

**SENTIMENT ANALYSIS IN PERSPECTIVE OF SHIFTING DIU CAMPUS TO
ASHULIA FROM DHANMONDI**

BY

Md. Nazmul Hasan
ID: 183-15-11881

Md. Ashikuzzaman
ID: 143-15-4524

This Report Presented in Partial Fulfillment of the Requirements for the
Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Ms. Asma Mariam
Lecturer
Department of CSE
Daffodil International University

Co-Supervised By

Md. Abbas Ali Khan
Assistant Professor
Department of CSE
Daffodil International University



DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2023

APPROVAL


This Project/internship titled “**Sentiment Analysis in Perspective of Shifting DIU campus to Ashulia from Dhanmondi.**” submitted by **Md. Nazmul Hasan**, ID: 183-15-11881 and **Md. Ashikuzzaman**, ID: 143-15-4524 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 05-01-2023.

BOARD OF EXAMINERS

Chairman

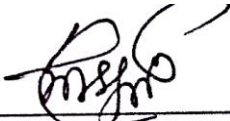
Dr. Touhid Bhuiyan
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner



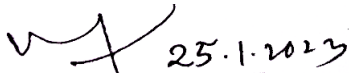
Dr. Md. Monzur Morshed
Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner



Dewan Mamun Raza
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner



Dr. Ahmed Wasif Reza
Associate Professor
Department of Computer Science and Engineering
East West University

DECLARATION

I hereby declare that this project has been done by us under the supervision of **Ms. Asma Mariam**, **Department of CSE** Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for an award of any degree or diploma.

Supervised by:

Asma

Ms. Asma Mariam
Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

Chan

Md. Abbas Ali Khan
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

Nazmul Hasan

Md. Nazmul Hasan
ID: 183-15-11881
Department of CSE
Daffodil International University

Ashikuzzaman

Md. Ashikuzzaman
ID: 143-15-4524
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, I express my heartiest thanks and gratefulness to almighty Allah for his divine blessing made me that possible to complete the final thesis successfully.

I am really grateful and wish my profound indebtedness to Ms. Asma Mariam, Department of CSE, Daffodil International University, Dhaka. Deep Knowledge & keen interest of my supervisor in the field of “Natural Language Processing and Machine learning” to carry out this thesis. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading and many inferior drafts, and correcting them at all stages have made it possible to complete this thesis.

I would like to express our heartiest gratitude to Professor Dr. Touhid Bhuiyan and Head, Department of CSE, for his kind help to finish my project and also to other faculty members and the staff of the CSE department of Daffodil International University.

I would like to thank our entire course mate in Daffodil International University, who took part in this discussion while completing the course work.

Finally, I must acknowledge with due respect the constant support and patients of my parents and brother.

ABSTRACT

Recently, the campus of Daffodil International University moved from Dhanmondi to Ashulia. Consequently, a variety of responses were observed among the students. Based on students' various reviews to campus changes, a sentiment dataset was created. Sentiment of the students are categorized into three types e.g., positive, neutral and negative. Different sentiment preprocessing techniques have been used to preprocess those collected dataset. By the help of this dataset, a machine learning based Random Forest model is generated to analyze the sentiment of the students. In consequence, this model can predict any sentiment of a particular student whether it is positive, negative or neutral. Eventually the accuracy of this sentiment analysis of the students using Random Forest algorithm for shifting the campus to Ashulia from Dhanmondi is 83.36%. Thus, this model can predict a particular student's sentiment with very good accuracy by which authority can take necessary steps to improve the sector where student's review was negative.

TABLE OF CONTENTS

CONTENTS	PAGE
Board of examiners	ii
Declaration	iii
Acknowledgments	iv
Abstract	v
CHAPTERS	
CHAPTER 1: INTRODUCTION	1-4
1.1 Introduction	1
1.2 Objective	2
1.3 Motivation	2
1.4 Rationale of the Study	2
1.5 Research Questions	3
1.6 Expected Outcome	3
1.7 Report Layout	4
Chapter 2: BACKGROUND	5-7
2.1 Introduction	5
2.2 Sentiment Analysis	5
2.3 Related Work	6
2.4 Research Summary	6
2.5 Scope of the Problem	7
2.6 Challenges	7

Chapter 3: RESEARCH METHODOLOGY	8-15
3.1 Introduction	8
3.2 Research Subject and Instrumentation	8
3.3 Data Collection Procedure	9
3.4 Implementation Requirements	9
3.4.1 Methodology for Sentiment Analysis	9
3.4.2 Data Analysis	10
3.4.3 Data Pre-Processing	10-14
3.4.4 Sentiment Analysis Procedure	15
Chapter 4: RESULT AND DISCUSSION	16-21
4.1 Introduction	16
4.2 Sentiment Prediction	16-19
4.2.1 Random Forest Algorithm	19
4.2.2 Working Procedure	20
4.3 Evaluation	21
CHAPTER 5: CONCLUSION AND FUTURE WORK	22-23
5.1 Summary of the Study	22
5.2 Conclusion	22
5.3 Future Work	23
APPENDIX	24
REFERENCES	25

LIST OF FIGURES

FIGURES	PAGE NO
Figure 1: Research Methodology	9
Figure 2: Data Preprocessing Procedure	10
Figure 3: Sample of Collected Data	11
Figure 4: Use of Mapping Function	12
Figure 5: Pre-processing Steps for Sentiment Analysis	13
Figure 6: Sentence & Word Tokenization	13
Figure 7: Frequency Distribution	14
Figure 8: Frequency Distribution Plot	14
Figure 9: Identifying Sentiment Analysis Prediction	17
Figure 10: Load the Dataset in Collab	18
Figure 11: Number of Reviews	18
Figure 12: Student's Review in Bar-Chart	19
Figure 13: Working Procedure of Random Forest Algorithm	20
Figure 14: Model Generation & Evaluation	21

LIST OF TABLES

TABLES	PAGE NO
Table 1: Accuracy of Random Forest Algorithm	21

CHAPTER 1

INTRODUCTION

1.1 Introduction

Recently Daffodil International University campus has been shifted to Ashulia from Dhanmodi. After moving the campus from Dhanmondi to Ashulia, several sentiments have been observed among the students. A lot lot of questions regarding this shifting have been asked to the students. The answers from students is used to create a dataset that would help us understand how they felt. With this dataset, we developed models employing a variety of machine learning algorithms. We ultimately settled on Random Forest as the best algorithm because it makes sentiment analysis of students simple and accurate. It is widely used to identify and classify opinions regarding a product, service, or idea or feeling. It might be at the level of a document, a sentence, or a feature. Sentences are taken from complete reviews. We gather information from understudies subsequent to moving the grounds from Dhanmondi to Ashulia for this review.

Hard copy of frequently asked questions and a google form is used to collect responses from students regarding their reactions to the campus moving from Dhanmondi to Ashulia in order to collect data. After putting together all of these responses, we create a dataset and use a machine learning algorithm to combine it with the Random Forest algorithm, which we find to be the most accurate and gives the best results. This indicates that students' Random Forest sentiment analysis can be carried out with ease and precision. By using this Random Forest algorithm we will predict a students review whether it is negative, positive or neutral.

1.2 Objective

- To observe the students' reactions after the campus was moved from Dhanmondi to Ashulia.
- To analyze the overall facilities in Ahulia campus compared to Dhabmondi campus.
- To differentiate between the environment of Ashulia campus compared to Dhanmondi campus.
- To find the best model and algorithm for deciding on a student's reaction that produces better results and accuracy than others.

1.3 Motivation

Recently, the campus of Daffodil International University moved from Dhanmondi to Ashulia. The students who had been attending the Dhanmondi campus in the past underwent a noticeable transformation as a result of this. Different kinds of responses were seen inside them for different purposes. We decided to use this reaction to collect all of the students' responses as a result. In order to accomplish this, we generate a few questions using a Google form and record each student's responses as answers. Using this data set and a number of different machine learning algorithms, we found that the Random Forest algorithm gives us the best results and the best accuracy. This indicates that Random Forest algorithm can easily and accurately analyze the sentiment analysis of the students.

1.4 Rationale of the Study

The process of determining the text's emotional tone is known as sentiment analysis [2]. In this study, sentiment analysis was carried out with the help of student reviews. We gathered responses from Daffodil International University students regarding the campus relocation from Dhanmondi to Ashulia for this study. For a variety of reasons, some students provide constructive feedback, while others provide constructive feedback. We create a dataset and use it in some machine learning algorithms after collecting all of their reviews. We find that Random Forest produces the best results, allowing us to easily and accurately determine the students' feelings.

1.5 Research Question

As part of our research, we chose the following questions.

- How do you feel about the Ashulia Campus's study environment?
- What do you think of the relationship between teachers and students?
- How are the Ashulia Campus's lab facilities?
- What changes occur in the outcomes of students after they leave campus?
- How are the facilities at Ashulia Campus in general?
- How is the campus of Ashulia's sports environment?
- How do you feel about the work environment at the club?

1.6 Expected outcome

We wanted to find out what the students thought about the campus relocation from Dhanmondi to Ashulia. The final results are as follows:

- The reviews provided by students make it simple to identify the study environment.
- Can learn about the relationships between teachers and students.
- Can compare lab facilities to the Dhanmondi campus and learn about lab facilities.
- Can assess student evaluations to determine the sporting environment.
- Effectively finds club working climate.

1.7 Report layout

The report layout includes a summary of all chapters. The following is a synopsis of all chapters:

Chapter 1: A discussion of the thesis motivation, the study's rationale, the research question, and the expected outcome.

Chapter 2: Introduction to the thesis, Sentiment Analysis, related work, research summary, the scope of the problems, and challenges.

Chapter 3: Discusses the methodology of my research work. Data mining and machine learning techniques in detail. The data collection processes are also briefly mentioned here.

Chapter 4: Discuss the details of the outcome and mention and out of that project's experiment and result.

Chapter 5: Discuss the scope of my research and how it can be carried out in the future.

CHAPTER 2

BACKGROUND

2.1 Introduction

The campus of Daffodil International University recently moved from Dhanmondi to Ashulia. Consequently, the students displayed a variety of reactions. After moving the campus from Dhanmondi to Ashulia, we gathered a lot of questions and answers from students to create a dataset that would help us understand how they felt. With this dataset, we made some models using machine learning algorithms, and we found that Random Forest was the best. With this algorithm, sentiment analysis of students can be easily and accurately done. It is widely used to identify and classify opinions regarding a product, service, or ideas and feelings. It could be at the level of a document, a sentence, or a feature. Sentiments can be gleaned from comprehensive reviews. After the campus moved from Dhanmondi to Ashulia, we collect data for this study from students. We used a Google form to collect responses from students regarding their reactions to the campus relocation from Dhanmondi to Ashulia in order to collect data. We create a dataset after collecting all of these responses. By combining this dataset with a variety of machine learning algorithms, we find that the Random Forest algorithm delivers the best results with the highest accuracy. This indicates that students' Random Forest sentiment analysis can be easily and precisely analyzed.

2.2 Sentiment Analysis

The process of determining the text's emotional tone is known as sentiment analysis. In this study, sentiment analysis was carried out with the help of student reviews. We gathered responses from Daffodil International University students regarding the campus relocation from Dhanmondi to Ashulia for this study. For a variety of reasons, some students provide constructive feedback, while others provide constructive feedback. We create a dataset and employ it in a number of machine learning algorithms after compiling all of their reviews. We discover that Random Forest delivers the best results. The majority of students, as we can see, provide positive feedback, while a few provide negative feedback and others are neutral.

2.3 Related work

For this research project, we read numerous sentiment analysis studies on a variety of subjects. Here, we can see that sentiment analysis has been the subject of a lot of research.

Vishal A. Kharde *et. al.* [1], discussed about a Sentiment analysis on different communities, or post messages on twitter across the world. Their survey is primarily focuses on sentiment analysis of Twitter data, which is useful for analyzing tweets containing opinions that are highly unstructured, heterogeneous, and either positive or negative—or even neutral in some instances. In this paper, they have used machine learning algorithms like Naive Bayes, Max Entropy, and Support Vector Machine and got accuracy of 74.56%, 74.94%, and 76.68% respectively to analyze the sentiments on different twitter data.

Palak Baid *et. al.* [2], proposed a Sentiment analysis based on movie reviews using machine learning techniques. In this paper, they have analyzed the Movie reviews using various techniques like Naïve Bayes, K-Nearest Neighbors and Random Forest. The accuracy the have got in this Sentiment analysis is 81.4% for Naïve Bayes, 55.30% for K-Nearest Neighbors and 78.65% for Random Forest.

Gurshobit Singh *Brar et. al.* [3], proposed a supervised machine learning based Sentiment analysis of movie reviews. In this paper, they have used supervised machine learning and feature-based opinion mining, to analyze the sentiment on movie reviews. Using nouns, verbs, and adjectives as opinion words, the primary goal of this paper is to ascertain the polarity of reviews. They have divided the sentiments into two categories such as positive and negative reviews. The average accuracy they got using those models for testing the review is 81.22%.

Federico Neri *et. al.* [4], presented a Sentiment analysis study which is performed on over than 1000 Facebook posts about newscasts, comparing the sentiment for Rai - the Italian public broadcasting service - towards the emerging and more dynamic private company La7. The overall performance of this Sentiment analysis is measured using Recall and Precision with 87% and 93% respectively.

2.4 Research Summary

The general objective of our exploration is to find out about the understudies' responses after the grounds moved from Dhanmondi to Ashulia. We gather their feedback and compile it into a data set for this purpose. Using this data set and a machine learning algorithm, we find that most students give positive reviews, some give negative reviews, and some are neutral. Students previously received more benefits on these grounds at the Dhanmondi campus, but in some instances, students receive more benefits than others. After the campus change, this research can quickly ascertain the positive and negative reviews left by students. This enables the authority to take the necessary measures to address students' complaints and issues following the campus change.

2.5 Scope of the Problem

We gathered responses from Daffodil International University students regarding the campus relocation from Dhanmondi to Ashulia for this study. For a variety of reasons, some students provide constructive feedback, while others provide constructive feedback. After collecting all of their reviews, we produce a dataset. The use of the dataset in these models has some drawbacks. By making use of the data set, we are able to group them together and separate the positive, negative, and neutral data for use in various machine learning algorithms; Random Forest delivers the best results. The majority of students, as we can see, provide positive feedback, while a few provide negative feedback and others are neutral.

2.6 Challenges

In our study of sentiment analysis, the most difficult obstacle we face is locating the optimal algorithm that can quickly and accurately determine the students' overall reactions to the switch from Dhanmondi to Ashulia. Another problem was that some students refused to give feedback, while others gave feedback that was neutral. However, a significant number of students provide their own feedback, which enables us to construct a data set and employ machine learning algorithms to select the one that delivers the most accurate results.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

The students' feelings and reactions to the campus's move from Dhanmondi to Ashulia are the basis for our research topic. For this review, we initially pre-arranged a few inquiries and gathered every one of the reactions from understudies about their responses after the shift. We used a dataset in a machine learning algorithm after gathering these reviews. To do this, first, we figure out how many students gave either positive, negative, or neutral feedback. We found that of all the machine learning algorithms, Random Forest gives us the best results. This means that we can use this algorithm to easily look at how students feel about the switch from Dhanmondi to Ashulia.

3.2 Research Subject and Instrumentation

Our research topic is based on the emotions and reactions of students after the campus moved from Dhanmondi to Ashulia. We first developed some questions for this study and gathered all of the responses from students regarding their reactions to the shift. We ordered these surveys into a dataset, which we then, at that point, took care of into an AI calculation. We start by figuring out how many students gave either positive, negative, or neutral feedback. We discovered that Random Forest outperforms all other machine learning algorithms, indicating that we can easily use this algorithm to analyze students' reactions to the switch from Dhanmondi to Ashulia.

3.3 Data Collection Procedure

For our study, we gathered data from the students. For this, we made some questions, and the students used a Google form to answer them. The following were our inquiries: What do you think of the Ashulia Campus's study environment? How do you feel about the relationship between students and teachers? How are the Ashulia Campus's laboratory facilities? How are students' outcomes changing after they leave campus? How are the facilities on the Ashulia Campus in general? How is the campus of Ashulia's athletic environment? How do you feel about the club's work environment? After we have gathered all of these responses, we make a data set and use it in a few machine learning algorithms to choose the one that delivers the best results.

3.4 Implementation Requirement

The following steps must be followed for the final implementation of our research and to determine the best model and algorithm using the collected dataset from the students:

3.4.1 Research Methodology for Sentiment Analysis.

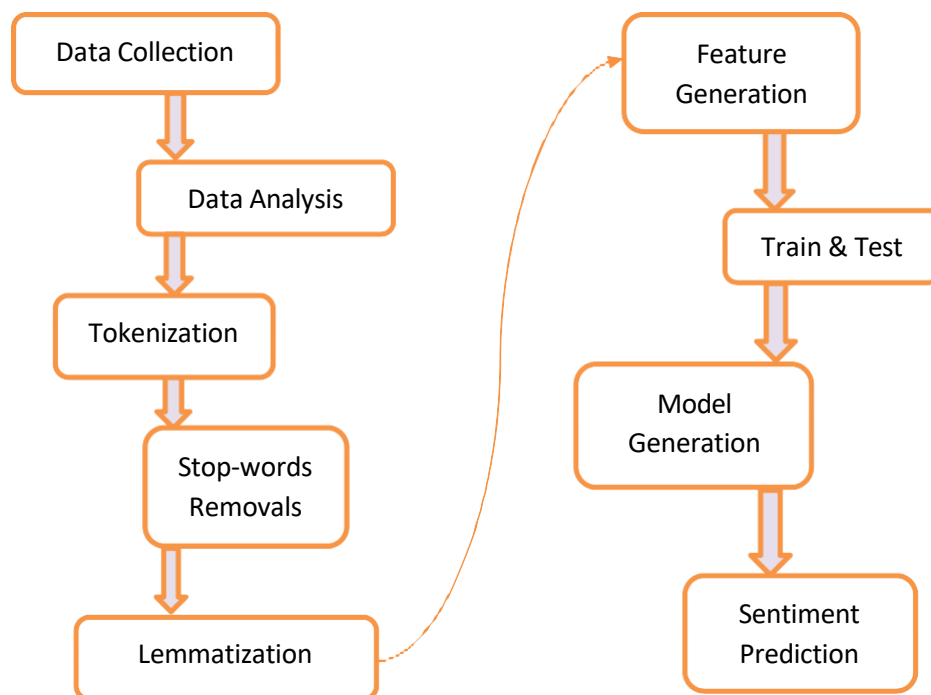


Fig. 1: Research Methodology of this Study

Fig. 1 exhibits the research methodology of our sentiment analysis process. Each step must be followed sequentially. Thus, we can analyze the sentiment of a particular student whether his/her sentiment is negative, positive or neutral.

3.4.2 Data Analysis:

For the purpose of this research, we gathered a substantial amount of information from students from various groups, campuses, departments, and genders. After collecting it, we put it into a data set. Then, at that point, we preprocess the informational index, break down it, lastly use it in a few AI calculations to find the best calculation with the most noteworthy precision [7].

3.4.3 Data Pre-Processing:

Data preprocessing is the process of preprocessing a data set before using it in an algorithm. Data preprocessing manages all of the important data in a data set while removing unnecessary items from it. Because our data set contains some unnecessary data, we must first preprocess it before applying it to some machine learning algorithms in order to achieve better results [6].

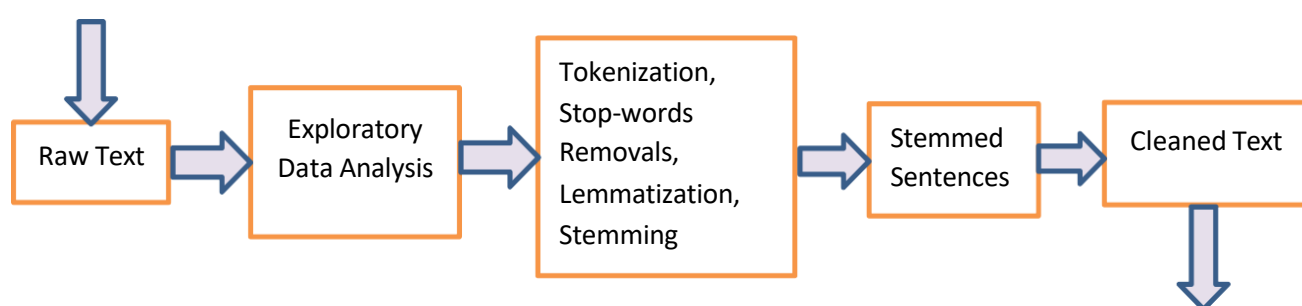


Fig. 2: Data Preprocessing Procedures

Fig. 2 displays the preprocessing procedure of our text data. In this section, a raw text is taken as input and then exploratory analysis of the texts will be applied. Then all the words will be tokenized using tokenizer and unnecessary texts will be erased by using stop-word removal. After this all the texts will be formatted into its original form using PortStemmer which is necessary for our sentiment analysis to produce stemmed sentences. Eventually

we will get the cleaned text by means of we can predict someone's sentiment whether it is positive, negative or neutral.

In this section, we will load our dataset which we have collected from the students. There are three columns consists of review, rating and sentiment. In which review column is considered as predictor class and the sentiment column is considered as targeted class.

Fig. 3 shows the sample of our collected data which we used to predict the sentiment of the students.

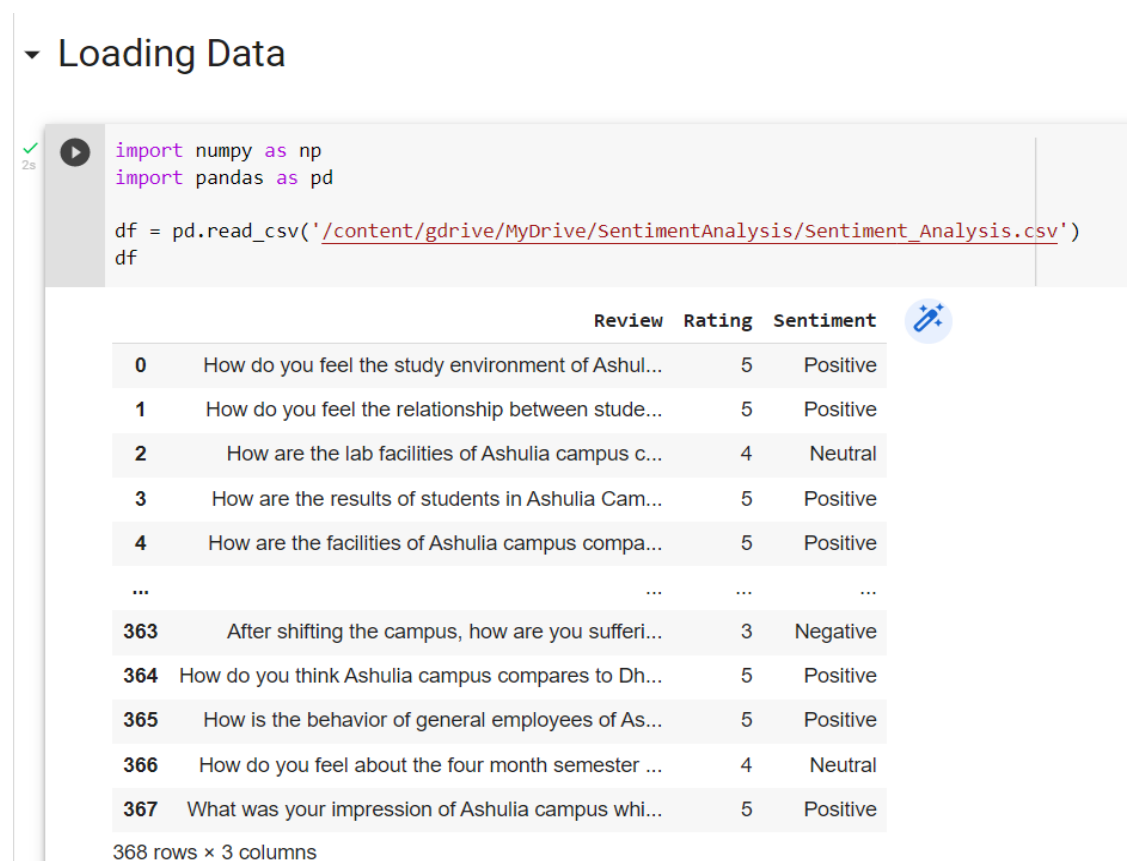


Fig 3: Sample of Collected Data

After loading the dataset, we need to explore the data which is exploratory data analysis. In this segment we will check if there any null value exists in our dataset or not. Then we need to mapping the targeted class. As the Sentiment class which we are going to predict is in categorical form, so we must need to convert it into numerical value. Since there are 3 classes in the Sentiment column, so it will be mapped as 'Positive' value is replaced by '1', 'Negative' value is replaced by '0' and the 'Neutral' value is replaced by '2'.

```
#Use of mapping function

mapping = {'Positive' :1 , 'Negative' : 0 , 'Neutral' : 2}
df['Sentiment'] = df['Sentiment'].map(mapping)
df
```

	Review	Rating	Sentiment
0	How do you feel the study environment of Ashul...	5	1
1	How do you feel the relationship between stude...	5	1
2	How are the lab facilities of Ashulia campus c...	4	2
3	How are the results of students in Ashulia Cam...	5	1
4	How are the facilities of Ashulia campus compa...	5	1
...
363	After shifting the campus, how are you sufferi...	3	0
364	How do you think Ashulia campus compares to Dh...	5	1
365	How is the behavior of general employees of As...	5	1
366	How do you feel about the four month semester ...	4	2
367	What was your impression of Ashulia campus whi...	5	1

368 rows × 3 columns

Fig. 4: Use of Mapping Function

Comparing **Fig. 3** and **Fig. 4** we can see that the 'Sentiment' column is reformed into a numerical value from categorical value which is the existing 3 classes in the Sentiment column, have been mapped as '1' where the sentiment is 'Positive', '0' where the sentiment is 'Negative', and '2' where the sentiment is 'Neutral'.

Fig. 5 exhibits the preprocessing steps of this sentiment analysis. There are three steps e.g., tokenization, stopwords removal and lemmatization.

▼ Pre-Processing Steps for Sentiment Analysis

- Tokenization
- Stopwords Removals
- Lemmatization

```
[13] import nltk
      nltk.download('punkt')
      from nltk import tokenize

[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data] Unzipping tokenizers/punkt.zip.
```

Fig. 5: Pre-Processing Steps for Sentiment Analysis

▼ Tokenization

Sentence Tokenization

```
[14] from nltk.tokenize import sent_tokenize
      text="How are the lab facilities of Ashulia campus compared to Dhanmondi campus?"
      tokenized_text=sent_tokenize(text)
      print(tokenized_text)

['How are the lab facilities of Ashulia campus compared to Dhanmondi campus?']
```

Word Tokenization

```
[21] from nltk.tokenize import word_tokenize
      tokenized_word = word_tokenize(text)
      print(word_tokenize(text))

['How', 'are', 'the', 'lab', 'facilities', 'of', 'Ashulia', 'campus', 'compared', 'to', 'Dhanmondi', 'campus', '?']
```

Fig. 6: Sentence & Word Tokenization

Fig. 6 describes the procedure of the sentence and word tokenization. As we can see there, we have taken a sentence and then the word_tokenize library reform this sentence into tokenizes words.

Here, we will discuss about the frequency distribution of our analyzed sentiment dataset. In a particular review of a student how many tokenizes words have been appeared frequently will be categorized and visualized. **Fig. 7** shows the frequency distribution of each word in the sentences.

Frequency Distribution

```
✓ 0s [▶] from nltk.probability import FreqDist
      fdist = FreqDist(tokenized_word)
      print(fdist)
      print("2 most common :- ",fdist.most_common(2))

[↵] <FreqDist with 12 samples and 13 outcomes>
      2 most common :- [('campus', 2), ('How', 1)]
```

Fig. 7: Frequency Distribution

Frequency Distribution Plot

```
✓ 0s [▶] import matplotlib.pyplot as plt
      fdist.plot(30,cumulative=False)
      plt.show()
```

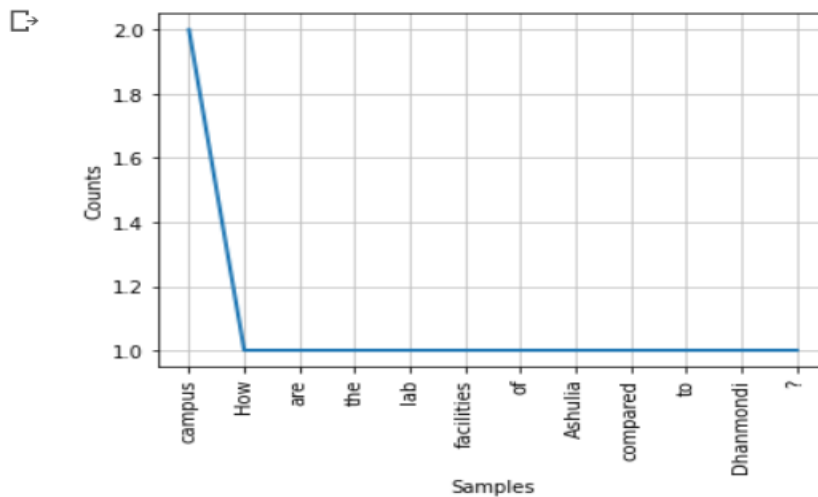


Fig. 8: Frequency Distribution Plot

Fig. 8 displays the frequency distribution plot of the previously generated frequency distribution of each word in a sentence.

3.4.4 Sentiment Analysis Procedure

The use of natural language processing, text analysis, and computational linguistics to determine a speaker's or writer's attitude toward a particular subject is referred to as Sentiment Analysis. Essentially, it aids in determining whether a text is expressing positive, negative, or neutral sentiments. This process is considered as a way to learn how people, particularly customers, feel about a particular idea, product, or topic is through sentiment analysis [5].

The process of determining the text's emotional tone is known as sentiment analysis. In this study, sentiment analysis was carried out with the help of student reviews. We gathered responses from Daffodil International University students regarding the campus relocation from Dhanmondi to Ashulia for this study. For a variety of reasons, some students provide constructive feedback, while others provide constructive feedback. We create a dataset and employ it in a number of machine learning algorithms after compiling all of their reviews. We discover that Random Forest delivers the best results. The majority of students, as we can see, provide positive feedback, some provide negative feedback, and others are neutral. The model was trained with 75% of the data in our prepared data set, and the model was predicted with the remaining 25%. With this dataset, we developed models employing a variety of machine learning algorithms. We ultimately settled on Random Forest as the best algorithm because it makes sentiment analysis of students simple and accurate.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

Students responded in a variety of ways to the campus relocation. After moving the campus from Dhanmondi to Ashulia, we gathered a lot of questions and answers from students to create a dataset that would help us understand how they felt. After the campus moved from Dhanmondi to Ashulia, we collect data for this study from students. We used a Google form to collect responses from students regarding their reactions to the campus relocation from Dhanmondi to Ashulia in order to collect data. After collecting all of these responses, we create a dataset and combine it with a variety of machine learning algorithms to find that the Random Forest algorithm delivers the most accurate results. This indicates that Random Forest can easily and accurately analyze the sentiment analysis of the students. In order to ascertain how students feel about the campus relocation from Dhanmondi to Ashulia, it is essential to determine the best algorithm for obtaining better results and accuracy from a large number of data sets.

4.2 Sentiment Prediction

The process of determining the text's emotional tone is known as sentiment analysis. Student reviews served as the basis for sentiment analysis in this research. For this study, we gathered responses from Daffodil International University students regarding the campus relocation from Dhanmondi to Ashulia. For a variety of reasons, some students provide negative feedback, positive feedback, and neutral feedback. We create a dataset and employ it in a number of machine learning algorithms after compiling all of their reviews; we discover that Random Forest delivers the best results. We can see here that most of the understudies give positive audits, some give negative surveys, and some are unbiased. The remaining 25% of our prepared data set was utilized for model prediction and training. With this dataset, we made a few models using machine learning algorithms, and we found that Random Forest was the best. With this algorithm, sentiment analysis of students can be easily and accurately done.

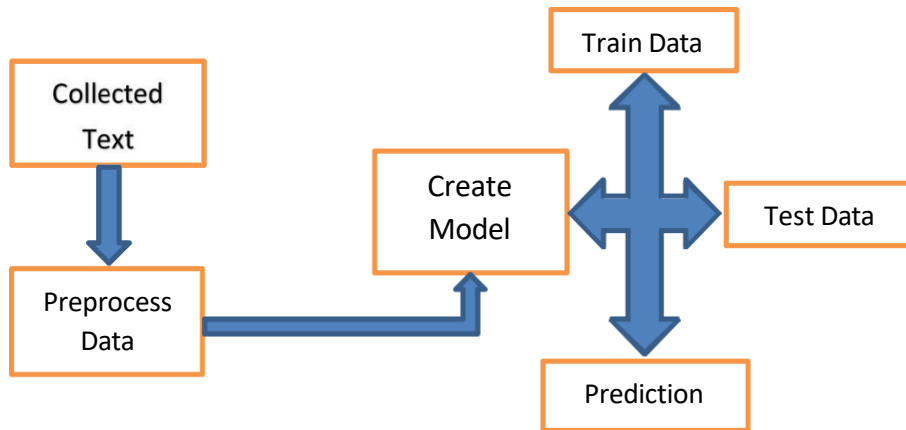


Fig. 9: Identifying Sentiment Analysis Prediction

Fig. 9 displays the whole procedure of the Sentiment analysis prediction. Collected dataset needs to preprocess and then a model is created using Random Forest. Those datasets will train into the dataset to be predicted whether a review of a particular student is positive, negative or neutral.

▼ Loading Data

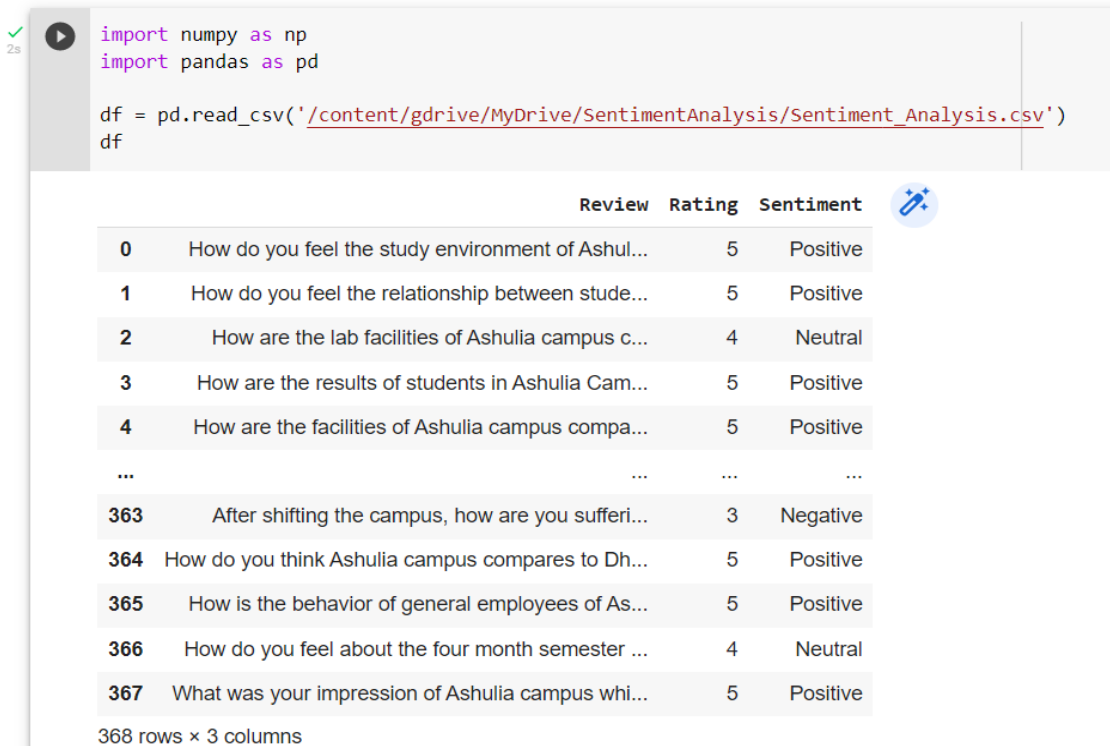


Fig. 10: Load the Dataset in Colab

To begin with this, we have loaded the dataset which is being collected from the students. This dataset is uploaded into the google drive and the using google colab we have loaded the dataset. **Fig. 10** displays how we can load our dataset from Colab.

▼ Exploratory Data Analysis

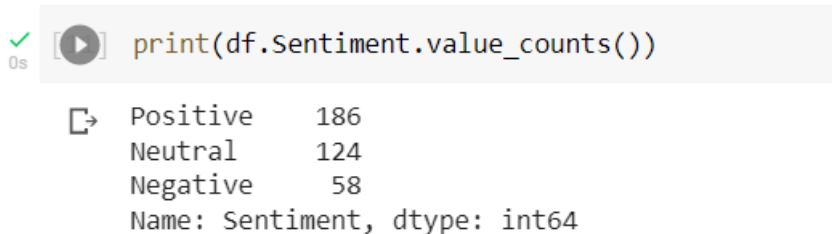


Fig. 11: Number of Reviews

Fig. 11 describes the number of reviews provided by the students and among them how many numbers of them are positive, how many are neutral and how many are negative.

Here by this exploratory data analysis, it is seen that Positive review rate is 186, Neutral is 124 and Negative review rate is 58.

▾ Visualize the Data Plot

```
[12] import matplotlib.pyplot as plt
      Sentiment_count=df.groupby('Sentiment').count()
      plt.bar(Sentiment_count.index.values, Sentiment_count['Review'])
      plt.xlabel('Review Sentiments')
      plt.ylabel('Number of Reviews')
      plt.show()
```

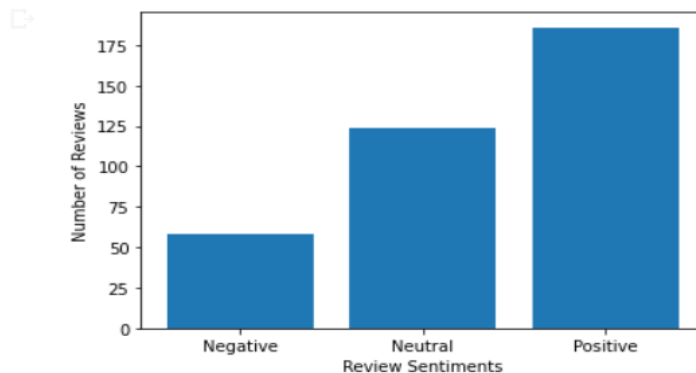


Fig. 12: Student's Review in Bar-Chart

Fig. 12 displays the bar chart of the student's different review based on their sentiment. As we can see the horizontal line shows the types of review which we have categorized to analyze the sentiment of the students. On the other hand, the vertical line shows the number of reviews for a particular type of sentiment of all students.

4.2.1 Random Forest Algorithm:

A well-known supervised machine learning algorithm for Classification and Regression problems is the Random Forest Algorithm. We as a whole realize that backwoods have many trees, and the more trees there are, the more grounded the timberland will be. Similarly, a Random Forest Algorithm's accuracy and ability to solve problems increase with the number of trees it contains. Irregular Woodland is a classifier that utilizes the normal of a few choice trees on various subsets of a given dataset to work on its prescient exactness. The idea of ensemble learning, in which multiple classifiers are combined to solve a complicated problem and boost the model's performance, serves as its foundation [6].

4.2.2 Working Procedure:

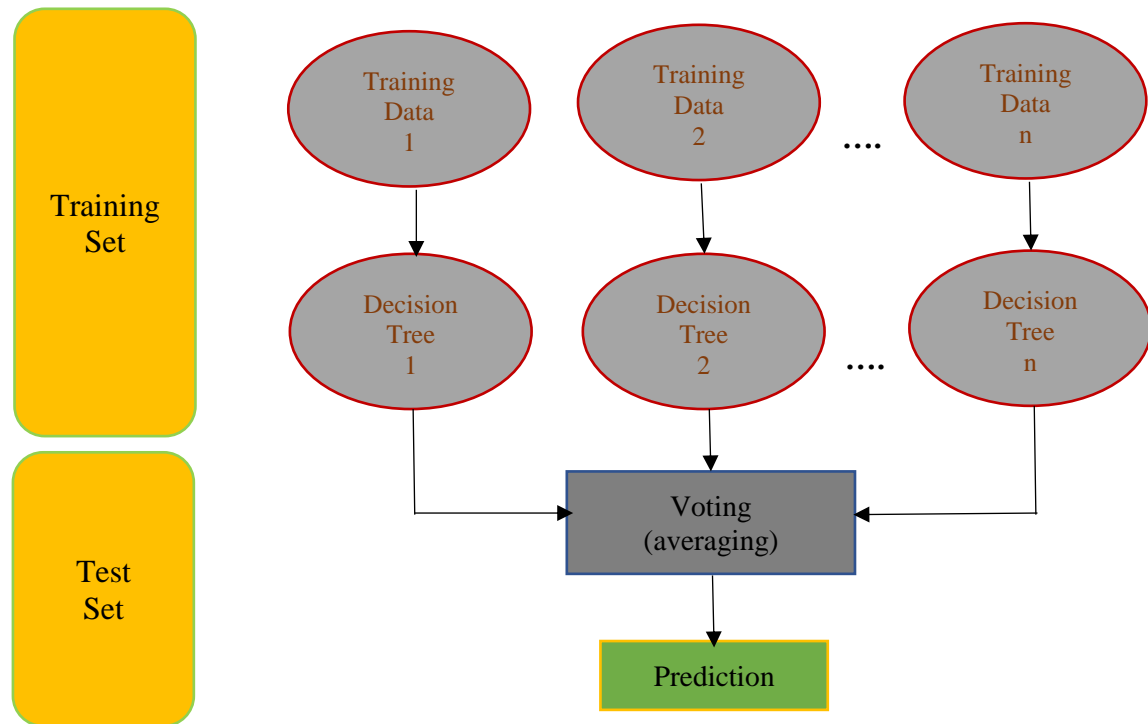


Fig. 13: Working Procedure of Random Forest Algorithm

Fig. 13 exhibits the working procedure of the Random Forest Algorithm.

The hyperparameters of a random forest are nearly identical to those of a decision tree or a bagging classifier. Fortunately, the random forest's classifier-class makes it unnecessary to combine a decision tree and a bagging classifier. Using the regressor of the algorithm, you can also deal with regression tasks with random forest. While the trees grow, random forest adds even more randomness to the model. When splitting a node, it searches for the best feature among a random subset of features rather than the most important feature. As a result, there is a lot of diversity, which usually leads to a better model [7].

As a result, the algorithm for splitting a node in a random forest only takes into account a random subset of the features. Instead of searching for the most optimal thresholds, as is the case with a standard decision tree, you can use random thresholds for each feature to make trees even more random.

4.3 Model Generation and Evaluation

In this section, we will generate our model using the Random Forest classifier. Then fit the train dataset into this model. Consequently, we will evaluate the performance of our Random Forest model. **Figure 14** displays the generation of our Random Forest classifier and the accuracy of predicting the sentiment of the students.

Model Generation

```
[31] from sklearn.ensemble import RandomForestClassifier
      from sklearn.model_selection import cross_val_score

      rf = RandomForestClassifier()
      scores = cross_val_score(rf,X_train,y_train.values.ravel(),cv=5)
```

Model Evaluation

```
[35] print(scores)
      scores.mean()

[0.75      0.89090909 0.83636364 0.87272727 0.81818182]
0.8336363636363636
```

Fig. 14: Model Generation & Evaluation

After model generation with Random Forest algorithm and then we train the model. We have analyzed the model using test data find accuracy of our model and algorithm. Here, the accuracy of the Random Forest algorithm is 83.36%.

Name of Algorithm	Accuracy
Random Forest	83.36%

Table 1: Accuracy of Random Forest Algorithm

Table 1 shows the accuracy of the Random Forest algorithm to predict the sentiment of the students. So, it can be said that this model can predict a students review whether it is positive, neutral or negative with the accuracy of 83.36%.

CHAPTER 05

CONCLUSION AND FUTURE WORK

5.1 Summary of the Study

The campus of Daffodil International University recently moved from Dhanmondi to Ashulia. Consequently, the students displayed a variety of reactions. Based on the various responses of students to campus changes, a sentiment dataset was created. With this dataset, we made some models using machine learning algorithms, and we found that Random Forest was the best for this sentiment analysis process. With this algorithm, the sentiment analysis of students can be easily and accurately done. The sentiment of a particular student based on his/her review can be predicted with high accuracy of 83.36%.

5.2 Conclusion

The process of determining the text's emotional tone is known as sentiment analysis. In this study, sentiment analysis was carried out with the help of student reviews. We gathered responses from Daffodil International University students regarding the campus relocation from Dhanmondi to Ashulia for this study. For a variety of reasons, some students provide constructive feedback, while others provide constructive feedback. Our principal objective is to track down the best outcome with the most elevated exactness. Because of this sentiment analysis, students can easily express their opinions about the campus change. This sentiment analysis can also be used by the authority to ascertain the necessary requirements and the overall outcome of the university environment. Moreover, authority can take necessary steps to improve the sector where student's review was negative.

5.3 Future Work

The shifting of Daffodil International University's campus from Dhanmondi to Ashulia is the primary focus of our study. We gather the opinions of the students and use them to create a data set to get a sense of their feelings. We then use that data set to find the best algorithm that gives us better, accurate results. Therefore, it is simple for anyone to ascertain how the students are feeling. So, if the authority notices problems among the students, they can take their clues and follow them to solve them. Using these pre-built models, anyone can quickly and easily identify any kind of sentiment in their collected data set and achieve better results.

APPENDIX

LIST OF ABBREVIATION

DIU – Daffodil International University

CSE – Computer Science and Engineering

ML – Machine Learning

RQ – Research Question

SA – Sentiment Analysis

REFERENCES

- [1] V. A. Kharde and S. Sonawane, "Sentiment Analysis of Twitter Data: A Survey of," *International Journal of Computer Applications (0975 – 8887)*, vol. 139, no. 11, pp. 5-15, 2016.
- [2] P. Baid, A. Gupta and N. Chaplot, "Sentiment Analysis of Movie Reviews using Machine Learning Techniques," *International Journal of Computer Applications (0975 – 8887)*, vol. 179, no. 7, pp. 45-49, 2017.
- [3] G. S. Brar and A. P. A. Sharma, "Sentiment Analysis of Movie Review Using Supervised Machine Learning," *International Journal of Applied Engineering Research*, vol. 13, no. 16, pp. 12788-12791, 2018.
- [4] F. Neri, C. Aliprandi, F. Capeci, M. Cuadros and T. By, "Sentiment Analysis on Social Media," in *International Conference on Advances in Social Networks Analysis and Mining*, Istanbul, Turkey, 2012.
- [5] S. MUKHERJEE, "Introduction to sentiment analysis," Kaggle, 2021. [Online]. Available: <https://www.kaggle.com/code/shyambhu/introduction-to-sentiment-analysis>. [Accessed 12 December 2022].
- [6] Simplilearn, "Random Forest Algorithm," Simplilearn, 27 March 2022. [Online]. Available: <https://www.simplilearn.com/tutorials/machine-learning-tutorial/random-forest-algorithm>. [Accessed 19 December 2022].
- [7] N. Donges, "Random Forest Classifier: A Complete Guide to How It Works in Machine Learning," Builtin, 22 September 2022. [Online]. Available: <https://builtin.com/data-science/random-forest-algorithm>. [Accessed 21 December 2022].

Sentiment

ORIGINALITY REPORT

22%

SIMILARITY INDEX

20%

INTERNET SOURCES

10%

PUBLICATIONS

0%

STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	17%
2	Mona Minakshi, Pratoool Bharti, Sriram Chellappan. "Identifying mosquito species using smart-phone cameras", 2017 European Conference on Networks and Communications (EuCNC), 2017 Publication	1%
3	Suvarna Patil, Neha More, Animesh Bhawtankar, Vishal Pratap Jagtap, Anjali Jadhav. "chapter 6 Artificial Intelligence and Machine Learning-Based Advanced Depression Detection System", IGI Global, 2022 Publication	1%
4	Tuhin Hossain, Ahmed Ainun Nahian Kabir, Md. Ahasun Habib Ratul, Abdus Sattar. "Sentence Level Sentiment Classification Using Machine Learning Approach in the Bengali Language", 2022 International	<1%