



**Tribhuvan University**  
**Institute of Science and Technology**  
**School of Mathematical Sciences**

## **SENTIMENT ANALYSIS ON EDUCATION SYSTEM**

**A Project Report**

**Submitted to**

**Office of the Dean**

**Institute of Science and Technology, Tribhuvan University**

**in**

**Partial Fulfillment of the Requirements for the  
Bachelors in Computer Science and Information Technology**

**Submitted By**

**Arjun Sherpa (T.U. Roll No. 20983/075)**

**Neupal Rai (T.U. Roll No. 20998/075)**

**Pema Gyalbu Lama (T.U. Roll No. 21000/075)**

**April 2023**



**Tribhuvan University**  
**Institute of Science and Technology**  
**School of Mathematical Sciences**

**DECLARATION**

I hereby declare that the project work entitled “Sentiment Analysis on Education Syatem” submitted to the Institute of Science and Technology, Tribhuvan University, is an original piece of work done in the form of college project for Bachelor in Computer Science and Information Technology Program under the supervision of Mr. Surya Bam.

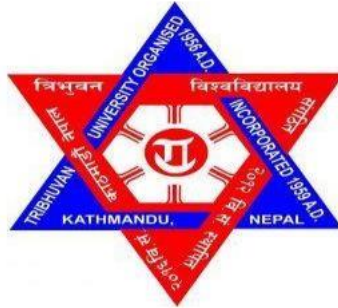
Signatures of candidate:

Arjun Sherpa

Neupal Rai

Pema Gyalbu Lama

Date: April 2022



**Tribhuvan University**  
**Institute of Science and Technology**  
**Asian School of Management and Technology**

## **LETTER OF RECOMMENDATION**

The project work entitled “Sentiment Analysis on Education System” submitted by Arjun Sherpa, Neupal Rai, Pema Gyalbu Lama of Asian School of Management and Technology, Gongabu, Kathmandu, is prepared under my supervision as per the procedure and format requirements laid by Institute of Science and Technology, Tribhuvan University, as partial fulfillment of the requirement for the degree of Bachelor of Computer Science and Information Technology. I, therefore, recommend their project work report is satisfactory to process for the future evaluation.

.....

Mr. Surya Bam

Date: April 2022



**Tribhuvan University**  
**Institute of Science and Technology**  
**Asian School of Management and Technology**

**LETTER OF APPROVAL**

This is to certify that the project work entitled “Sentiment Analysis on Education System” done by team of Arjun Sherpa, Neupal Rai, Pema Gyalbu Lama is satisfactory in the scope and generality as a project in the partial fulfillment of the requirement for Computer Science and Information Technology.

Evaluation Committee

.....

Mr. Surya Bam

Supervisor

.....

Mr. External Examiner

.....

Mr. Program Coordinator

## **ACKNOWLEDGEMENT**

We would like to express our deepest appreciation to all those who provided us the possibility to complete this project. A special gratitude we give to our final year project Supervisor, Mr.Surya Bam, whose suggestions and encouragement, helped us to coordinate our project especially in writing this report.

Furthermore, we would also like to acknowledge with much appreciation the crucial role of the staff of Asian College, who gave the permission to use all required equipment and the necessary materials to complete the task. We are thankful and fortunate enough to get constant support from our seniors and every teaching staff of B.Sc. CSIT department which helped us successfully complete our project.

We would also like to extend our regards to all the non-teaching staff of B.Sc. CSIT department for their timely support. We have to appreciate the guidance given by other supervisors as well as the panels especially in our project presentation that has improved our presentation skills thanks to their comment and advice. Our thanks and appreciations also go to each and every one of our colleagues for their encouragement and support in developing the project.

With respect,

Arjun Sherpa

Neupal Rai

Pema Gyalbu Lama

April 2022

## **ABSTRACT**

The topic of the project was “Sentiment Analysis on Education System”. The main objective of the project is to understand the sentiment of student opinions. It would be a great help for student to present their views for the improvement of education system. Our project is based on the web platform, our project works in real time and can provide the result with more accuracy. Our project classify the given sentiments on the basis of trained dataset. Sentiment analysis is based on machine learning where the machine has to be trained at first.

## **TABLE OF CONTENTS**

| <b>Content</b>                           | <b>Page</b> |
|------------------------------------------|-------------|
| <b>DECLARATION</b>                       | <b>ii</b>   |
| <b>LETTER OF RECOMMENDATION</b>          | <b>iii</b>  |
| <b>LETTER OF APPROVAL</b>                | <b>Iv</b>   |
| <b>ACKNOWLEDGEMENT</b>                   | <b>v</b>    |
| <b>ABSTRACT</b>                          | <b>vi</b>   |
| <b>TABLE OF CONTENTS</b>                 | <b>vii</b>  |
| <b>LIST OF TABLES</b>                    | <b>ix</b>   |
| <b>LIST OF FIGURES</b>                   | <b>ix</b>   |
| <b>LIST OF ACRONYMS /ABBREVIATIONS</b>   | <b>x</b>    |
| <br>                                     |             |
| <b>CHAPTER I: INTRODUCTION</b>           | <b>1-4</b>  |
| 1.1 Background of the Study              | 1-2         |
| 1.2 Statement of the Problem             | 2           |
| 1.3 Objectives of the Study              | 2           |
| 1.4 Scope and Limitation                 | 3           |
| 1.5 Development Methodology              | 3           |
| 1.6 Report Organization                  | 3-4         |
| <br>                                     |             |
| <b>CHAPTER II: LITERATURE REVIEW</b>     | <b>5-7</b>  |
| 2.1 Background Study                     | 5           |
| 2.2 Literature Review                    | 6-7         |
| <br>                                     |             |
| <b>CHAPTER III: SYSTEM ANALYSIS 8-14</b> |             |
| 3.1 System Analysis                      | 8-14        |
| 3.1.1 Requirement Analysis               | 8-9         |
| 3.1.2 Feasibility Analysis               | 9-11        |

|                                                         |              |
|---------------------------------------------------------|--------------|
| 3.1.3 Analysis                                          | 11-14        |
| <b>CHAPTER IV: SYSTEM DESIGN</b>                        | <b>15-22</b> |
| 4.1 Design                                              | 15-16        |
| 4.1.1 System Architecture and Overview                  | 16-17        |
| 4.2.2 Interface and Dialogue Design                     | 17-18        |
| 4.2 Algorithm Used                                      | 18-22        |
| 4.2.1 Naïve Bayes Algorithm                             | 17-18        |
| <b>CHAPTER V: IMPLEMENTATION AND TESTING</b>            | <b>23-26</b> |
| 5.1 Implementation                                      | 23-24        |
| 5.1.1 Tools Used                                        | 24           |
| 5.2 Testing                                             | 24-26        |
| 5.2.1 Unit Testing                                      | 25           |
| 5.2.2 System Testing                                    | 26           |
| <b>CHAPTER VI: CONCLUSION AND FUTURE RECOMMENDATION</b> | <b>27</b>    |
| 6.1 Conclusion                                          | 27           |
| 6.2 Future Recommendation                               | 27           |
| <b>REFERENCES</b>                                       | <b>28</b>    |



# **Chapter 1**

## **INTRODUCTION**

### **1.1 Background of the study**

Sentiment analysis is the process of using Natural Language Processing in order to detect whether the opinions can be positive, negative or neutral on text. However, manually analyzing the sentiment of text and reviews are impossible because there is large amount of students opinions. Nowadays, social media has become one of the most effective communication media available. As a result, a large amount of data can be collected and sentiment analysis can be done to analyze the opinions of student on education system effectively and efficiently. Sentiment analysis is critical because it need to understand the emotions and sentiment of students. Using sentiment analysis, computers can automatically process text opinions and understand it just as human does.

Student feedback is important because it can helps the college to understand students learning behaviour. Sometimes students do not understand what lecturer is trying to express thus by providing feedback, student can indicate this to lecturer. Student feedback can also help in understanding different issues that students have including the students not understanding the education standard. The process of sentiment analysis on education system reviews involves data collection to collect data, preprocessing to remove some non-meaningful words, feature extraction and use of Naïve Bayes Algorithm.

Sentiment Analysis on Education System Reviews helps to make improvement in teaching techniques. The main goal of our project is to collect the student's opinion on Education System and use it to obtain the meaningful information and semantics from text determining student's opinions.

### **1.2 Statement of the problem**

Education system review is one of the most important factors to improve the teaching environment between teachers and students. There is a lack of interaction between students and teacher on education system review, it is because teachers usually come to college to teach whereas students come to learn something new from

teachers. There is no any mediator to listen the reviews of education system provided by students.

Students are come from different cultural background and it is important to collect the reviews of students on education system. It may be difficult or impossible to understand the sentiments of all student opinions on education system manually. It may takes a lot of time to manually take the students opinions on education system, further it takes time to understand the sentiments of students opinions. Our project helps to understand the opinions of students on education system and generate the output of the student opinions on education system. Our system helps to understand the sentiments of students within a short period of time.

### **1.3 Objectives of the study**

The main objectives of our project are:-

- To effectively and accurately classify the sentiments collecting student opinions on education system.
- To understand the feedback and reviews of the students in a large number without any human interference.

### **1.4 Scope and limitation**

Sentiment analysis is a uniquely powerful tool for education system that are looking to measure attitudes, feelings and opinions regarding their teaching methodology in the organization. By investigating and analyzing students sentiments, organization are able to understand the overall feedback of the students and ultimately better serve their students with the best styles, services and experiences they offer to improve their learning quality. The future of sentiment analysis is going to continue in many business organizations too.

Sentiment analysis is one of the hardest tasks in natural language processing because even humans struggle to analyze sentiments accurately. Data scientists are getting better at creating more accurate sentiment classifiers, but there's still a long way to go. Some of the main limitations of sentiment analysis are:-

1. Incorrectly targeted sentiments

2. Subjectivity and Tone
3. Context and Polarity
4. Sarcasm
5. Comparisons
6. Defining Neutral
7. Human Annotator Accuracy

## 1.5 Development Methodology

To develop our system successfully, we used an agile development methodology based on iterative and incremental processes. During the development of our project, we divided our project workflow to a different phases with each sprint to be completed within a particular given period of time. While developing our project, we manage by involving constant collaboration within team and working in iterations.

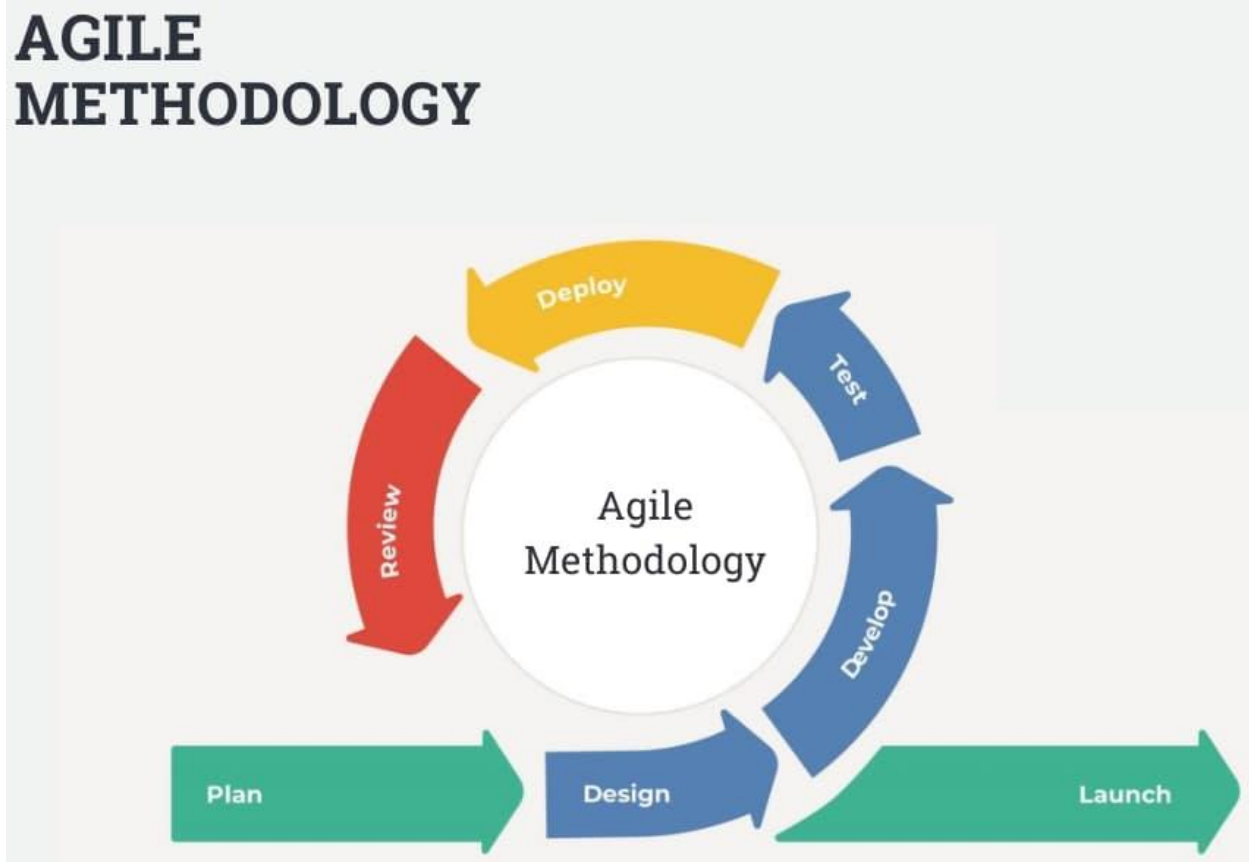


Figure: Agile Methodology

## **1.6 Report Organization**

The Report Organization includes the contents about how the study is being organized and carried out. The report is being divided into 6 chapters and the headings of the chapter includes:

- Chapter one contains the background of the study, statement of the problem, objectives of the study, scope and limitations of the study, development methodology and organization of the study.
- Chapter two contains background study and review of related literature.
- Chapter three contains system analysis including requirement analysis and feasibility study, the way of data modeling, process modeling of the system
- Chapter four includes design of the system including system architecture, database design, interface design, Algorithm used for designing system.
- Chapter five contains implementation of the system using different tools and finally testing of the system as a unit and as a whole.
- Chapter six contains conclusions and future recommendations from the study.

## **Chapter 2**

### **LITERATURE REVIEW**

#### **2.1 Background Study**

Student feedback system is one of the main problem which is ignored by most of the educational organization. Sentiment analysis on education system studies the student's feelings, opinions, attitude and assessments towards educational system of organization. The main aim of sentiment analysis is to detect positive, negative or neutral feelings from text provided by students.

#### **2.2 Literature Review**

The research on sentiment analysis on product reviews are still going on in this world. It is improving and updating regularly by integrating the new features with more accuracy. Many successful research had been done on this topic.

Some of the research on this field are done by Esparza et al. [1] utilized sentiment analysis to investigate the views of students on teacher performance, and they discovered that the difficulty of assessing teacher evaluation feedback could be solved with the use of social mining. Wan Y. and Gao. [2] who applied different classification methods such as NB, SVM, decision tree and random forest algorithms. They experimented with six classification techniques and the given ensemble approach was trained and tested using data set of 12864 tweets with 10 folds evaluation. The performance of proposed ensemble approach gave an output of these individual classifiers in the airline service Twitter dataset.

Similarly, Rajput et al. [3] applied sentiment analysis to examine the textual teacher evaluation provided by students. The study finding's highlighted that the classification of student's feedback could enhance teaching and annual evaluation process quality. In this way, Iskander and Bati [4] applied Sentiment Analysis to compare universities rankings from the university and student's social media posts and comments like Twitter. Besides, Abdelrazeq et al. [5] used Sentiment Analysis to analyze Twitter data (16488 tweets) as source of information for evaluating universities performance.

Another study by Balanchandran and Kirupananda [6] used Sentiment Analysis techniques StandfordCoreNLP to help students find the right HE institution. Also

Santos et al. [7] carried out a study to investigate international student's online review about their HE institutions based on social platform. Furthermore, Aung and Myo [8] and Kaewyong et al. [9] used lexicon based Sentiment Analysis to identify student's positive and negative opinions for predicting teacher performance and evaluating quality of teaching process. In their approach, any teacher's opinion level score is automatically extracted and evaluated from student's feedback comments. The student's feedback opinions were automatically classified into strongly positive, moderately positive, strongly negative, moderately negative, weakly negative, weakly positive or neutral category based on two lexicons. In the same sector, Altrabshes et al. [10] used Sentiment Analysis methods to analyze student's feedback to identify their positive or negative feelings towards teaching process. Similarly, by Kouloumpis E., Wilson T and Moree J. evaluate the usefulness of existing lexical resources as well as features that capture information about informal and creative language used in microblogging.

## **Chapter 3**

### **SYSTEM ANALYSIS**

#### **3.1 System Analysis**

System analysis includes evaluation of a system to achieve certain objectives during the development phase. Including analyzing its functional requirement which is an overview of how a system operates or should behave in any particular situation along with non-functional requirements regarding its usability, availability as well reliability. System analysis also emphasizes feasibility analysis to determine if a system is feasible or not. It basically includes looking at the wider system and figuring out how the system works in order to achieve a specific goal.

##### **3.1.1 Requirement Analysis**

In this project, we will focused on rule-based sentiment analysis algorithm. With the use of this algorithm, we can automatically perform sentiment analysis based on set of manually crafted rules. This algorithm include various NLP techniques such as stemming, tokenization, parsing, lexicons and part of speech tagging. More advanced techniques can be used and new ruled can be added to support new expression and vocabulary. However, adding new rules may affect previous results and may affects whole system which makes it complex. So, it require fine tuning and regular investments.

There are two types of requirements:

- Functional Requirement
- Non Functional Requirement

##### **3.1.1.1 Functional Requirement**

Our system can classify the opinions of student at the real time. Our system is able to handle a large amount of opinions collected from students. Students can easily submit their thoughts through the online form. The output of our system is user friendly. Sentiments generated by our system is easy to evaluate.

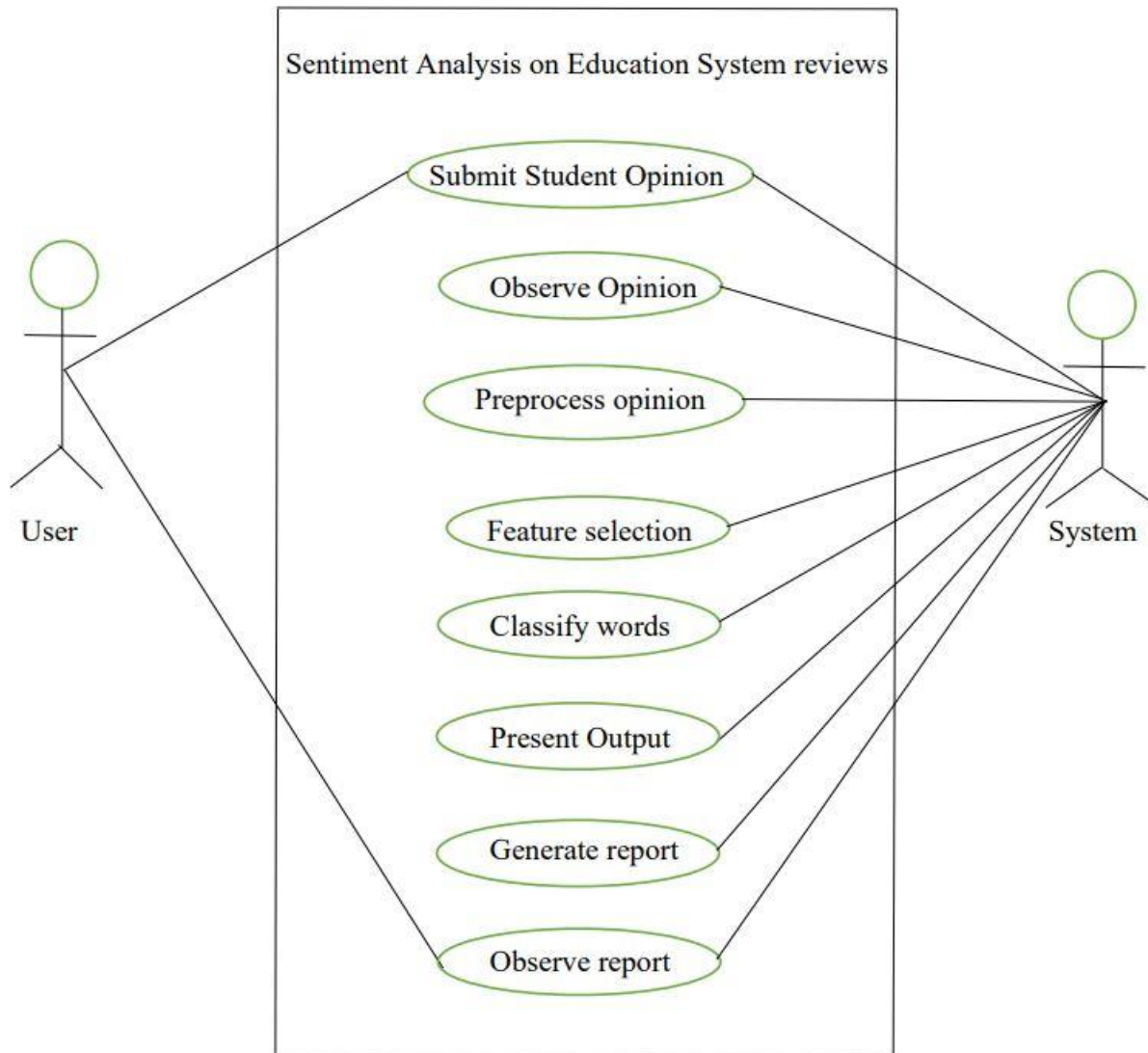


Figure: Use case diagram of sentiment analysis on Education system reviews.

### 3.1.1.2 Non-Functional Requirement

A non-functional requirement is the software testing technique that verifies the attribute of system such as speed, performance and robustness of system. The non-functional requirements of our project are given below:



- Performance

Our system can provide the required output correctly at real time.

- Security

Our system can protect the data and opinions of students in secure way. There will be no chance of leakage of students opinion.

- Availability

The collected data are available that can be used to perform sentiment analysis by our system.

- Usability

Our system is easy to use. It is user friendly. Anyone who have the knowledge about technology can operate our system.

### **3.1.2 Feasibility Analysis**

Feasibility study helps to understand the system before installation, it is beneficial or not. Our system is feasible from 4 perspectives. They are explained below:

#### **3.1.2.1 Technical Feasibility**

Our system is technically feasible because it can operate within 8GB RAM that is now available in most of the computer. So, there is no need to have complex integration with system hardware components. Our system is mainly based on online that use the Naïve Bayes algorithm. It will be upgraded regularly by improving the accuracy of output of customer feedback on product review.

#### **3.1.2.2 Operational Feasibility**

Nowadays, students are familiar with technologies and they can give their opinions through online. Our system can operate with average memory space and processor. Our system can predict the output at real time. The manpower who has the knowledge about technology can operate our system properly.

### 3.1.2.3 Economic Feasibility

By conducting the survey, we can get extra fund from organization. It does not require any experts to operate our system. So, we can save our budget to conduct our project survey. If it is successfully implemented then there will be high chance of purchasing our system by business parties to conduct their survey. Directly, it will be profitable for our team. In this way, it is economically feasible.

### 3.1.2.4 Schedule Feasibility

We have to build our system completely within a three months interval of time. In three months interval of time, we have separate the duration of time for the tasks which requires at the time of development.

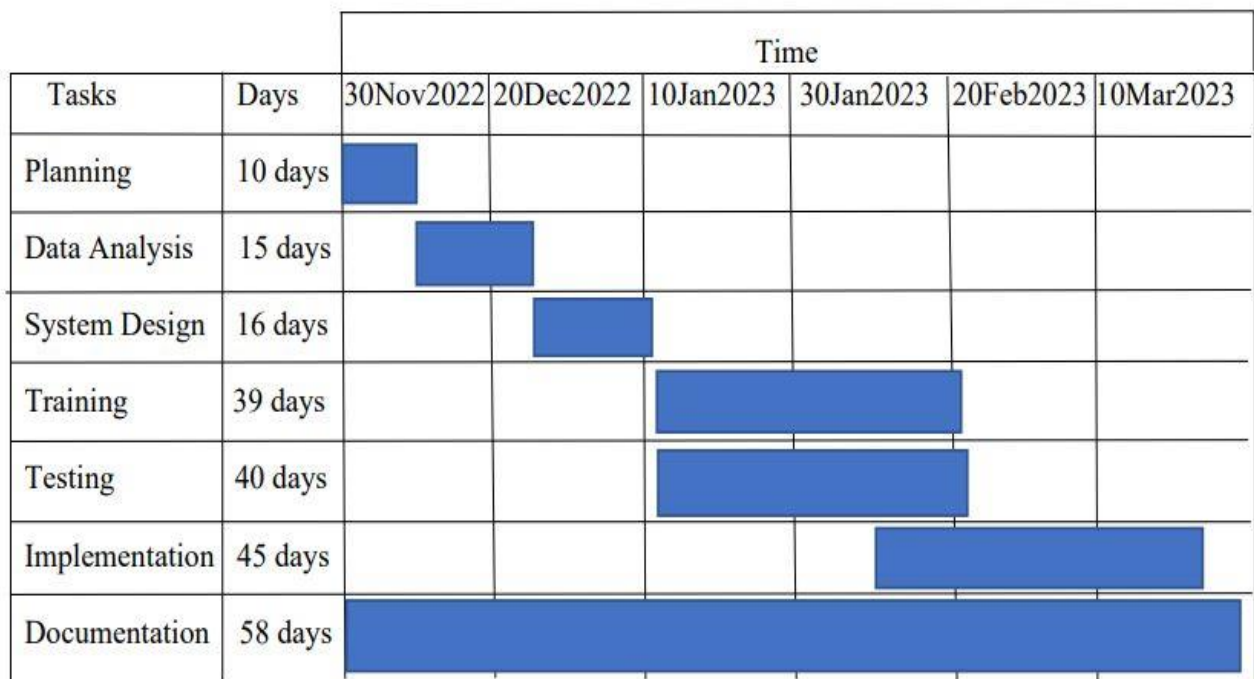


Figure: Gantt chart on sentiment analysis of students' opinions on education system.

### 3.1.2 Analysis

System is based on a structured approach which focuses on modeling data using Entity-Relation diagrams, representing system activities or system flow using flowchart and process modeling using data flow diagrams. It helps to make the proper plan and assist to achieve the desired goals.

#### 3.1.2.1 Data modeling using ER diagram

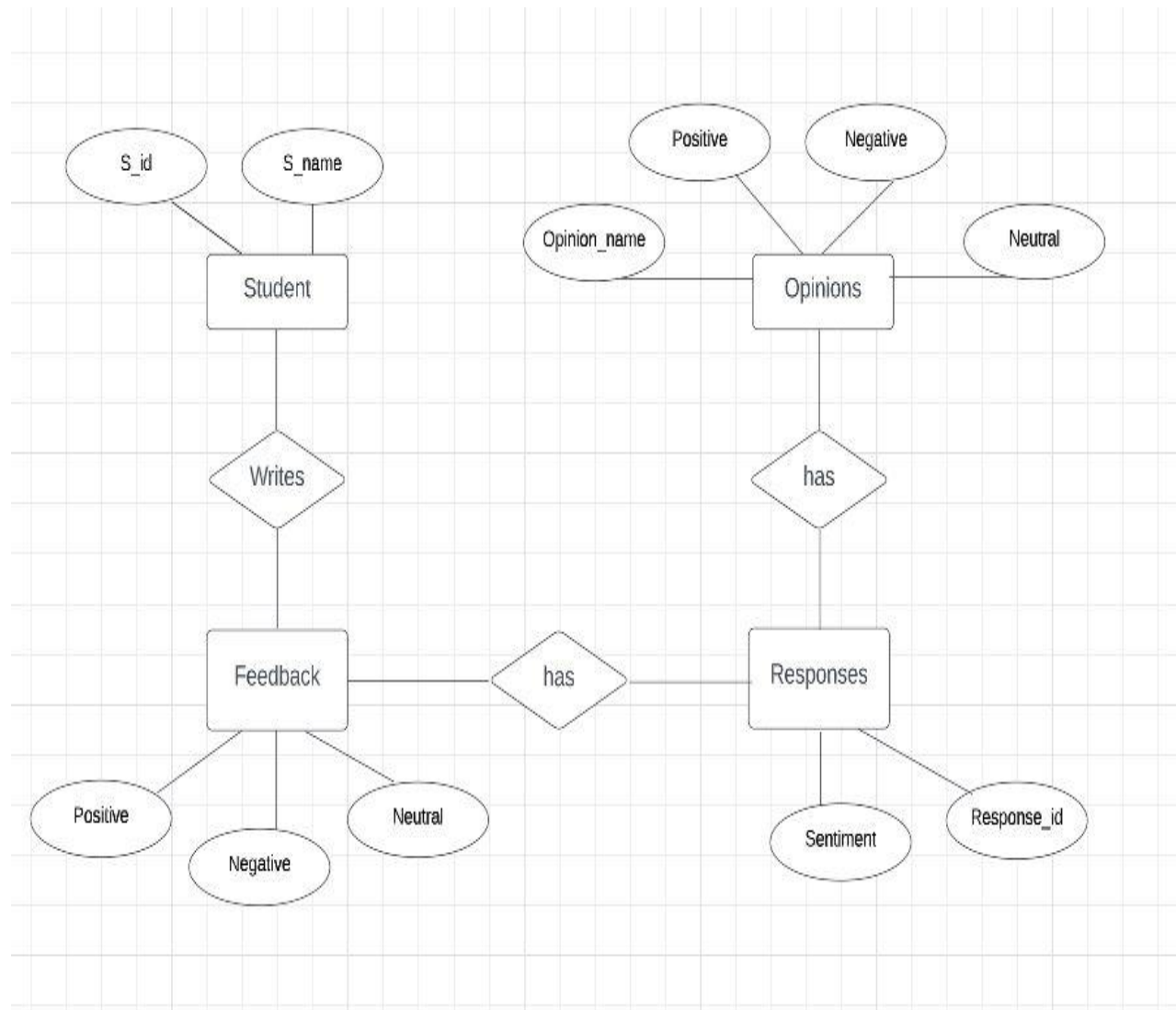


Figure 3.3: ER Diagram of Sentiment Analysis on student feedback.

### 3.1.3 Data Modelling

Data modeling is the process of diagramming data flows. When creating a new or alternate database structure, the designer starts with a diagram of how data will flow into and out of the database. This flow diagram is used to define the characteristics of the data formats, structures, and database handling functions to efficiently support the data flow requirements. After the database has been built and deployed, the data model lives on to become the documentation and justification for why the database exists and how the data flows were designed.

#### 3.1.3.1 Class Diagram

A class diagram in the Unified Modeling Language is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations and the relationships among objects.

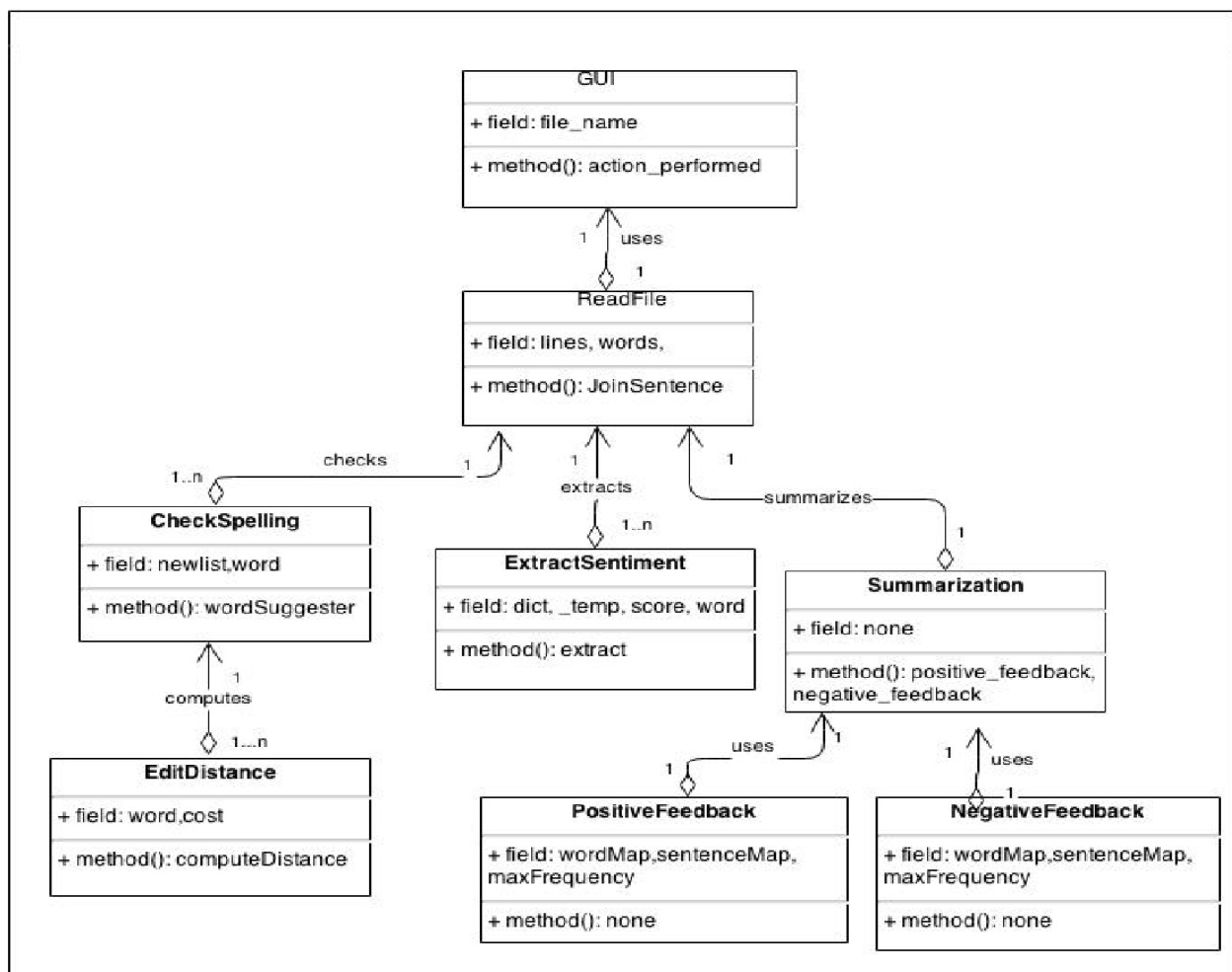


Figure: Class diagram of sentiment analysis.

### 3.1.4 Process Modelling

Process modeling is the graphical representation of business processes or workflows. Like a flow chart, individual steps of the process are drawn out so there is an end-to-end overview of the tasks in the process within the context of the business environment.

#### 3.1.4.1 Sequence diagram

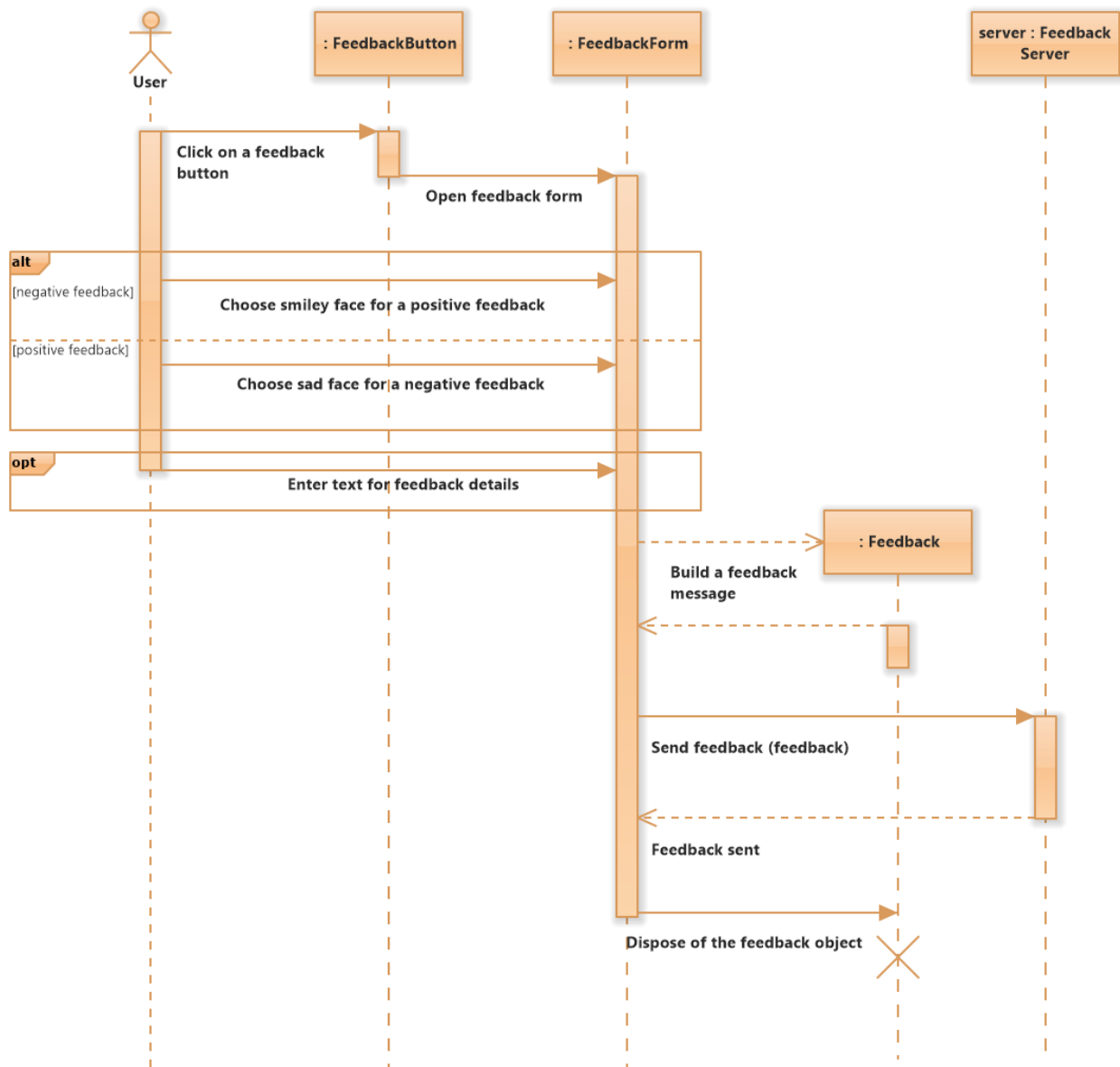


Figure: Sequence diagram of sentiment analysis.

# CHAPTER 4

## SYSTEM DESIGN

### 4.1 System Architecture

A system architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system.

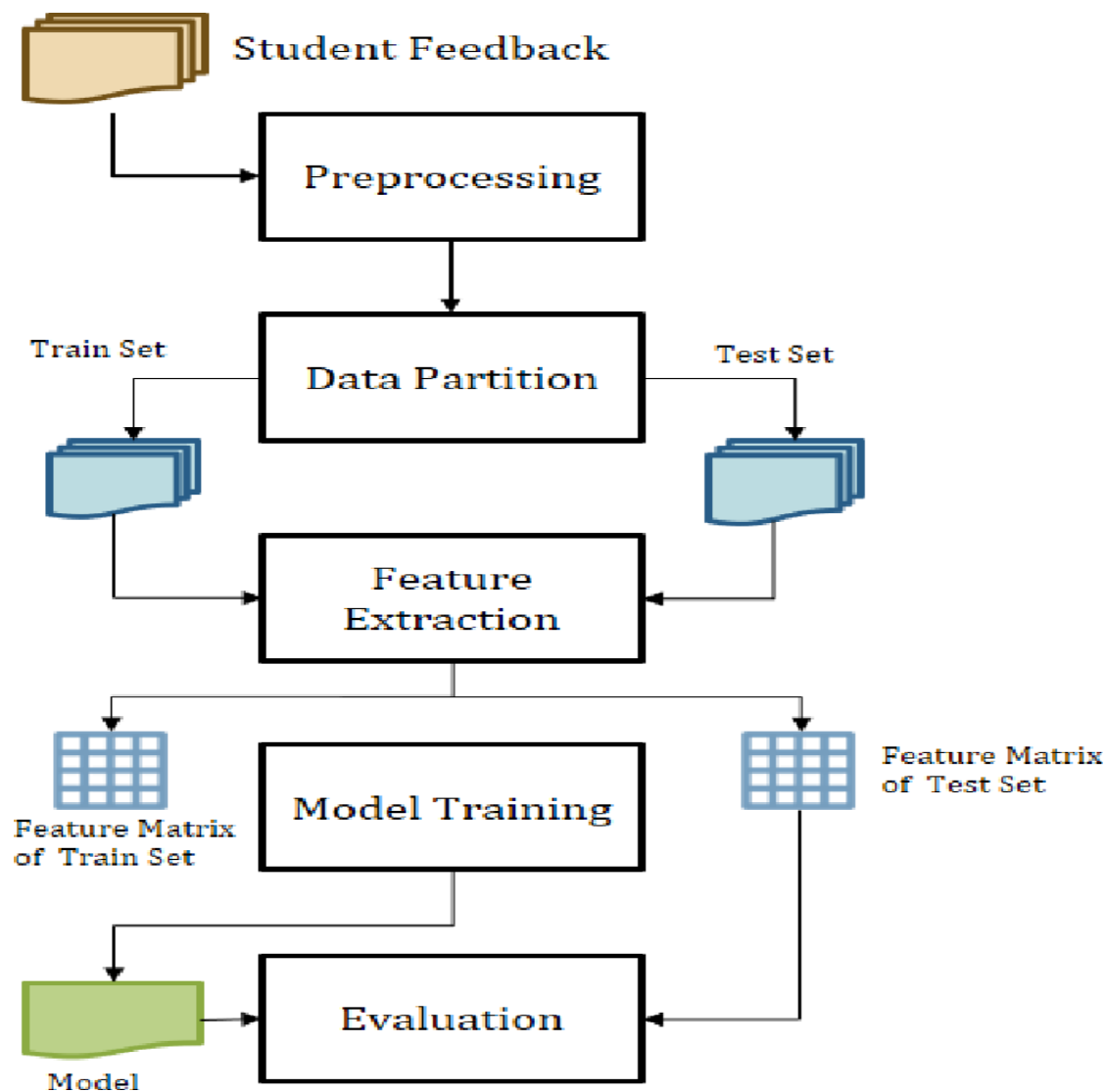


Figure: Architecture of sentiment analysis.

## Input Output Design

Data was given to the system in the form of text provided by students.

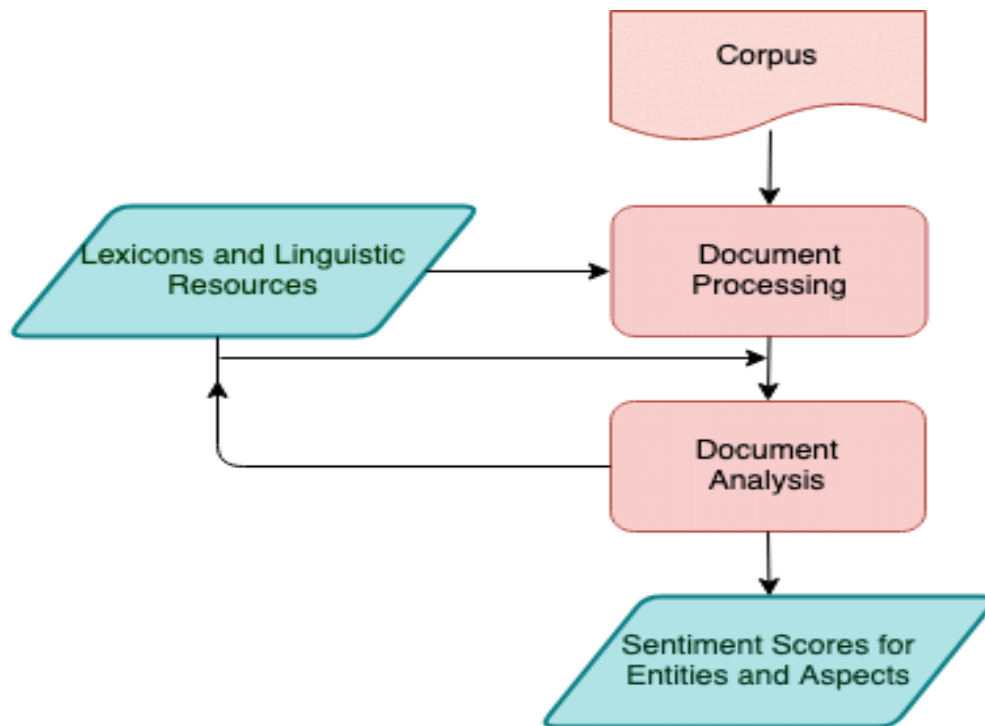


Figure: Input output design.

## Interface design

User interface design is responsible for a product's appearance, interactivity, usability, behavior, and overall feel. For the better GUI, we use HTML and CSS. Django is used to connect the GUI with the backend. For backend, we use python.

## Process design

Process design is a method to understand its processes and how it can improve them. A simple process design is presented which provides the information about sequential flow with the given flowchart.

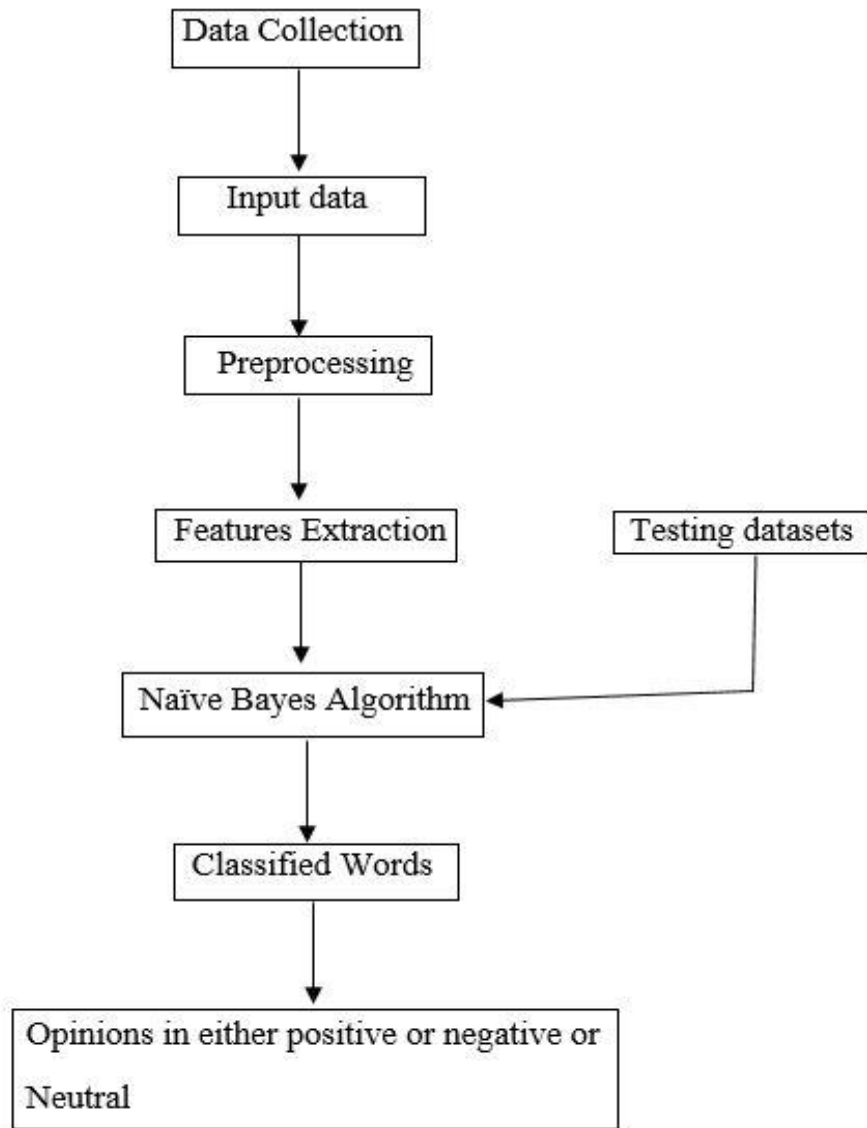


Figure: Flowchart of sentiment analysis on education system.



# **CHAPTER 5**

## **IMPLEMENTATION AND TESTING**

### **5.1 Implementation**

Systems implementation is a set of procedures performed to complete the design contained in the approved systems design document and to test, install, and begin to use the new or revised Information System. Systems implementation is the process of: defining how the information system should be built, ensuring that the information system is operational and used, ensuring that the information system meets quality standard.

#### **5.1.1 Tools Used**

##### **Front End**

##### **HTML**

HTML is the standard markup language for creating Web pages. HTML elements are the building blocks of HTML pages. With HTML constructs, images and other objects such as interactive forms may be embedded into the rendered page.

##### **CSS**

Cascading Style Sheets is a simple mechanism for adding style to Web documents. These pages contain information on how to learn and use CSS and on available software.

##### **IDE**

##### **Visual Studio**

The Visual Studio IDE is a creative launching pad that you can use to edit, debug, and build code, and then publish an app. Over and above the standard editor and debugger that most IDEs provide, Visual Studio includes compilers, code

completion tools, graphical designers, and many more features to enhance the software development process.

## **Backend**

### **Python**

Python is a high-level, general-purpose programming language. Its design philosophy emphasizes code readability with the use of significant indentation via the off-side rule. Python is dynamically typed and garbage-collected. It supports multiple programming paradigms, including structured, object-oriented and functional programming.

### **Django**

Django is a high-level Python web framework that encourages rapid development and clean, pragmatic design. Built by experienced developers, it takes care of much of the hassle of web development, so you can focus on writing your app without needing to reinvent the wheel.

## **5.1.2 Data Collection**

The data was collected through students in which internet has become the main source for users to express their opinions. At first, required data was collected from students by creating online form about education system. Students able to summarized their opinions in text form by covering their reviews few sentences.

## **5.1.3 Preprocessing**

After obtaining the data from the source, the data were pre-processed before the sentiment analysis stage this is to increase accuracy and to reduce error in the data. Data preprocessing is a technique to transfer a raw data in a useful and efficient format. Preprocessing can be discovering emoticons, words in upper case, removing stop words, finding exclamation marks or question marks, and removing inconsistent casing of letters. Preprocessing was done in order to clean the data which can be done by data cleaning process. It helps to convert the raw data into useable form. Tokenization process was done replacing sensitive data into unique identification symbols that retain all the essential information about the data without

compromising its security. Tokenization was done from the collected data by breaking a text of stream into words, symbol and other meaningful elements.

```
def preprocess_text(text):  
  
    print("%%%%%%%%%%>> I am preprocess Function <<%%%%%%%%%%")  
    print("----->> Before preprocessing <<-----")  
    print(text)  
    print("----->> After preprocessing <<-----")  
  
    text = text.lower() # Convert to lowercase  
    text = re.sub(r'http\S+', '', text) # Remove URLs  
    text = re.sub(r'\d+', '', text) # Remove digits  
    text = re.sub(r'[\W\s]', '', text) # Remove punctuation  
    text = re.sub(r'\s+', ' ', text) # Replace multiple spaces with single space  
    text = text.strip() # Remove leading and trailing whitespace  
    return text
```

Figure: Before preprocessing

```
=====>> Applying preprocessing technique <<=====>>  
  
----->> Before preprocessing <<-----  
I do not like @Nepal Education system.  
  
----->> After preprocessing <<-----  
['i', 'do', 'not', 'like', 'nepal', 'education', 'system']
```

Figure: After preprocessing

## 5.1.4 Feature Selection

After the completion of preprocessing, unnecessary data were then removed from the database. The selected words of the preprocessed texts were categorized into the respective groups which is known as feature categorization. The feature consists of similar nature of words that got significant role for identifying relevant attribute and increasing classification accuracy.

### 5.1.5 Description of Algorithm Used

The Naïve Bayes algorithm have been proved to work well with sentiment data and have been used. Naive Bayes algorithm gave good accurate results when implemented on reviews. Naïve Bayes algorithm is supervised algorithm. It is based on Bayes theorem and used for solving classification problems. Naïve Bayes algorithm is used to find the probabilities of the opinion words and classify them to respective classes by training, testing and evaluating the model. We can use this algorithm for text classification that includes high dimensional datasets. It is probabilistic classifier in which it predicts on the basis of probability of an object. It is based on Bayes' theorem which is used to determine the probability of a hypothesis with prior knowledge which depends on conditional probability.

The Bayes Theorem is defined as following equation

$$P(A|B) = P(B|A) * P(A) / P(B)$$

Where,

$P(A|B)$  represents the probability of event A happening given that B is true.

$P(B|A)$  represents the probability of event B happening given that A is true.

$P(A)$  and  $P(B)$  are the probabilities of observing A and B without any prior conditions. These are referred to as prior probabilities.

To classify the sentence as either positive or negative. We would first determine the probability that the text is positive and the probability that the text is negative. We would then compare the two probabilities. To calculate the probability of sentence being positive, we can rewrite Bayes' theorem as follows. Similarly, we would do the same thing for the negative class, except we would replace all occurrences of Positive with Negative.

$$P(\text{Positive} | \text{Sentence}) = P(\text{Sentence} | \text{Positive}) * P(\text{Positive}) / P(\text{Sentence})$$

$P(\text{Positive} | \text{Sentence})$  represents the probability of the text being positive.  $P(\text{Sentence} | \text{Positive})$  represents the probability of a positive text being sentence. To calculate this, we can count the number of occurrences of sentence in the positive texts and divide it by the total number of samples labeled as positive.  $P(\text{Positive})$  is the prior probability of a training sample being classified as positive. For example, if there are 10 sentences in our training data, and 7 sentences are positive, then

$P(\text{Positive})$  would be  $7/10$ . We can discard the denominator since the denominator will be the same for each class.

Naive Bayes assumes that every word in the sentence is independent of the other words. Instead of counting the occurrences of specific sentences in the training data, we are now looking at the frequency of individual words. We multiply the result by the prior probability of the class. After that, all we do is compare which class has the greatest probability, that class is the output of the classifier.

```
# Compute the class prior probability of each class
def compute_prior(data):
    print("~~~~~>> I am compute_prior Function <~~~~~")
    n = len(data)
    positive_count = sum(1 for row in data if row[1] == 'positive')
    negative_count = sum(1 for row in data if row[1] == 'negative')
    neutral_count = n - positive_count - negative_count
    return {'positive': positive_count / n, 'negative': negative_count / n, 'neutral': neutral_count / n}
```

Figure: Prior probability

```
# Compute the conditional probability of each word given each class
def compute_likelihood(data):
    print("~~~~~>> I am compute_likelihood Function <~~~~~")
    word_count = {'positive': {}, 'negative': {}, 'neutral': {}}
    word_total = {'positive': 0, 'negative': 0, 'neutral': 0}

    for ix, row in enumerate(data):
        print()
        print()
        print()

        print(ix, row[0])
        text = preprocess_text(row[0])

        words = text.split()
        print(words)

        print("")

        for index, word in enumerate(words):
            #print(word)

            if word not in word_count[row[1]]:
                word_count[row[1]][word] = 0
                #----->> As word count have "keys(either positive, negative or neutral)"

            word_count[row[1]][word] += 1
            word_total[row[1]] += 1
            #----->> weather if condition is valid or not but this statement execute
            print("---->>", index, row[1], word_count[row[1]])
```

Figure: Conditional probability

```

158 def predict(feedback, prior, called_from_msg, likelihood):
159     print("===== I am main predict Function <<=====")
160     #print(called_from_msg)
161     #print(feedback)
162     feedback = preprocess_text(feedback)
163     words = feedback.split()
164     log_prob = {'positive': log(prior['positive']), 'negative': log(prior['negative']), 'neutral': log(prior['neutral'])}
165     for word in words:
166         if word in likelihood['positive']:
167             #print(" found ", word, "=", likelihood)
168             log_prob['positive'] += log(likelihood['positive'][word])
169         else:
170             #print("Not found ", word, "=", likelihood)
171             log_prob['positive'] += log(1e-10)
172         if word in likelihood['negative']:
173             log_prob['negative'] += log(likelihood['negative'][word])
174         else:
175             log_prob['negative'] += log(1e-10)
176         if word in likelihood['neutral']:
177             log_prob['neutral'] += log(likelihood['neutral'][word])
178         else:
179             log_prob['neutral'] += log(1e-10)
180     return max(log_prob, key=log_prob.get)
181

```

Figure: Main Predict Function

## 5.1.6 Classified Words

The sentence of the student's opinion were classified to the respective class either positive or negative or neutral class.

## 5.2 Testing

System testing is the process in which a quality assurance team evaluates how the various components of an application interact together in the full, integrated system or application.

### 5.2.1 Unit Testing

Unit testing was done with several tasks to complete the whole system. We performed unit testing on every activity from preprocessing to the final results. Finally, whole system worked fine.

### **5.2.2 System Testing**

First of all, we tested web application and classification tasks and then integrate them to form the whole system. Then, we tested the system which gave positive results.

### **5.3 Result Analysis**

Our project main goal is to classify the given sentence into respective class on the basis of trained datasets. We have separate the entire datasets into 2 parts. Training and testing dataset.

To analyze the accuracy of our project, we split out datasets into 80-20 on testing and training. So the accuracy is 75%.

# **CHAPTER 6**

## **CONCLUSION AND FUTURE RECOMMENDATION**

### **6.1 Conclusion**

Our system is successfully designed to detect the sentiment of sentences provided by students in the form of texts. We have created a large set of databases that are used to classify the opinions of students. The main goal of our Project Work is to scan the words of sentence and classify them and provide the output more efficiently in the form of positive, negative and neutral. We have trained the datasets to make it more understandable so that it finally provided output with more accuracy.

### **6.2 Future Recommendation**

The future of sentiment analysis is going to continue to dig deeper. Sentiment analysis is already evolving rapidly from positive, negative or neutral to more granular and deep understanding. At Revealed Context, we also have intensity and confidence scores, as well as emotion and more. Sentiment analysis can also be used in various sectors with more accuracy in social media, customer product. Sentiment analysis is getting better because social media is increasingly more emotive and expressive.

In context of our country Nepal, Sentiment analysis on education system can be used to understand the opinions of students whether students are satisfied or not with the current education trends so that improvement can be done for better education system.



## REFERENCES

- [1] Esparza GG. A sentiment analysis model to analyze students reviews of teacher performance using support vector machines. In: International Symposium on Distributed Computing and Artificial Intelligence. 2017. Springer.
- [2] <https://ieeexplore.ieee.org/document/7395820>
- [3] Rajput Q, Haider S and Ghani S 2016 Lexicon-based sentiment analysis of teachers' evaluation Applied Computational Intelligence and Soft Computing
- [4] İskender E and Batı G B 2015 Comparing Turkish Universities Entrepreneurship and Innovativeness Index's Rankings with Sentiment Analysis Results on Social Media Procedia-Social and Behavioral Sciences 195 pp 1543-1552
- [5] Abdelrazeq A, et al. 2016 Sentiment analysis of social media for evaluating universities Automation, Communication and Cybernetics in Science and Engineering 2015/2016 Springer pp 233-251
- [6] Balachandran and Kirupananda A 2017 Online reviews evaluation system for higher education institution: An aspect based sentiment analysis tool 2017 11th international conference on software, knowledge, information management and applications (SKIMA) IEEE
- [7] Santos C L, Rita P and Guerreiro J 2018 Improving international attractiveness of higher education institutions based on text mining and sentiment analysis International Journal of Educational Management
- [8] Aung K Z and Myo N N 2017 Sentiment analysis of students' comment using lexicon based approach 2017 IEEE/ACIS 16th international conference on computer and information science (ICIS) IEEE
- [9] Kaewyong P, et al. 2015 The possibility of students' comments automatic interpret using lexicon based sentiment analysis to teacher evaluation 3rd International Conference on Artificial Intelligence and Computer Science (AICS2015)

[10] Altrabsheh N, Cocea M and Fallahkhair S 2014 Sentiment analysis: towards a tool for analysing real-time students feedback 2014 IEEE 26th international conference on tools with artificial intelligence IEEE

# APPENDICES

## Sentiment Analysis on Nepal Education System

Enter your opinion:

I like Nepal education system.

Classify

Your Opinion belongs to below class



**Positive**

## Sentiment Analysis on Nepal Education System

Enter your opinion:

Nepal Education system is not good.

Classify

Your Opinion belongs to below class



**Negative**

## Sentiment Analysis on Nepal Education System

Enter your opinion:

I do not dislike Nepal education system.

Classify

Your Opinion belongs to below class



**Neutral**

# LOG

| Supervisor: Surya Bam |                                                                  |
|-----------------------|------------------------------------------------------------------|
| Date                  | Activity                                                         |
| 28 December,2022      | Meeting with supervisor for project title                        |
| 8 January,2023        | Proposal submission                                              |
| 15 January,2023       | Re-submission of proposal                                        |
| 24 January,2023       | Proposal accepted                                                |
| 28 January,2023       | Discussion about Naïve Bayes Algorithm                           |
| 8 February,2023       | Meeting with supervisor about datasets                           |
| 24 February,2023      | Proposal defense                                                 |
| 5 March,2023          | Meeting with supervisor in Google Meet                           |
| 8 March,2023          | Discuss with supervisor about output and accuracy of our project |
| 23 April,2023         | Submitted softcopy of project report                             |

