ÇANKAYA UNIVERSITY
FACULTY OF ENGINEERING
COMPUTER ENGINEERING DEPARTMENT

CENG 407
Sentiment Analysis of the Feedback from Airplane Passengers

SRS REPORT

Merve Şirin - 201626404
İrem Şahingöz - 201611051

2020-2021

Contents

1. Introduction

There are many comments made by individuals to evaluate a subject, product, or service on social media accounts. Organizations, manufacturers, or customers improve their services or products based on these comments. Customers can have an idea about the product or service by reading these reviews. During the purchasing phase, these comments are of great benefit to the users in the decision-making phase. Sometimes there are few reviews that can be read and evaluated by the client or the company. In some cases, these forums, polls, social media comments contain too many comments. At this point, sentiment analysis comes into play. These comments are categorized with various machine learning methods and classified as positive, negative, and neutral. Thanks to the Sentiment analysis, institutions, organizations, companies, and customers can make this analysis automatically and save time, and ensure correct inferences.

1.1. Purpose

We aim to determine the level of satisfaction of customers by analyzing the comments made by American airport passengers on the Twitter as a result of the service they receive. Using Python codes and machine learning algorithms, we will evaluate the service-taking comments of airport passengers and classify them as positive, negative, and neutral. By classifying the words such as "good, regret, amazing, bad, lovely, terrible" in the text, we will obtain results such as positive, negative, or neutral.

By analyzing the emotions of its passengers with sentiment analysis, it will guide American airports to improve or develop services by airport according to the results of this analysis.

1.2. Scope

In this project, we aim to analyze customer satisfaction based on the comments of American airport passengers. In this way, airport operators will evaluate customer satisfaction without having to read all the comments and direct their services accordingly. It is ensured that businesses can serve their service structure better with correct evaluation results. These evaluation results will be classified as positive, negative, and neutral, helping customers and airport operators make better decisions for their benefit. With this project, it is possible to develop services in the right direction for airports and to make better decisions for customers.

1.3. Glossary

Terminology	Definition
SRS	Software Requirements Specifications
NLP	Natural Language Processing
User	The person who wants to analyze emotions according to tweets written by passengers of American airlines.
Object	An entity that can be a product, service, individual, organization, event, or topic.
Attribute	An object has some attributes: delays, quality of service, etc.
Opinion	A person or organization that expresses a positive or negative sentiment on a particular attribute of an object at a certain time.
Opinion Holder	The person or organization that expresses the opinion.
Opinion Orientation	i.e., polarity, e.g., positive, negative or neutral.
Opinion Strength	Level, scale, intensity of opinion indicating how strong it is, e.g., contented, happy, joyous and ecstatic, whose strength are incremental.

1.4. References

[1] IEEE Std 830™-1998(R2009) Recommended Practice for Software Requirements Specifications.

1.5. Overview

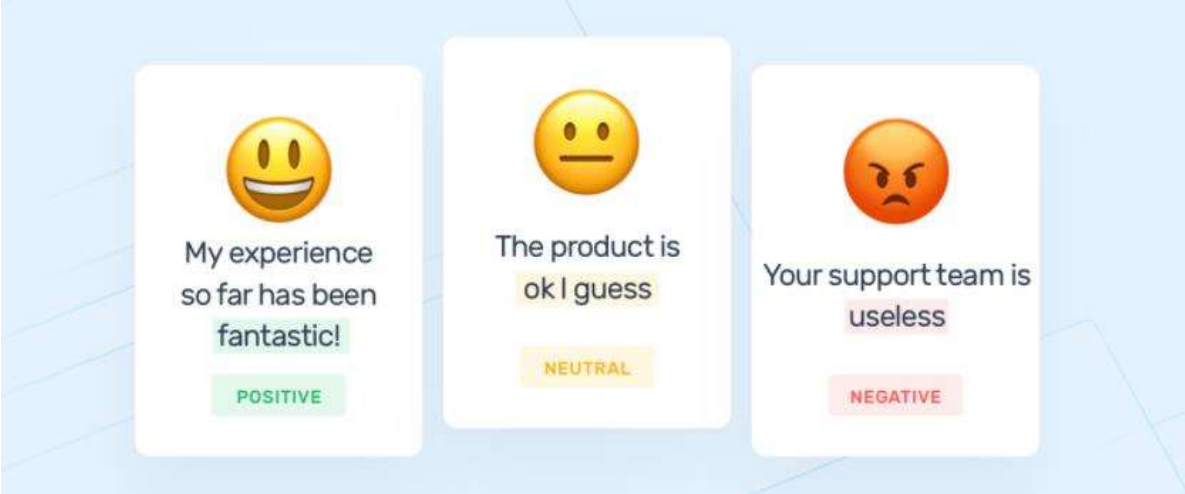
The continuation of this SRS document is as follows: In Section 2, the features of the project we will realize are mentioned for the users. Also, the limitations and risks of the project are detailed in this section. Chapter 3 discusses the requirements of the project we will realize using technical terms. Also in this section, the functions and features of this project are detailed for users. In Chapter 4, the use of the project we will develop is explained by visualizing it with diagrams.

2. Overall Description

In this section, various aspects of the project will be discussed.

2.1. Product Perspective

With sentiment analysis, it can be determined which way of thinking about a certain subject by evaluating people's comments depending on the comments or texts. In this project, the comments made by the American airport passengers about the service they receive on their Twitter accounts will be analyzed and classified. This project provides benefits such as providing a better service to its passengers by improving the services of the airport in line with the results of this analysis by making emotional analysis for the comments received from airport passengers.



2.2. User Characteristics

Sentiment analysis will be determined by opinion orientation according to the opinions expressed about some attributes of an object belonging to an opinion stated by the opinion holder. While doing this analysis, opinion strength is an important factor for a more accurate analysis. Our sentiment analysis is a process that automates the mining of opinions and emotions from comment tweets through NLP. Opinions are classified as positive, negative, and neutral with sentiment analysis. This project aims to analyze people's feelings, attitudes, opinions, and feelings about the service and quality they receive from the American airport.

2.3. Constraints

Individuals' sensitivity to the service or the situation may be affected for different reasons. He can make a negative comment when he's having a bad day and about a topic he is neutral. With sufficiently large data, outliers can be diluted in these many interpretations to obtain more accurate results. It is also important to evaluate these interpretations in terms of time, as the mood of people may change over time according to the events that occur in the world. Context matters when it comes to mockery under the NPL. It is necessary to have a large sample of natural language data that provides clues to determine whether a statement a user uses is ironic or not.

2.4. Risks

The successful implementation of the project depends on careful examination and meeting of some risk conditions. These conditions for our project are listed below:

- Tweeter data should be collected.
- Tweets do not have a specific structure and free/colloquial language is used.
- A tweet may reflect satisfaction and/or dissatisfaction at the same time.

2.5. Assumptions and Dependencies

Emotion analysis for service comments is to obtain service characteristics and sensitivity based on the feature based on the assumption that these comment texts contain a hidden feature and emotional structure. In our project is assumed that sensitivity based on the context of the service feature can be captured by learning the corresponding features by word marking. Sentiment analysis suitable for the subject is based on controlled text classification, which can be used with decision trees and polarity dictionaries, to determine the correct connection points for words containing polarity and to determine the relevant dependencies in each sentence in the comments.

3. Requirements

3.1. Specific Requirements

3.1.1. User Interface

There is no user interfaces for this project since there is no website or application will be developed.

3.1.2. Hardware Interface

Hardware interface is not needed for this project.

3.1.3. Software Interface

Twitter API will be used to analyse tweets in twitter. For high accuracy, we need big data so the system needs libraries such as Keras, Tensorflow.

3.1.4. Communication Interface

Communication interface is not needed for this project.

3.2. Functional Requirements

- Twitter data should be collected
- System should be able to process new tweets stored in twitter's database.
- System should be able to analyze data and classify tweets' polarity

3.3. Non-Functional Requirements

Non-functional requirements require the following:

- It should use big dataset for high accuracy
- Tweets do not have a specific structure and free/colloquial language is used
- It should be user friendly.
- It should provide efficient analysis process in very large data in a short time.
- It should provide the high accuracy and most reliable feedback.

3.4. Software System Attributes

The system need to Windows 7 or higher operating system. Also, the system requires Text Editor, 8GB or more RAM and 2.5 GHz or more speed is needed for this project.

3.4.1. Reliability

We will safely shoot our target through the Twitter API.

3.4.2. Avaibility

We will work with big data to ensure high efficiency. We will make use of Keras, Tensorflow open libraries for this.

3.4.3. Performance

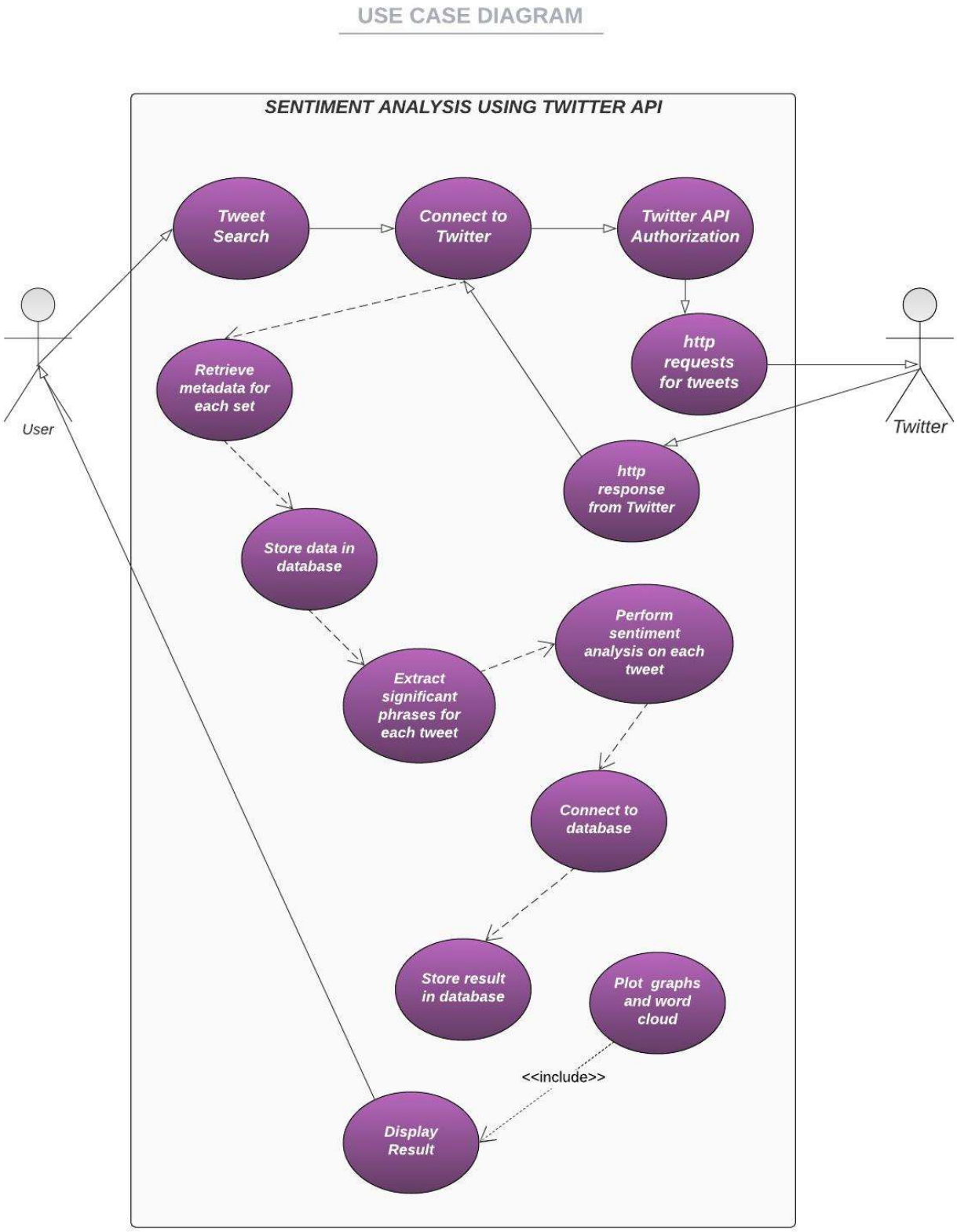
It should provide the high accuracy and most reliable feedback.

3.4.4. Portability

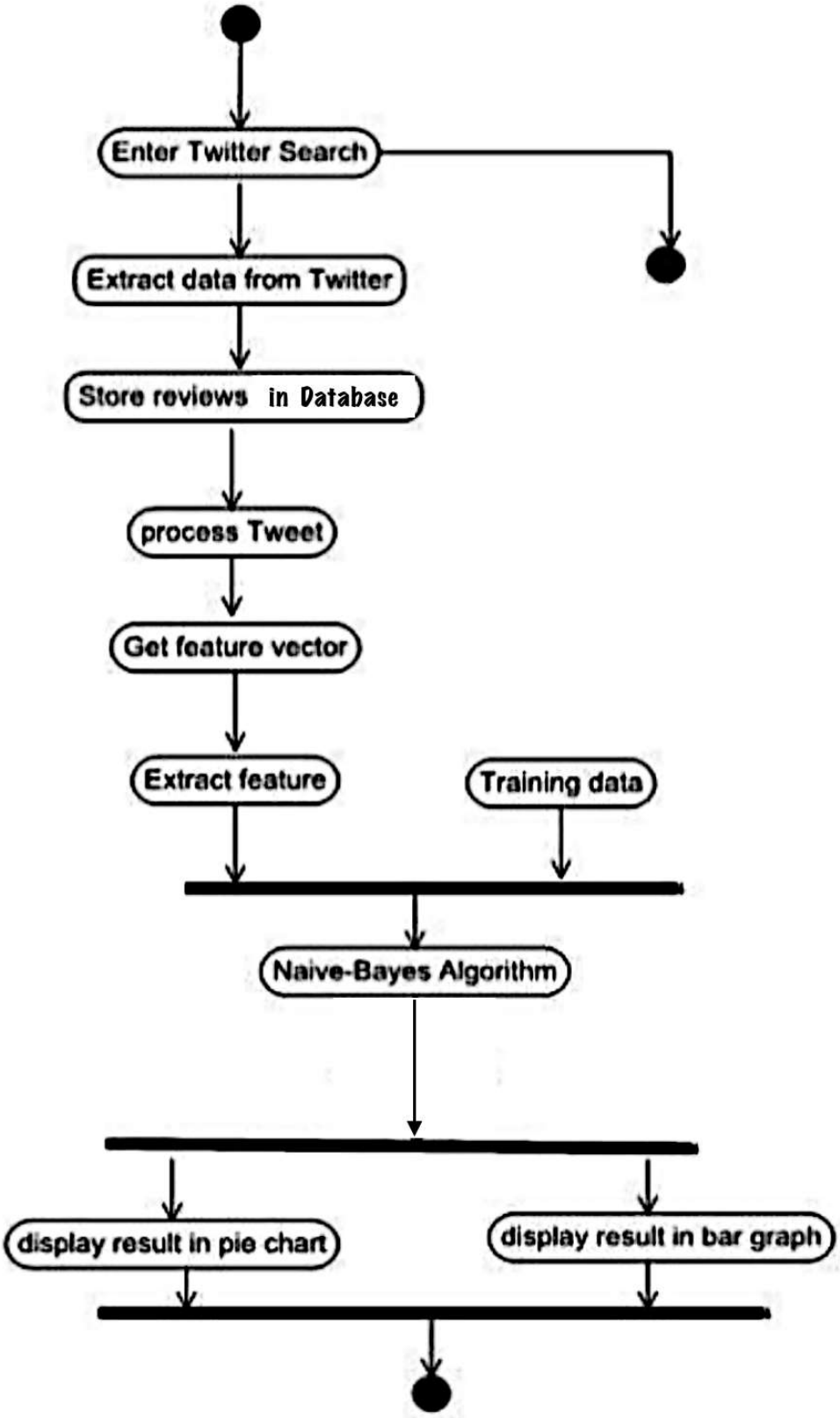
The system will be useful in both Linux and Windows.

3.5. UML Analysis Model

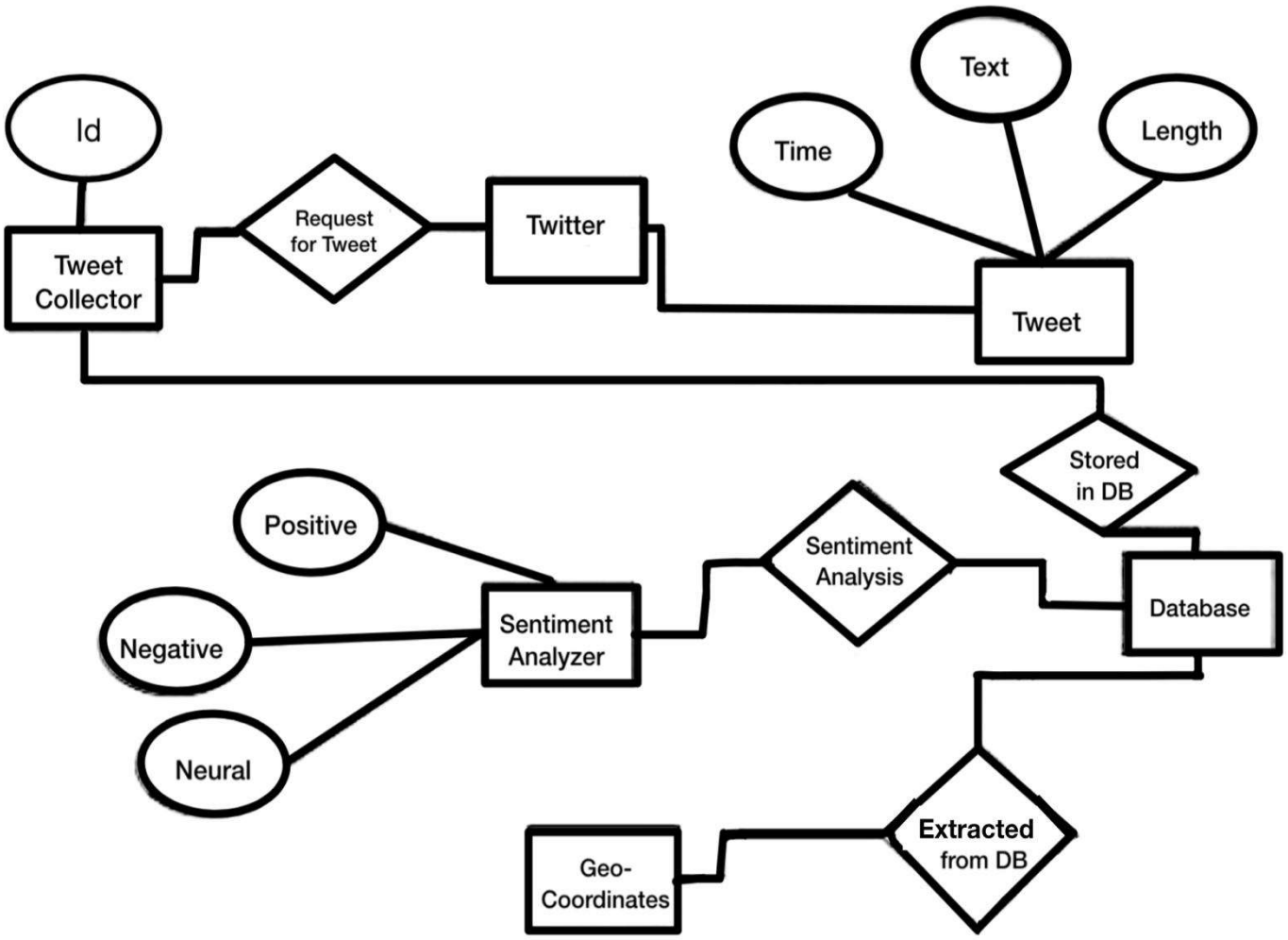
3.5.1. User Case Diagram



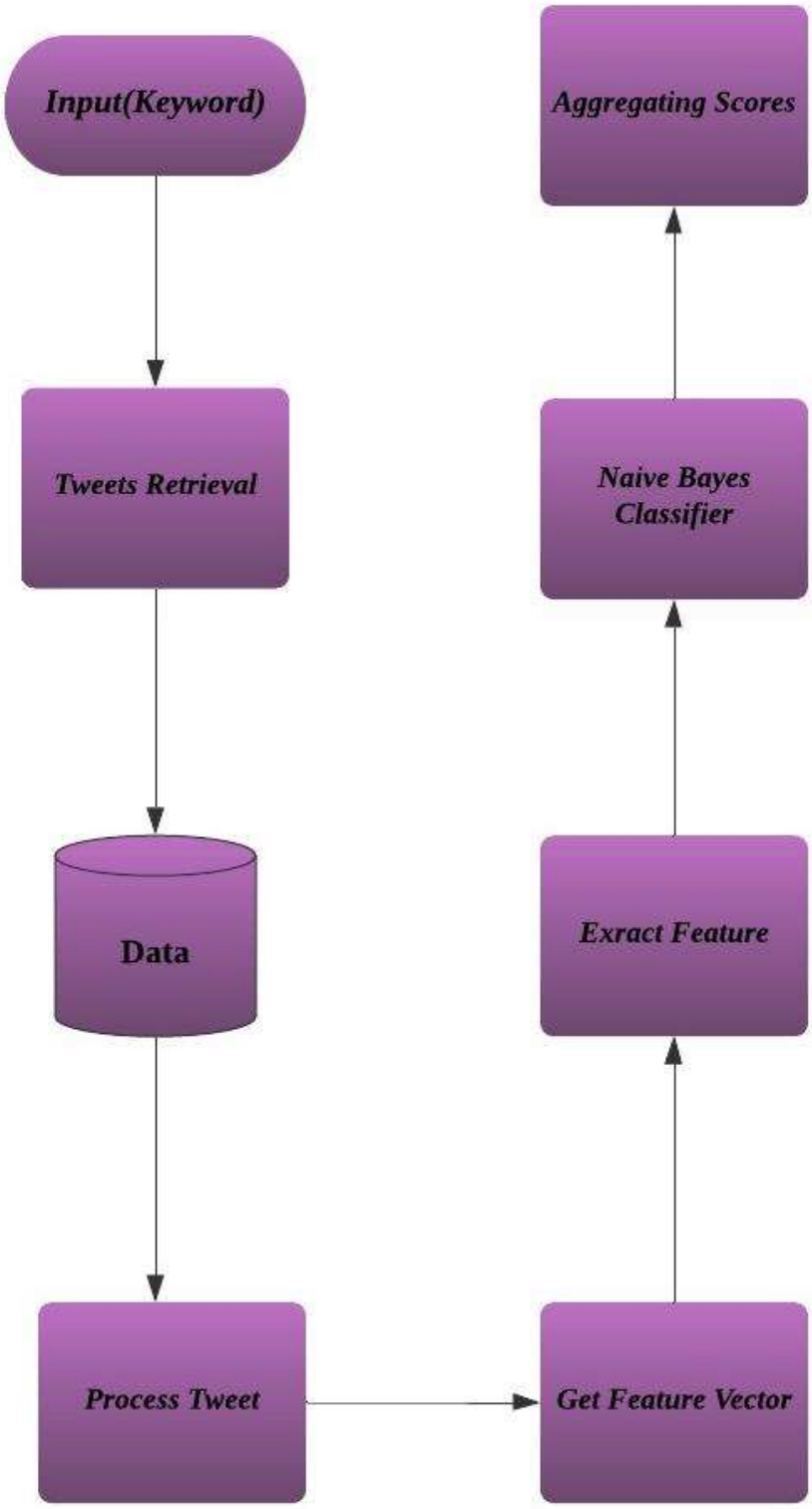
3.5.2. State Chart Diagram



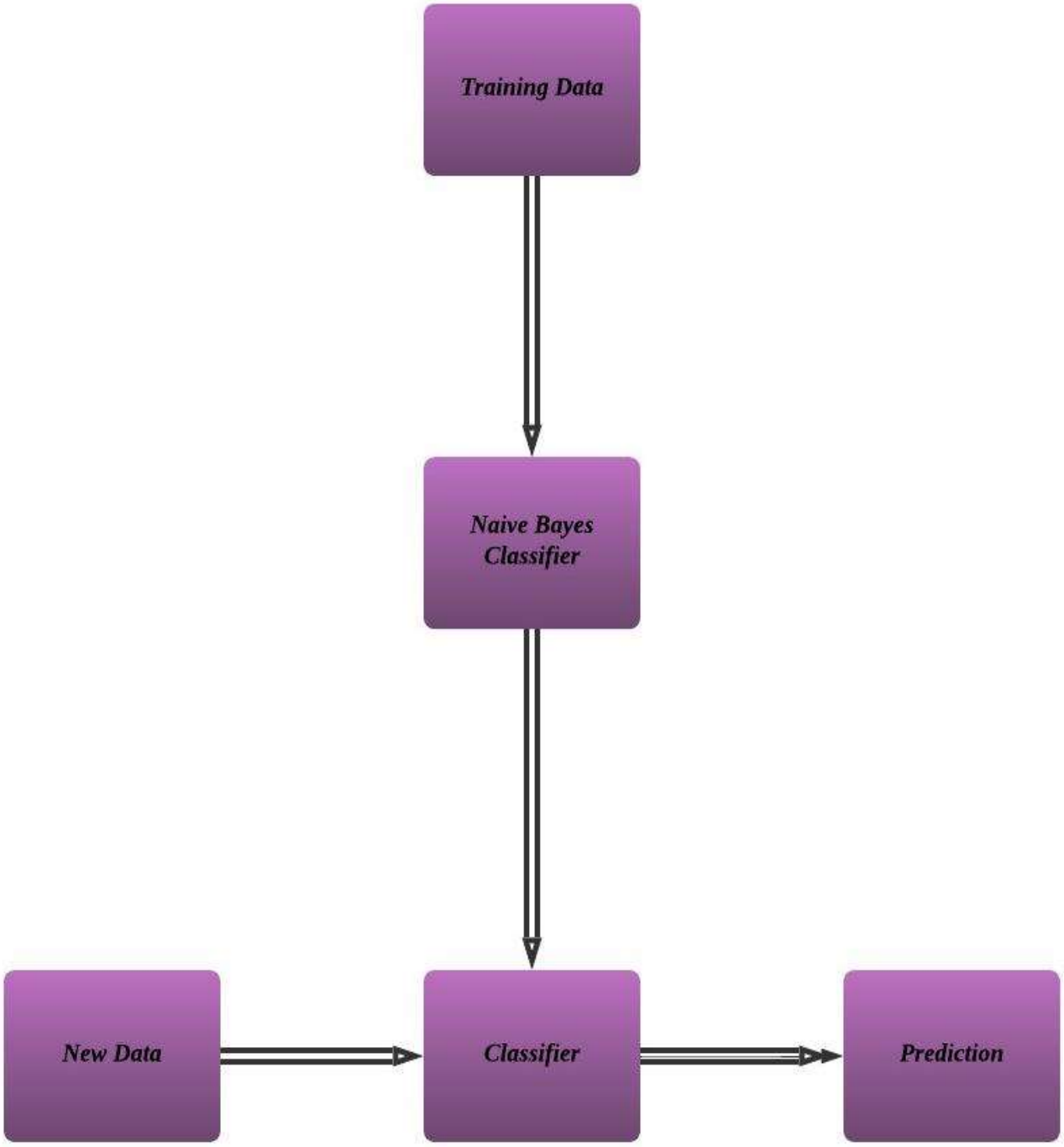
3.5.3. ER Diagram



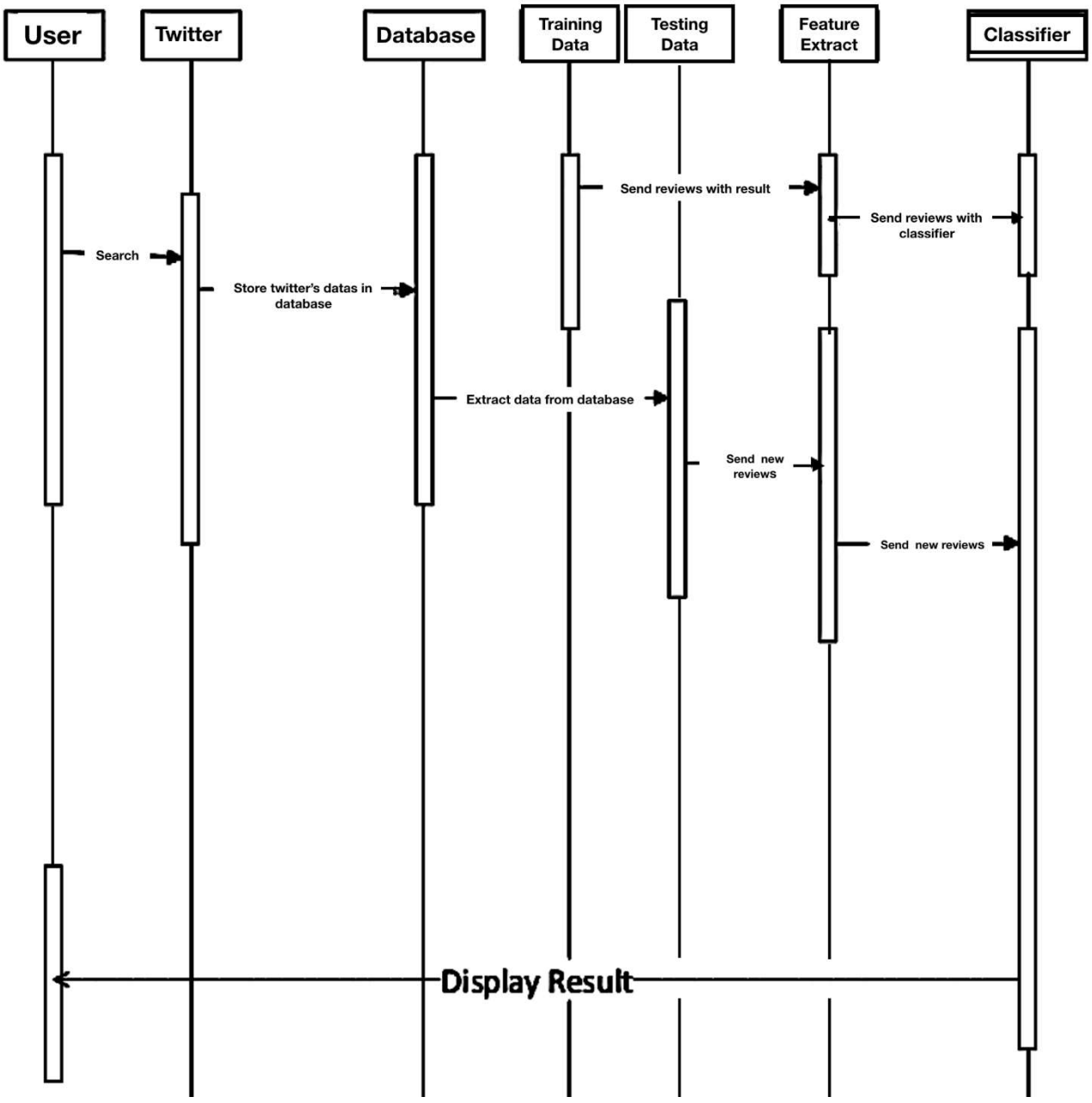
3.5.4. Data Flow Diagram



3.5.5. System Flow Diagram



3.5.6. Sequence Diagram



4. References:

[1] Bing Liu. Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers, May 2012.

[2] Zhai, Zhongwu, Bing Liu, Hua Xu, and Peifa Jia. Clustering Product Features for Opinion Mining. in Proceedings of ACM International Conference on Web Search and Data Mining (WSDM-2011). 2011. 385. Zhai, Zhongwu, Bing Liu, Hua Xu, and Peifa Jia. Constrained LDA for Grouping Product Features in Opinion Mining. in Proceedings of PAKDD-2011. 2011.

[3] Zhang, Lei and Bing Liu. Extracting Resource Terms for Sentiment Analysis. in Proceedings of IJCNLP-2011. 2011a.

[4] Liu, Bing. Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data, 2006 and 2011: Springer.

[5] 296. Sarawagi, Sunita. Information extraction. Foundations and Trends in Databases, 2008. 1(3): p. 261-377.