. Specific Requirements

3.1 External interface requirements

3.1.1 Hardware interfaces

Laptop or Desktop computer

USB dongle or ADSL router for connect to the internet.

3.1.2 Software interfaces

Python spaCy – Library for advanced Natural Language Processing(NLP).

NLTK 3.4.1 - Will be used to build the lexicon library building system.

POS-Tagger - Get pos tags for the sentence.

Stanford Parser – Extract relationships among different words of a sentence.

TextBlob library - Identify the polarity of the aspect terms.

Wordnet – A lexical database for the English language.

Django - Framework for python.

React – Will be used for frontend development.

MongoDB – Will be used for backed development.

Visual Studio Code - IDE

Edraw Max – To draw diagrams

Any web browser.

GitHub – To integrate the whole system.

3.1.3 Communication interfaces

Internet connection and a web browser are required in order to make use of several functions and to be executed.

3.2 Classes/Objects

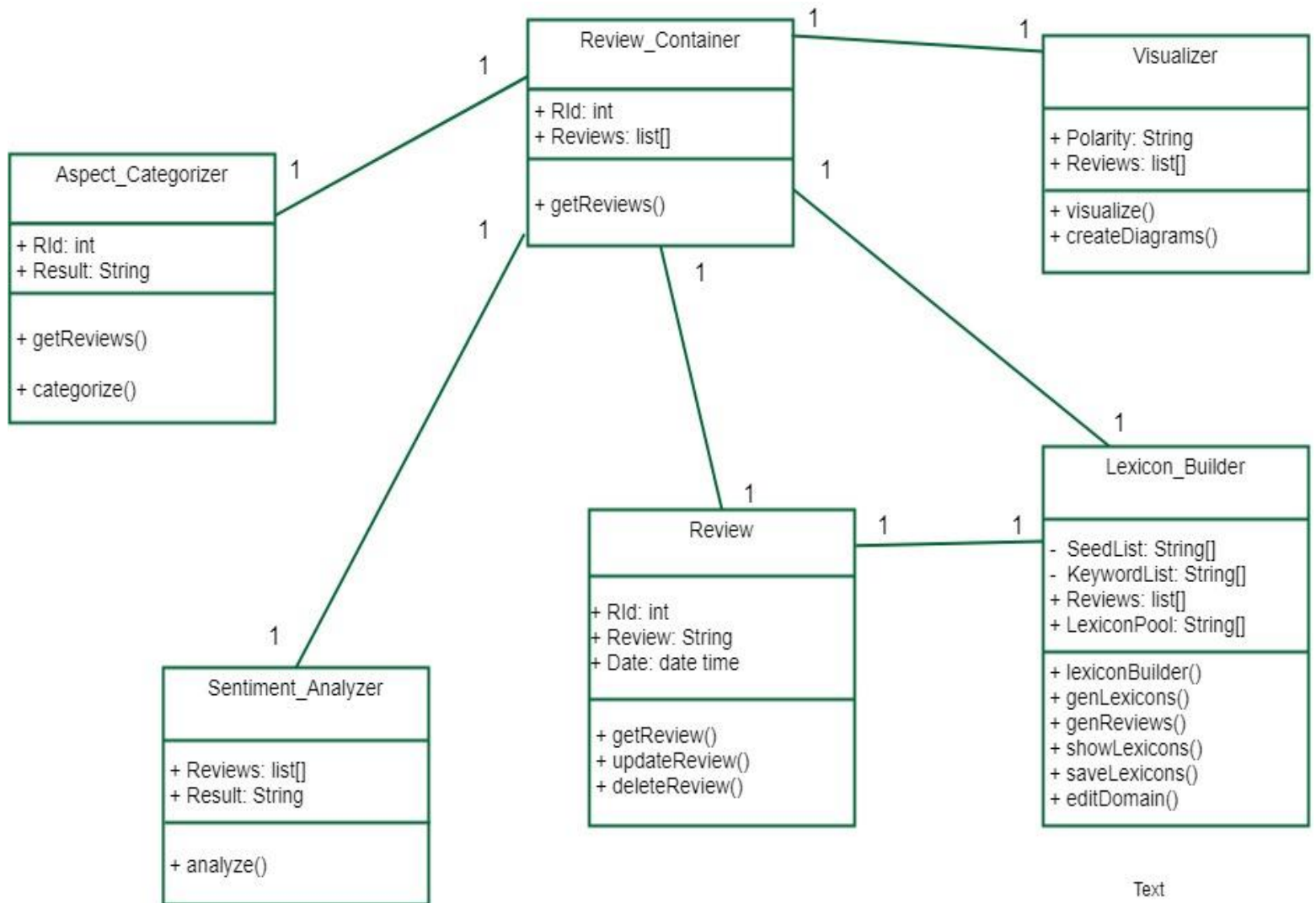


Figure 1. 6 Class diagram

3.3 Performance requirements

- 2GB ram
- Python3
- 500MB from disk space
- Following performance requirements will be handle by the system.
 - Response time
 - Scalability
 - Accuracy of the system

Multiple consumers can have access the same time to the system and can make reviews on products and also can view existing reviews with polarity.

When consumers visit to a product, they can view the aspect-based opinions of other customers as soon as possible without waiting. So, the time consuming is less. Performance is high.

3.4 Design constraints

Load time of the results should be less than 4 seconds.

All data should be stored in a well-organized way

New reviews for products should be analyzed once every two weeks.

Accuracy of domain-based polarity should be greater than 80%.

3.5 Software system attributes

3.5.1 Reliability

System should not be occurring any corruptions during the usage of the system. And also, the possibility of failure should be decreased. When there are any corruptions in showing the output, it is disturbing to the users. So, the system should be Reliable.

The software will meet all the functional requirements without any unexpected behavior

3.5.2 Availability

System should be available at all the time without any breakdown. Data cannot be lost and without any crashes. Aspect-based polarity of reviews will be shown graphically when consumers refer to a product.

3.5.3 Security

Users should prevent from irrelevant reviews of other users and the system should maintain a proper way to provide privacy related security of reviewers.

Data accessibility is defined by administrator for users.

3.5.4 Maintainability

Maintenance of the system is the most essential thing. System cannot be crashed when users are working with that. And also, when some errors occurred or the system is not working properly, to identify those things developed code must be clear. So, when the development of the system, there should be a comments under the each methods of coding. Hence, it can be easy to understanding bugs easily.

4. Supporting requirements

References

- [1] <https://monkeylearn.com/blog/aspect-based-sentiment-analysis/>
- [2] <https://medium.com/@pmin91/aspect-based-opinion-mining-nlp-with-python-a53eb4752800>
- [3] <https://github.com/topics/aspect-based-sentiment-analysis>
- [4] https://www.codementor.io/vishnu_ks/aspect-based-sentiment-analysis-dk3jup75h
- [5] <https://www.kaggle.com/itratrahman/nlp-tutorial-using-python>
- [6] <https://towardsdatascience.com/a-practitioners-guide-to-natural-language-processing-part-i-processing-understanding-text-9f4abfd13e72>
- [7] Naveen Kumar Laskari, Suresh Kumar Sanampudi, “Aspect Based Sentiment Analysis Survey” , Mar-Apr. 2016, [Online]. Available: https://www.researchgate.net/publication/304252751_Aspect_Based_Sentiment_Analysis_Survey
- [8] Swapn Joshi, “Aspect based sentiment analysis for United States of America Airlines”, 11th December 2017, [Online]. Available: https://www.researchgate.net/publication/322958673_Aspect_Based_Sentiment_Analysis_for_United_States_of_America_Airlines

[9] Bo Wang, Min Liu, “ Deep Learning for Aspect-Based Sentiment Analysis”, [Online]. Available:

<https://cs224d.stanford.edu/reports/WangBo.pdf>

[10] Dohaiha, H. H., Prasad, P., Maag, A., & Alsadoon, A. (2018). Deep Learning for Aspect-Based Sentiment Analysis: A Comparative Review. Expert Systems with Applications. doi:10.1016/j.eswa.2018.10.003