



Daffodil
International
University

**Sentiment Analysis Through Evaluation of logistics support company
Customer Reviews**

Submitted By

MD RABIUL AWAL SHUVO

ID: 193-35-487

Department of Software Engineering

Daffodil International University

Supervised By

Mr. Md. Rittique Alam

Lecturer

Department of Software Engineering

Daffodil International University

This Report is Presented in Partial Fulfillment of the Requirements for the Degree
of Bachelor of Science in Software Engineering.

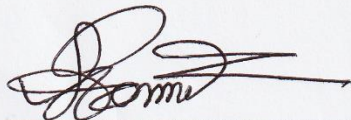
Spring – 2023

Copyright © 2023 by Daffodil International University

APPROVAL

This thesis titled on “Sentiment Analysis Through Evaluation of logistics support company Customer”, submitted by **MD RABIUL AWAL SHUVO (ID: 193-35-487)** to the Department of Software Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Software Engineering and approval as to its style and contents.

BOARD OF EXAMINERS



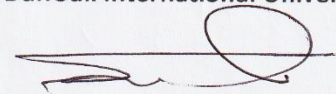
Nuruzzaman Faruqui
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Chairman



Khalid Been Budruzzaman Biplob
Lecturer (Senor Scale)
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 1



Md Khaled Sohel
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 2



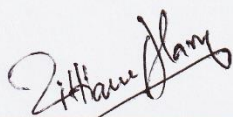
MD. Tanvir Quader.
Solution Architect Engineer,
Technology Team, a2i Programme

External Examiner

DECLARATION

I hereby declare that, this thesis report is done by me under the supervision of **Mr. Md. Rittique Alam**, Lecturer, Department of Software Engineering, Daffodil International University, in fulfillment of my original work. I am also declaring that according to the best of my knowledge, neither this thesis nor any part therefore has been submitted else here for the award of M.Sc. or any degree.

Supervised by:



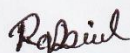
Mr. Md. Rittique Alam

Lecturer

Department of SWE

Daffodil International University

Submitted by:



MD RABIUL AWAL SHUVO

ID: 193-35-487

Batch: 30th

Department of SWE

Daffodil International University

ACKNOWLEDGEMENT

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

Then I really grateful to my research supervisor, **Mr. Md. Rittique Alam** who guided me throughout the whole research activities.

I wish to express my special thanks to **Dr. Imran Mahmud**, Head of the Faculty for providing all the necessary facilities for the research purpose. I am also thankful to all the lecturers, Department of Software Engineering who sincerely guided me at my difficulty. I am thankful to my friend who supported me throughout this venture.

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

ABSTRACT

Abstract: The development of modern logistics has become a powerful and powerful enterprise for the construction of smart cities. Supervision of fine suppliers can help change sales, and the development of logistics companies can improve the city's public goods. In this study, we plan an optimal assessment of logistics service providers mainly through a combination of real-time analysis perspective and traditional assessment perspective. With the help of sensitivity analysis, this version aims to extract male or female characteristics of service delivery quality from user opinions and to obtain information about customers' use of certain products. By improving the SERVQUAL version, we are creating a new logistics product quality measurement system. The tool used the assessment tool to my content rating data by removing features and creating family members in content rating and rating. We also use opinion-based evaluation, with features shown to clearly compare employees' world-class work. Experimental analysis shows that the developed models and methods are very accurate.

Keywords: smart city; sentiment analysis; logistics service quality; feature extraction

TABLE OF CONTENTS

CONTENTS	PAGE
Approval	ii
Declaration	iii
Acknowledgments	iv
Abstract	v
CHAPTER 1 : INTRODUCTION	1-4
1.1 Introduction	1
1.2 Motivation	2
1.3 Rationale of the Study	2
1.4 Expected Output	2
1.5 Research Questions	3
1.6 Report Layout	3
CHAPTER 2: BACKGROUND	5-9
2.1 Related Works	5
2.2 Comparative Analysis and Summary	5
2.3 Research Summary	7
2.4 Scope of the Problem	8
2.5 Challenges	8
CHAPTER 3: RESEARCH METHODOLOGY	9-19
3.1 Introduction	9

3.2 Data collection	10
3.3 Data Preprocessing	12
3.4 Statistical Analysis	14
3.5 Applied Mechanism	14
CHAPTER 4: EXPERIMENTAL ANALYSIS AND DISCUSSION	19-31
4.1 Introduction	19
4.2 Experimental Analysis	19
4.3 Experimental Result	25
4.4 Summary	30
CHAPTER 6: IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY	31-33
5.1 Impact on Society	31
5.2 Ethical Aspects	32
5.2 Impact on Environment	32
5.4 Sustainability	33
CHAPTER 6: CONCLUSION & FUTURE WORK	34-35
6.1 Conclusion	34
6.2 Recommendation	34
6.3 Future Work	35
REFERENCES	36-37

LIST OF FIGURES

Figure Name	Pages
Figure 3.1: Types of the sentiment in logistics review dataset	11
Table 3.2: Dataset demo table for logistics review.	12
Figure 3.3: found emoji in logistics review dataset.	13
Table 3.4: Clean Dataset table logistics review.	14
Figure 3.5. LSTM Architecture	15
Figure 3.6. CNN Architecture	16
Figure 3.7: Random forest classification	18
Figure 4.1: Dataset	19
Figure 4.2: import library	20
Figure 4.3: Remove emoji	21
Figure 4.4: Final data after preprocessing	21
Figure 4.5 Embedding data	21
Figure 4.6: Word cloud	22
Figure 4.7: word count	22
Figure 4.8: separate all word	23
Figure 4.9: Label Encoding	23
Figure 4.10: Features columns 1	24
Figure 4.11: Features columns 2	24

Figure 4.12: Model Summery	25
Figure 4.13: SNN model accuracy	26
Figure 4.14: SNN model accuracy Graph	26
Figure 4.15: CNN model accuracy	27
Figure 4.16: CNN model accuracy Graph	27
Figure 4.17: LSTM model accuracy	28
Figure 4.18: LSTM model accuracy Graph	28
Figure 4.19: Logistics Regression model accuracy	29
Figure 4.20: Logistics Regression F1 Score	29
Figure 4.21: Logistics Regression Heatmap	30

LIST OF Tables

Tables Name	Pages
Table 2.1 Comparative Analysis and Summary	7
Table 4.1: Different Algorithm results	31

CHAPTER 1

INTRODUCTION

CHAPTER 1

INTRODUCTION

1.1 Introduction

As the infrastructure at the bottom degree of society, logistics not exceptional connects manufacturing and consumption but also supports the pillar industries and extraordinary industries of the city. currently, the rapid popularity of purchasing has extensively. Promoted the improvement of the logistics issuer enterprise. Logistics offerings, which are vital components of online buying, are without delay associated with consumer satisfaction and affect purchase intentions [1,2]. To awareness on improving their non-public carrier stage and to attract extra customers, many e-trade agencies usually outsource their logistics offerings to one/3-party agencies due to the truth they may be now not middle aggressive. but, customers have the propensity to treat the whole way of purchasing as a whole, and their attitude in the route of the fine of logistics offerings without delay impacts their attitudes inside the route of the issuer and commodity pleasant, thereby affecting the earnings of commodities [3]. Consumers' attitudes closer to logistics service have ended up being a vital element affecting the improvement of e-trade.

In the current logistics service, the competition between logistics service companies is very fierce and the price structure of logistics services is becoming more and more the same. Choosing a high-service shipping company is a concern of customers [4,5]. To build trust with customers, most e-commerce businesses pay close attention to customer feedback on quality delivery. Most studies allow users to evaluate the quality of logistics services by creating a benchmarking system. According to the evaluation model [6-9], the quality level of the logistics service can be evaluated by analyzing the scores of the users and appropriate logistics service can be offered to the customers. However, good logistics services have many aspects, and different users will have different experiences. Index-based evaluation can only score and evaluate the quality of logistics services. Chapter However, it is difficult to use certain factors to explain the advantages and disadvantages of users Chapter. Sentiment analysis, also known as sentiment mining and directional analysis, is an important branch of natural language processing. The main function of analytical thinking is to assist people who use computer resources to obtain, organize, and analyze information about the subject, including intelligent analysis of data files, processing, writing content and reasons [10]. Methods based on sentiment analysis have been used in many fields such as social networking [11], e-commerce [12], online demographic analysis [13], and others. In the operation of logistics services, e-commerce platforms have collected a lot of customer evaluation data on the quality of logistics services. Reviews often describe various aspects of logistics services

in the literature that can affect the quality. It is difficult to manually process data on such a large scale; therefore, automated machine analysis methods should be used. With the help of thinking technology research, the emotional characteristics of the words related to logistics service quality can be extracted to help the platform analyze user behavior regarding various logistics services and explain the quality of logistics services. This method is very important to help users choose the best logistics service and improve the service level of businesses.

1.2 Motivation

In the process of uncovering reviews and whether your product is getting good or bad reviews, social media research can provide a lot of additional information about your product or service. Use this tool to gather and analyze public opinion about your cause to determine if they are positive or negative for your cause. Just because your post or video receives positive or negative reviews does not mean that the majority of comments are positive or negative or positive. Content on many social media platforms is smaller and less complex than blog content and news media, Twitter's content is also popular because of typos and slang, as well as some of the special characters we listed in the first stage. little. Based on our findings, we believe that more research is needed in this area and we did not find many research articles that directly address the topic of gathering user opinions and thoughts from Twitter entries. Our findings should allow us to infer the level of opinion or disagreement that users express when sending messages or "tweets" based on our test results.

1.3 Rational of the study

Social media also allows people to connect with each other through electronic sites and applications. Facebook, Twitter, and LinkedIn are places for personal-to-person communication. The importance of shared fun is still in the sand, because its results so far are important new developments necessary for rapid change. Predictive analysis is an analytical technique used in research and language planning to analyze different types of information or information from a fragmented text. Analysts have done a lot of work on this in recent years. Many online services, such as e-come, are slowly becoming popular with people. We can feed our decision tables on the web differently, and sometimes people search based on that view. These studies are often based on irrelevant data that has not been studied further. This control is based on eliminating unnecessary data to understand customers' behavior and interactions on online platforms. First, we collect data from websites to do this test.

1.4 Research Questions

- Why Sentiment Analysis is Important?
- Why we need logistic support company
- How to use sentiment analysis in Business purpose?

1.5 Expected Output

Social media organizations also allow people to connect with each other through electronic sites and applications. Facebook, Twitter, and LinkedIn are places for personal-to-person communication. The importance of connecting organizations like sand is still sand, and for what it's worth, Twitter has come so far for users to analyze and view the audit dataset organization by crying out information, but so far there's no better answer than asking many. the number of tweets that left room for investigation. This theory is based on the difference between the expression of different writing strategies and the creation of another final statement to control tweet demand about shipping companies. Four methods are presented in this report, including logistic regression, RNN, LSTM, simple RNN, and a method involving random forest classifiers. The social disclosure method is to use the majority of the voting process to determine the results of the game plans of the four classifications and select the final evaluation. It provides an assessment of the different ideas of theory, processes and experiments. For the approach to AI's final game plan, the result is a plan that is modeled after a decision has been made. However, for some systems, such as dictionary-based philosophies, the AI system is not included and does not require planning. I'm talking about success in my business here. Here, the ML representation computation is used to represent the hypothesis, the estimation depends on the skewness of the statistical analysis. They are characterized by the fact that the number represents positive or negative emotions, but does not contradict this. In recent years, analysis areas such as skewness detection, test mining, hypothesis mining and skewness extraction have received more attention. Apparently, online reviews have become an important standard for judging the nature of business. This article provides a comprehensive study of methods of conducting market surveys using the large survey data provided by the Yelp dataset. In this study, we present a method for visual representation tasks using two inference strategies and four cognitive models. A similar report on the feasibility of an opinion-based audit illustrates this. For rapid change, huge new developments are needed. Predictive analysis is an analytical technique used in research and language planning to analyze different types of information or information from fragmented texts. Analysts have done a lot of work on this in recent years. Many online services, such as e-come, are slowly becoming popular with people. We can feed our decision tables on the web differently, and sometimes people search based on that view. These studies are often based on irrelevant data that has not been studied further. This control is based on eliminating unnecessary data to understand customers' behavior and interactions on online platforms. First, we collect data from websites to do this test.

1.6 Report Layout

The derivative of this agreement is as follows. The first section describing the motivation introduces the purpose and formulation of the proposal. Chapter 2 discusses the relevant study and presents the main methods of comparing related studies. Part III presents data refinement, preliminary data, and exhibits. In area 4, the motivation for product evaluation is explained and

the results are analyzed. District Fifth organizes three assessments, official assessments and inspections. As in 6, the end is set and my role is drawn.

CHAPTER 2

BACKGROUND STUDIES

CHAPTER 2

BACKGROUND STUDIES

2.1 Introduction

Text analysis is also a subfield of data collection. However, the product identification problem has some similarities with the data collection process. Standard documents often require apps to track numbers or facial features, but messaging games make apps to track content that connects to text, words, or voices. The most common way to use the input method for content requests is to convert the physical data into digital data and then use the descriptive data. For example, we can turn each word occurrence in the search data into an attribute and each data entry into an attribute vector that represents the word's occurrence in the document. However, the size of the existing transfer case may be too large for the identification task. Of course, even data research products can have over 1,000 arguments such as speech and longer grammar. The main operating system is used in many areas such as spam exclusion, e-mail return, message validation, content collection and request evaluation. Due to the development of electronics and education, the electronic data file has become so large that it cannot be measured accurately. It is problematic and opens the door to the development of language processing techniques such as content applications. Data acquisition systems can use real or probabilistic numbers to create large electronic datasets and perform dynamic operations. Special attention is required in the system. We select the features in the product information when evaluating the composition, remember that the calculated decision depends on the purpose of the product profile. The dimensionality of the application data can be reduced to an appropriate size by selecting the main features that make up the task. The product description and end-of-action class were highly sought after, but Twitter's evaluation criteria for product review data were different. In addition to using the requested suggestions to manage tweets for product reviews, it would be interesting to create other ways to manage the development of the reality of the game plan.

2.2 Related Works

This chapter focuses on user sentiment mining research, especially for sentiment analysis in online services. Various machine learning studies and dictionary-oriented methods were also researched and analyzed to form the basis of this study.

2.2.1 Customer Opinion Mining in online services

The development of technology and the widespread use of social media have created opportunities to obtain information from unstructured data. Opinion mining in big data is used to classify the opinions of customers with different opinions and to measure the needs of people. At the same time, thought mining has produced surprising results according to many reviews on the web. Customers, restaurants, schools, hospitals, resorts, etc. where they share their views on products and services. The value of user reviews, comments or ratings for a particular product or service is their opinion, judgment, opinion or opinion about their quality, appearance or price. Thoughts can be positive or negative depending on the person's point of view.

One of the limitations of machine learning-based methods is that it depends on the size of the training data, which must be tagged and large enough. However, registration information is often scarce, especially in some narrow specialties. Research groups often have to spend time and money collecting data.

2.2.2 Machine learning-based customer sentiment analysis

Sentiment and thoughts are the subject of interest and research for many researchers (Akter et al., 2016; Lugović et al., 2016). Therefore, there is a difference of opinion. Depending on the nature of the emotions, emotions can be divided into two groups: positive emotions and negative emotions. Depending on how and where we express them, we can divide emotions into six categories: happiness, sadness, anger, surprise, disgust, and fear. Under the influence of different situations, different environments, and different emotions, people's emotions sometimes get mixed up, mixed, and combined. This makes its owner the owner of other thoughts.

In general, emotional analysis is defined as "the scientific study of thoughts, feelings and emotions represented in texts" (Nagpal et al., 2020). In other words, opinion mining, as a way to get the opinions of the people who created a particular document, has recently become the most popular research topic in relation to relationship (Pang et al., 2008; Ohana et al., 2009). With the rise of social media such as reviews, forums, and social media, opinion polls have grown in importance.

2.2.3 A classification algorithm

SVM is a machine learning taxonomy that uses kernel functions to map the location of nonlinearly allocated data points to new locations with misclassifications. We refer the reader to Burges (1998) for an explanation of SVM and its design principles. Detailed information on the use of this model for text classification can be found in Joachims (2002).

SVM is essentially an optimization problem; The purpose of the algorithm is to find the location of F and the overall decision f that minimizes the classification error of F . Let the model set $\{(x_1, y_1), (x_2, y_2), \dots, (x_f, y_f)\}$ and $x_i \in \mathbb{R}^n$ in two ways as follows: $y_i \in \{-1, 1\}$, class anchor tag x_i (-

1 for class I, 1 for class II). We have the vector x_i in space of the equation hyperplane: $x_i \cdot w + b = 0$

$$\text{Set } f(x_i) = \text{sign}(x_i \cdot w + b) = \begin{cases} +1, & x_i \cdot w + b > 0 \\ -1, & x_i \cdot w + b < 0 \end{cases}$$

Table 2.1. Comparative Analysis and Summary

Reference	Algorithm / Classifier	Data collection	Accuracy	Comparism Proposed Methodology
Van-Ho Nguyen	SNN	Social Media	99%	SNN 0.99%
Fahrettin HORASAN	CNN	Website	0.875%	CNN 0.85%
Chao Wei	LSTM	websites	90%	LSTM 90%,
Dingkai Zhang	Logistic Regression	website	0.79%	Logistic Regression 0.79%

2.3 Research Summary

Product delivery has been a major concern for the rapidly growing online business over the past two decades. Product organizations use customer analysis models, so these models are very boring and boring. Facebook, Twitter, and website databases can be great places to gather customer insights and conduct sensitivity tests. In this paper, we trace data including analysis of the main supply chain and perform multivariate analysis. The concept begins with a preplanned process to clean up words that are then used as vectors to assess the level of clarity using basic learning techniques. Analysis was performed using 4 special algorithms in Random Forest, Logistic Regression, Simple RNN, LSTM and RNN. The system is set to use 80% of the data and tries to

use 20% of the remaining data. According to the results, the accuracy of the evaluation of each collection method was determined and it was found that all evaluations of the delivered products could be combined to predict all.

2.4 Scope of the Problem

This research presents and discusses tests to evaluate the hypothesis that Twitter exposure is similar to posting research. The purpose of this test is to determine whether reviews can be claimed to present positive, negative opinions. Online sector allows users to communicate and share their records, thoughts and predictions through joint communication. Operator reviews are useful and are used as data to review customer concerns. In this paper, experts build models by identifying common constraints using techniques such as classifiers, LSTMs, Simple RNNs, Logistic Regression, and Random Forest. Experts wrote reviews on the latest publications of their respective organizations. Experts also chose the customer review to be clear, unbiased, negative, and analyzed more, not like the final review, to anticipate its benefits.

2.5 Challenges

Our plan shows that preventive measures and over-monitoring are important in improving our results. The use of custom decision methods reproduces the best feature set and reduces the amount of information required to construct a classification. But it reduces the skewed position of classes in many small datasets without causing too much interference. Our results confirm that the proposed model achieves high accuracy in predicting the three event classes (good, bad). It can be seen that using some classifiers beats others. For example, random forests, logistic regression, RNNs, and LSTMs exhibit greater reliability and security when applied to all data. While others have had commendable success on all measures. Finally, after many variations and selections, we come to a point where the model can be chosen well. For this reason, we can say that the most difficult and time-consuming task related to this study is the data collection and selection proce

CHAPTER 3

RESEARCH METHODOLOGY

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

On this phase, we consider the study in details. Each test can be solved with different strategies. Collecting information online is our first step. For our research, it is necessary to remove some redundant values and data arrays after processing the data. Then choose a machine learning algorithm. Before we build the model and fit the algorithm, we need to build the data because we've determined that we're using four different types of machine learning. Then use this information to specify the model. This is how feature selection works.

There is no limit to the nuance that posts and messages can bring, and these little words are always used to give people the opportunity to understand and appreciate the whole environment. Thus, Twitter, Facebook trend corpus efforts and advances are designed to support research that supports a unique understanding of transmission evaluation in words and messages. The use of such electronic devices presents problems for the management of business language, because the language used is emotional, thinking about writing, and more importantly, writing bad letters, slang, neologisms, URLs and clear phrase patterns and abbreviations. Other types of electronic telecommunication of personal information, such as Twitter messages, contain a wealth of information about the participants in the conversation.

Subject of study and research components:

Hardware and Software:

- Intel Core i5 8th generation
- 1 TB Hard Disk Drive
- 256GB m.2 SSD
- 12 GB RAM
- 6 GB GPU

Development Tools:

- Windows 10

- Jupyter Notebook
- Google Colab
- Python
- Pandas
- NumPy
- Matplotlib
- Seaborn
- Sklearn
- Scikit-Learn

3.2 Data collection

The collection of data is the initial step in the pre-processing process. As a basis, I began by gathering information through the Facebook Review section, Logistics department website review. I am collected around 10,000 data through different logistics website.

Following that, each process must use the dataset. Since each of the algorithms we use behaves differently, as we have already demonstrated, we must preprocess the data in order to make it fit into our model correctly.

Dataset columns name are:

- **Sentiment**
- **Review**

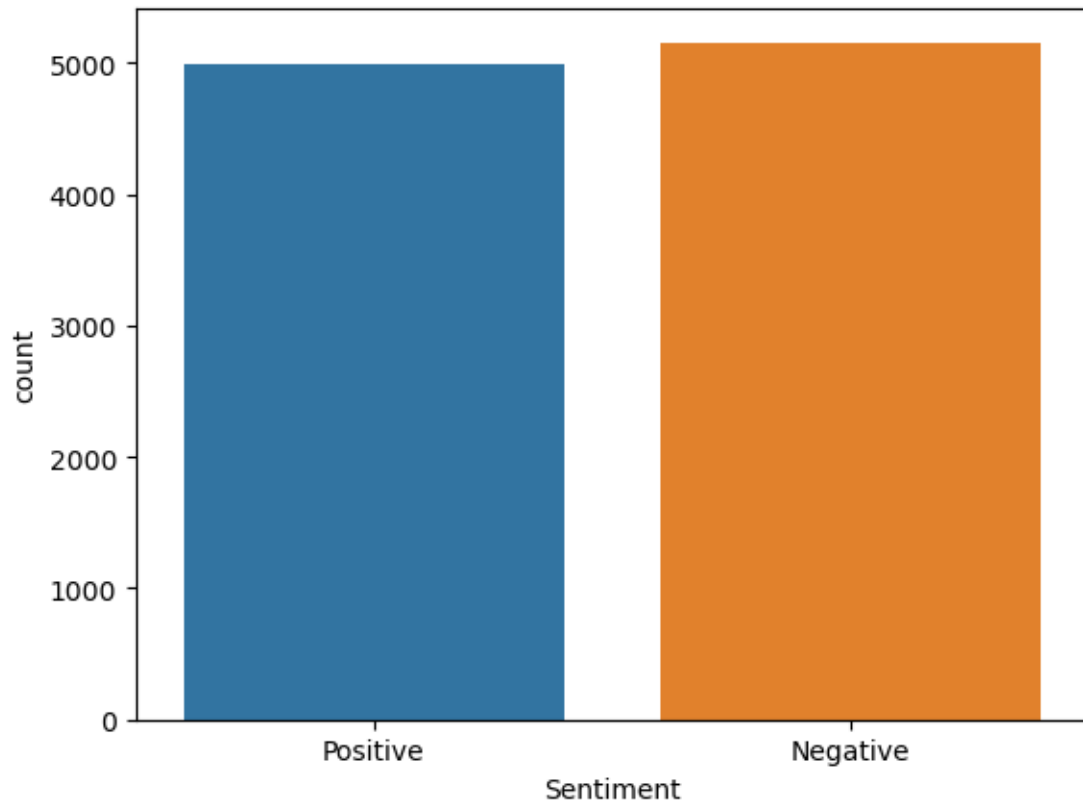


Figure 3.1: Types of the sentiment in logistics review dataset.

I deliver a demo of the dataset underneath in table 2.2.3. Text segment contains crude comments in the dataset.

	Review	Sentiment
0	Good and fast service And Recommend	Positive
1	Disgusting service, i Hat this courier company	Negative
2	Very very bad service	Negative
3	Excellent service	Positive
4	Their service is so good I am satisfied	Positive
5	Very bad service Specially for home delivery	Negative
6	worst one no customers services, rude behaviour	Negative
7	its make a important rule in our life	Positive
8	need to improve customer service	Positive
9	Worst service ever Their employees are illite...	Negative
10	Worst service ever Every single employee is a...	Negative

Table 3.2: Dataset demo table for logistics review.

3.3 Data Preprocessing

3.3.a Preprocessing

The data preprocessing stage is one of the most important points in the study. In this way, working with pre-processed data can provide much better performance than processing the data in raw form. In general, information on the Internet is popular. In this case, if the data is not the same, the requested operation cannot be completed. Some areas may require data deletion, error correction, data reduction and calibration. Depending on the data, noise normalization and appropriate adjustments may be required for the preferred model to perform well. In this case, regular expression techniques are used first to remove unwanted characters from the analyzed data. Also, duplicate charges are removed. It is then checked using the NLP library and the noise quality is checked. In this data have some emoji, Hashtag, colon, and many more we need to drop and preprocess this for uppercase letter to lower case letter. There I am found some emoji which are not predictable.


```
max(tokenizer.word_index)
```

'😡😡'

Figure 3.3: found emoji in logistics review dataset.

3.3.b Regular expression

Overhauled full support for Perl-like expressions in Python in the Python module. re module re upgrades private data. An error occurs when an error occurs while writing or using a translation model. I will discuss two important constraints that govern sentence coherence. But first, an easily overlooked detail: There are many types of Chinese characters, and they are the most important when they use language patterns.

3.3.c Stop word remove

Stop words are an ordinarily utilized word that a web list has been revamped to ignore, both when mentioning fragments for glancing and remembering that recovering them as the postponed outcome of a fascinating solicitation. I would not need these words to devour room in our information base, or including the colossal managing time. For this, I can clear them sufficiently, by dealing with an outline of words that we consider to stop words. NLTK in python has a quick overview of stopwords put aside in 16 unique vernaculars. We can discover them in the nltk library. Words to block is a commonly used word where network names are changed to be ignored, both when referencing snippets for scanning and remembering that editing them is a delay in request expansion. I don't need these messages to take up space on our storage or require too much admin time. I can clear them up enough by making explanations for what we think are abandoned words. NLTK in Python provides a quick overview of stop messages in 16 custom messages. We can find them in the nltk library.

Finally, we found our clean and raw data for predict accuracy.

	Review	Sentiment	liked
0	good fast service recommend	Positive	1
1	disgusting service hat courier company	Negative	0
2	bad service	Negative	0
3	excellent service	Positive	1
4	service good satisfied	Positive	1

Table 3.4: Clean Dataset table logistics review.

3.4 Statistical Analysis

1. There are 2 columns in total
2. There are 10153 rows in total
3. 80% of the data were used to train and 20% to test our model.
4. The dataset is saved as a csv file.

3.5 Applied Mechanism

3.5.a Simple Neural Networks

Simple Neural networks are an artificial intelligence method that teaches computers to process data inspired by the human brain. It is a machine learning process called deep learning that uses cells or neurons connected in a hierarchical structure similar to the human brain. It creates an adaptive system that computers can use to learn from their mistakes and continually improve. Therefore, neural networks try to solve complex problems such as gathering information or recognizing faces with higher accuracy. A simple neural network has three layers of artificial neurons interconnected:

Input Layer: Information from the outside world enters the neural network through the input layer. Input nodes process, analyze or classify data and transmit it to the next layer.

Hidden Layers: hidden layers provide input from the input layer or other hidden layers. Neural networks can have many hidden layers. Each layer analyzes the output of the previous layer, processes it further and forwards it to the next layer.

Output Layer: Output operation provides the final result of all data processed by the neural network. It can be single or multiple. For example, if we have a binary classification problem (yes/no), the output operation will output a 1 or a 0. However, if we have a multi-class classification problem, the output method has multiple outputs.

3.5.b Long Short-Term Memory (LSTM)

LSTM networks are deep learning neural networks that allow continuous data processing. It is a special type of neural network that can solve the gradient problem solved by RNNs. LSTM was developed by Hochreiter and Schmidhuber to solve problems caused by traditional RNN and machine learning algorithms. LSTMs can be implemented in Python using the Keras library. For example, you remember the previous event while watching a video, or you know what happened in the previous chapter while reading a book. RNNs work similarly; they remember past information and use it to form current ideas. The disadvantage of RNNs is that they cannot remember long-term prospects due to loss. LSTMs are specifically designed to avoid long-term dependency issues.

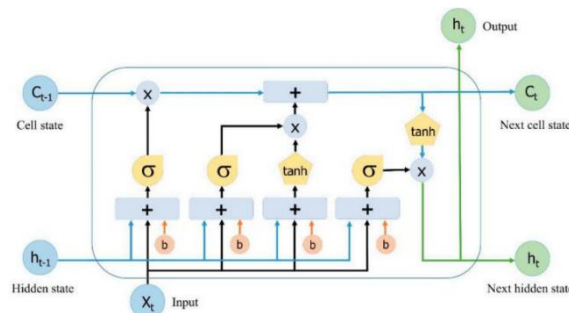


Figure 3.5. LSTM Architecture

The RNN approach is embedded in LSTM by adding LSTM cells to the RNN architecture. LSTMs take longer to calculate than RNNs because LSTMs store weights. On long data, LSTM may outperform RNN because the LSTM cell has a self-loop. There are three main elements in the LSTM processing stage: the memory gate, the input gate, and the output gate. Figure 3 is the LSTM architecture.

The formula for the forget gate (f_t) is written in formula 1.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f).$$

Furthermore, the two parts are combined to update the

$$C_t \cdot i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

For the formula to update the old cell state (C_{t-1}) to the new cell state (C_t) in formula 4.

$$c_t = (i_t \cdot \tilde{C}_t + f_t \cdot c_{t-1})$$

Second, pass cell state (C) in formula 6 through \tanh to produce a value between -1 and 1. Then multiply it by the sigmoid gate, resulting in the decided value.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t \cdot \tanh(c_t)$$

3.5.c Convolutional Neural Networks (CNN)

Convolutional neural networks were first used in image processing or computer vision. In 2015, research was conducted on Natural Language Processing (NLP) using CNNs to classify text. Apply convolutional techniques to sentences, paragraphs or whole text for text classification using CNNs. Deconvolution techniques work by splitting a matrix of text representations into windows or filters. Filters are added together to create a new representation of the text, which can be called a custom map. The maximum value is derived from each specification using the maximum pooling technique. The CNN architecture is shown in Figure 4, including sentence representation, convolutional layers, maximum pooling, full coupling, dropout, and SoftMax. Sentence representation is the sentence entered in matrix form, this matrix is used as input to the convolution layer. Convolutional layers are made with filters that create feature maps.

Convolutional layers also set parameters. The most important thing is the number of cores and the core size.

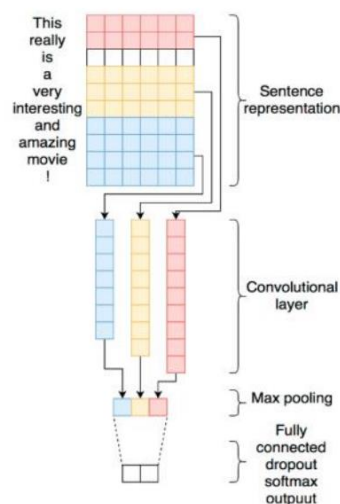


Figure 3.6. CNN Architecture

n of the data. SVM is recognized for working well with problems that have a lot of characteristics and little training data. SVM may, however, be memory and computationally-intensive for large datasets. The parameters and choice of the kernel may also have an impact on SVM. Numerous practical applications of SVM are available, such as bioinformatics, text classification, and image classification. SVM may be used with other machine learning algorithms to improve the performance of the system. KNN (K-Nearest Neighbors) is an illustration of a supervised learning algorithm used for classification and regression analysis. KNN is a non-parametric algorithm, therefore it doesn't base its decisions on how the data are distributed. The Manhattan distance or the Euclidean distance are two distance measures that the KNN technique uses to identify the K data points that are closest to the query point. The K nearest neighbors are then used to identify the class or value of the query point. A majority of the query point's K nearest neighbors determine how to classify it. If $K = 5$ and the three nearest neighbors are in class A and the other two are in class, the query point is placed in class A, for example. In a regression, the query point's value is the mean or median of the K nearest neighbors. KNN has the advantage of being clear-cut and easy to understand. When the training set contains a lot of data as well as when the dataset is little, KNN may be effective. KNN may, however, be sensitive to the data distribution, K number of neighbors, and the choice of distance measure. KNN provides a wide range of practical applications, such as text and picture classification, recommendation systems, and anomaly detection.

3.5.d RANDOM FOREST

The random forest distribution method of co-learning uses multiple decision trees to improve the robustness and accuracy of the model. Random Forest is a supervised learning method for classification and regression analysis. Section Harassment Each decision tree is constructed from a variable feature set and a random subset of training examples. As a result, overfitting is reduced and the capability of the model is improved. When a new sample is wanted to be classified, it is sent to all decision trees in the forest and the most supported class is selected as the protected class for the sample limit. There are many advantages to comparing random forest paths with self-determining trees. It is not easy to overfit and can accept noise and missing data. It also has the ability to manage files that are large in size and have many variables.

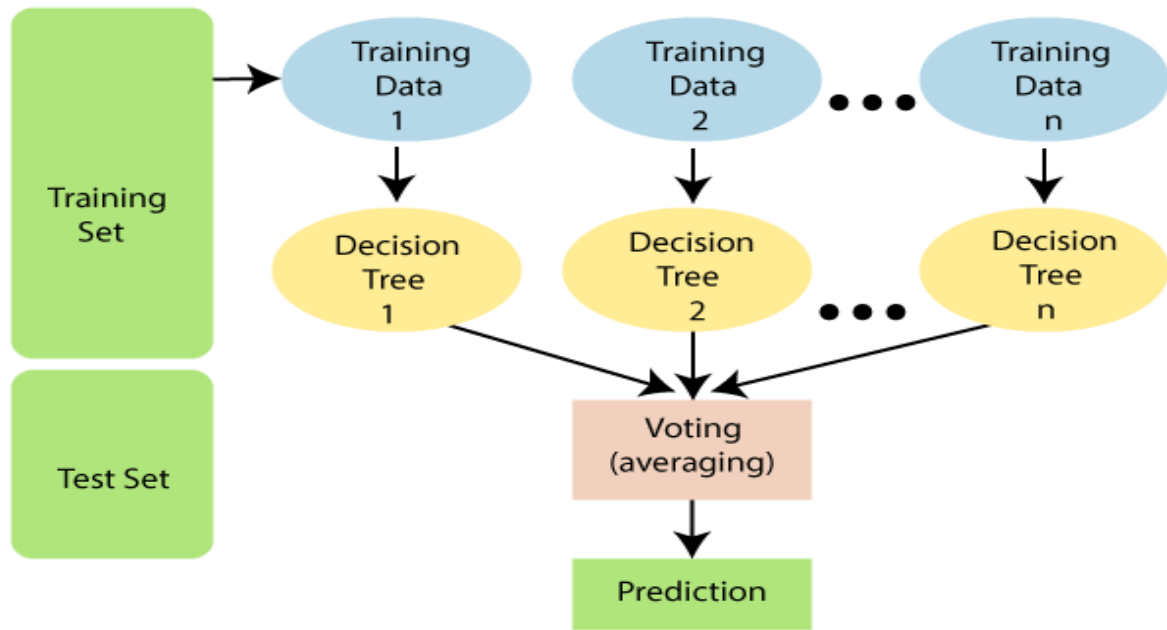


Figure 3.7: Random forest classification

CHAPTER 4

EXPERIMENTAL ANALYSIS AND DISCUSSION

CHAPTER 4

EXPERIMENTAL ANALYSIS AND DISCUSSION

4.1 Introduction

The classification of the data is determined by our model. We've selected four frequently used classification methods for this situation. We used the algorithms Simple Neural Networks, Long Short-Term Memory (LSTM), Convolutional Neural Networks, Random Forest. We tested each output strategy in our model training to see which was best. This segment, provide the overview of the research and exam result. what's more, the locating of the order document of applied calculations. complete 80% of the statistics is set for the training and for test information all out 20% records is utilized.

4.2 Experimental Analysis

4.2.1 Dataset Collection

	Review	Sentiment
0	Good and fast service And Recommend	Positive
1	Disgusting service, i Hat this courier company	Negative
2	Very very bad service	Negative
3	Excellent service	Positive
4	Their service is so good I am satisfied	Positive
5	Very bad service Specially for home delivery	Negative
6	worst one no customers services, rude behaviour	Negative
7	its make a important rule in our life	Positive
8	need to improve customer service	Positive
9	Worst service ever Their employees are illite...	Negative
10	Worst service ever Every single employee is a...	Negative
11	Good service All Country	Positive
12	Never send food items via them They eat the f...	Negative
13	The worst courier service of our country	Negative
14	good delivery surveys,24-48 ghonta delivery , ...	Positive
15	Worst service ever They even don't know how t...	Negative

Figure 4.1: Dataset

4.2.2 Import library

```
# Importing essential libraries and functions

import pandas as pd
import numpy as np
import nltk
nltk.download('stopwords')
from nltk.corpus import stopwords
from numpy import array
import seaborn as sns
import plotly.express as px
from sklearn.model_selection import train_test_split as tts
from keras.preprocessing.text import one_hot, Tokenizer
from tensorflow.keras.preprocessing.sequence import pad_sequences
from keras.models import Sequential
from keras.layers.core import Activation, Dropout, Dense
from keras.layers import Flatten, GlobalMaxPooling1D, Embedding, Conv1D, LSTM
from sklearn.model_selection import train_test_split
import matplotlib.pyplot as plt
import parse
import multiprocessing
from bs4 import BeautifulSoup
import re, string
from nltk.tokenize import word_tokenize
from nltk.stem.porter import PorterStemmer
from nltk.stem import WordNetLemmatizer
from sklearn.feature_extraction.text import TfidfVectorizer
from keras_preprocessing.sequence import pad_sequences
import unicodedata
import html
from sklearn.preprocessing import LabelEncoder
from keras.models import Model
from keras.layers import LSTM, Activation, Dense, Dropout, Input, Embedding
from keras.optimizers import RMSprop
from keras.preprocessing.text import Tokenizer
from keras.preprocessing import sequence
from tensorflow.keras.utils import to_categorical
from keras.callbacks import EarlyStopping

from sklearn.metrics import accuracy_score, confusion_matrix
from sklearn.naive_bayes import MultinomialNB
from sklearn.linear_model import LogisticRegression
from sklearn.svm import SVC

from sklearn import naive_bayes
from sklearn.metrics import roc_auc_score

import warnings
warnings.filterwarnings('ignore')
```

Figure 4.2: import library

Import several libraries to prepare for our dataset. The section contains a list of every tool we use. Python language is being used to recognize all of the used techniques. Sklearn is utilized for finished the version train. Portrayal settings in line with estimation were set observationally resulting to acting tries at numerous things with exclusive settings.

4.2.3 Data Preprocessing

This chapter analyzes our data using a wide range of techniques. We looked at there have any Hashtag, emoji, html tag, any of colon, comma, full-stop, and change uppercase letter to lower case letter.

```
max(tokenizer.word_index)
```

'🤔🤔'

Figure 4.3: Remove emoji

In our dataset, there are some emoji 4.3. There are so many rows of comma, semicolon, full-stop data in the data. For preprocessing, it was extremely huge

	Review	Sentiment	liked
0	good fast service recommend	Positive	1
1	disgusting service hat courier company	Negative	0
2	bad service	Negative	0
3	excellent service	Positive	1
4	service good satisfied	Positive	1

Figure 4.4: Final data after preprocessing

After removing this value, we get 4.4 figures chart. Here we can see there are extra values in our dataset.

4.2.4 Embedding Layer

I am using Text data and I want to analysis the sentiment of the given data. So, I need to embedding the text data. There I am using tokenizer, corpus for embedding that text.

```
# Create Embedding Matrix having 100 columns
# Containing 100-dimensional GloVe word embeddings for all words in our corpus.

embedding_matrix = zeros((vocab_length, 100))
for word, index in word_tokenizer.word_index.items():
    embedding_vector = embeddings_dictionary.get(word)
    if embedding_vector is not None:
        embedding_matrix[index] = embedding_vector
```

```
embedding_matrix.shape
```

```
(12766, 100)
```

Figure 4.5 Embedding data

4.2.5 Data Visualization

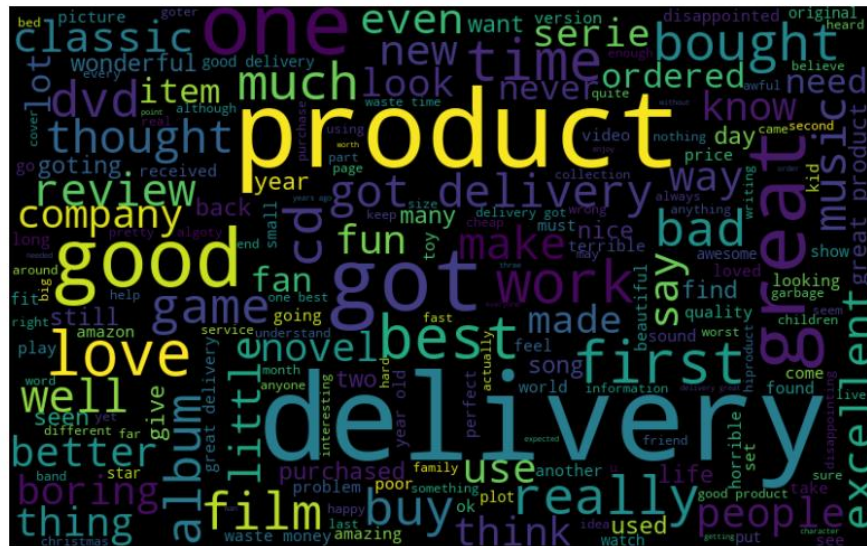


Figure 4.6: Word cloud

There are some groups by text in this dataset. It provides some text ranges.

```
In [ ]: tokenizer.word_counts
```

```
Out[75]: OrderedDict([('good', 1051),  
                        ('and', 4310),  
                        ('fast', 84),  
                        ('service', 141),  
                        ('recommend', 76),  
                        ('disgusting', 6),  
                        ('', 5937),  
                        ('i', 6068),  
                        ('hat', 2),  
                        ('this', 5393),  
                        ('courier', 17),  
                        ('company', 144),  
                        ('very', 933),  
                        ('bad', 348),  
                        ('excellent', 261),  
                        ('their', 140),  
                        ('is', 3528),  
                        ('so', 671),  
                        ('am', 263),  
                        ('best', 611),  
                        ('...', 10000)])
```

Figure 4.7: word count

To take a gander on the plan ampleness and precision measures were used. Request behind schedule results of every dataset have proven up in desk, where the dataset name is shown up

beneath each plan consequences collecting, at the same time as precision and effects are proven up in modified figures, restricted for each thing warranty level. Inside the discern, all pre-owned calculations are given with their presentation for opinion exam. Sentiments are calculated through the delivery overview using the customer score. The score factors are divided into five stars. Each start consists of the food quality of the patron review.

```
In [ ]: text = logistics_reviews["Review"][0]
        print(text)
        print("<=====>")
        print(word_tokenize(text))

Good and fast service And Recommend
<=====>
['Good', 'and', 'fast', 'service', 'And', 'Recommend']
```

Figure 4.8: separate all word

4.2.6 Data preprocessing

Permit's import a few important facts technology tools like pandas and NumPy before transferring on to label encoding in Python. Pandas will subsequent observe the Excel-formatted records record to determine whether or not the import was a hit. A state datatype is absolutely an object. Label encoding is the subsequent situation.

4.2.7 Label Encoding

```
from sklearn.preprocessing import LabelEncoder
import re
```

Figure 4.9: Label Encoding

Specific variables are transformed into numerical labels the usage of the label encoding technique in system mastering and records analysis. It's important to take into account that label encoding should most effective be used sparingly, specifically while running with variables whose numerical names indicate a herbal hierarchy or order. In those situations, label encoding ought to supply a misleading feeling of ordinality and lead the version of beam. it's miles often most well-known to symbolize categorical variables in a device studying version the use of one-warm encoding or different appropriate approaches if the specific variable lacks any discernible order.

4.2.8 Train Test Split

```
] : # Embedding layer expects the words to be in numeric form
# Using Tokenizer function from keras.preprocessing.text library
# Method fit_on_text trains the tokenizer
# Method texts_to_sequences converts sentences to their numeric form
|
word_tokenizer = Tokenizer()
word_tokenizer.fit_on_texts(X_train)

X_train = word_tokenizer.texts_to_sequences(X_train)
X_test = word_tokenizer.texts_to_sequences(X_test)
```

Figure 4.10: Features columns 1

```
# Padding all reviews to fixed length 100

maxlen = 100

X_train = pad_sequences(X_train, padding='post', maxlen=maxlen)
X_test = pad_sequences(X_test, padding='post', maxlen=maxlen)
```

Figure 4.11: Features columns 2

The dataset is split usually to evaluate how nicely the trained version works on sparkling records. by way of testing the model on a exclusive dataset, you could advantage a greater correct evaluation of its overall performance and find out how well it generalizes to novel, untested situations. keep in mind that further to the education and testing units, a separate validation set must be applied to great-track the model and optimize parameters all through the education phase.

4.3 Experimental Results

```
# Model compiling
snn_model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['acc'])
print(snn_model.summary())
```

Model: "sequential"

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 100, 100)	1276600
flatten (Flatten)	(None, 10000)	0
dense (Dense)	(None, 1)	10001

=====
Total params: 1,286,601
Trainable params: 10,001
Non-trainable params: 1,276,600
=====
None

Figure 4.12: Model Summery

```
: # Model Performance  
  
print("Test Score:", score[0])  
print("Test Accuracy:", score[1])  
  
Test Score: 0.03782855346798897  
Test Accuracy: 0.9995076060295105
```

Figure 4.13: SNN model accuracy

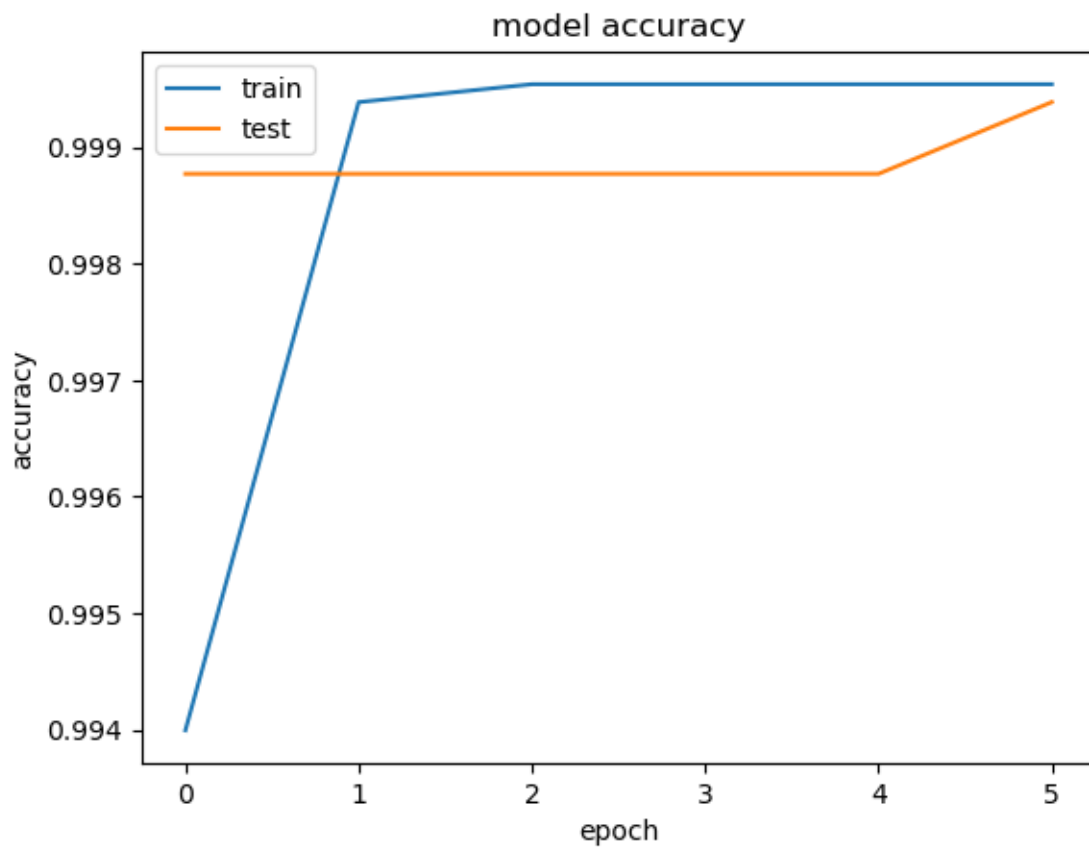


Figure 4.14: SNN model accuracy Graph

```
: # Model Performance  
  
print("Test Score:", score[0])  
print("Test Accuracy:", score[1])
```

Test Score: 0.3356787860393524
Test Accuracy: 0.8579000234603882

Figure 4.15: CNN model accuracy

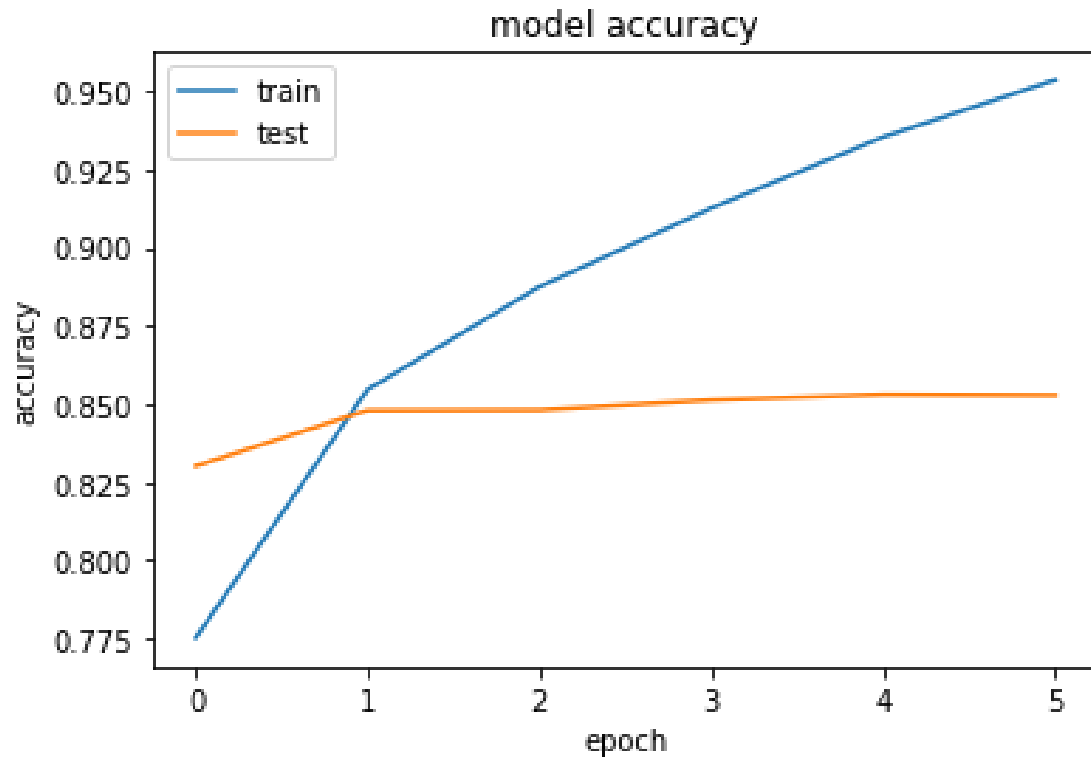


Figure 4.16: CNN model accuracy Graph


```
# Model Performance  
  
print("Test Score:", score[0])  
print("Test Accuracy:", score[1])
```

Test Score: 0.31936636567115784
Test Accuracy: 0.864300012588501

Figure 4.17: LSTM model accuracy

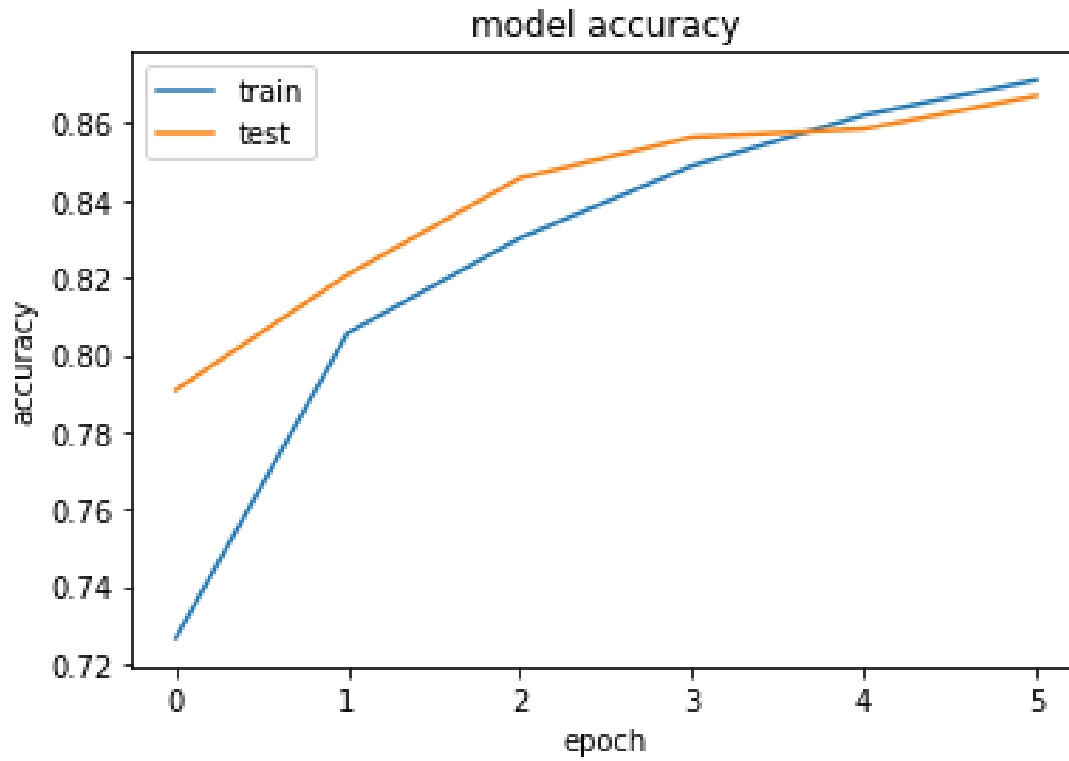


Figure 4.18: LSTM model accuracy Graph

```
] : ypred=model.predict(xtest)

# model score
accuracy_score(ypred,ytest)

]: 0.7778653012997243
```

Figure 4.19: Logistics Regression model accuracy

```
# f1 score
recall=A[0][0]/(A[0][0]+A[1][0])
precision=A[0][0]/(A[0][0]+A[0][1])
F1=2*recall*precision/(recall+precision)
print(F1)

0.7879699248120301
```

Figure 4.20: Logistics Regression F1 Score

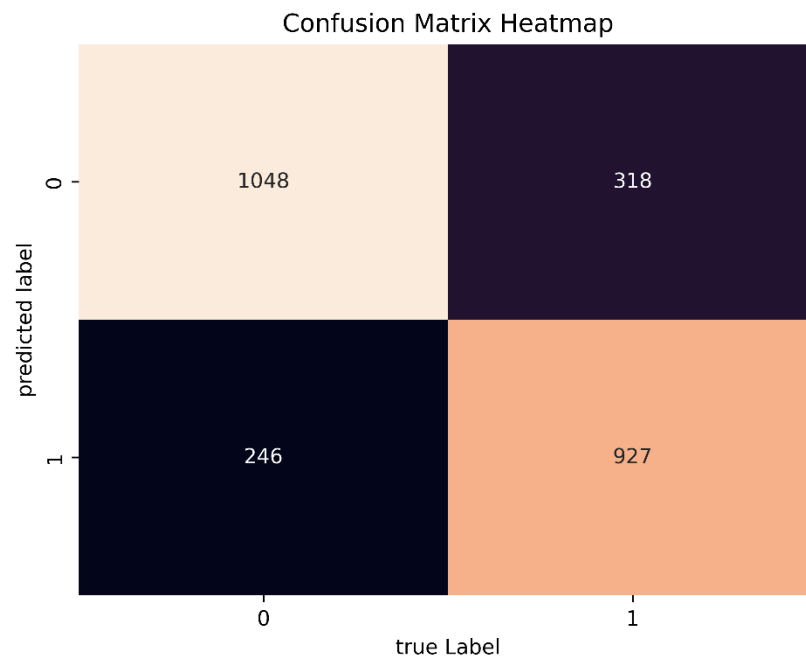


Figure 4.21: Logistics Regression Heatmap

Algorithm	Test Score	F1-Score	Accuracy
Simple Neural Network	0.37	0.98	0.99
CNN	0.33	0.84	0.85
LSTM	0.31	0.65	0.86
Logistic Regression	0.74	0.79	0.77

Table 4.1: Different Algorithm results

In our projects, we utilized each of these four separate algorithms. The accuracy of each method is displayed in the table 4.1 that was previously discussed. The algorithm known as the decision tree produces the most fruitful results. Its accuracy is higher than that of other algorithms, which range from 77% to 100%.

4.4 Summary

there may be a ton of expansions to go after this version which fuses point primarily based portrayal and emoticon, retweet records, and many others. to be considered as functions to make a continuously summarized gaining knowledge of pointer depend.

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

5.1 Impact on Society

Online existence can be relatively persuasive on society in each superb and poor manners. It gives individuals a technique to handling to preserve in contact with folks that live some distance away. It allows individuals to proportion a laugh, spellbinding, and informative substance. It offers affiliations an approach to dealing with drawing in with clients. one of the troubles, in any case, is that everybody can proportion whatever, along with material that won't be precise. every now and then, genuine mischievousness is done whilst individuals spread scorching, unconfirmed, or all-around fake statistics. this may hurt personal human beings, as when somebody is tormented at the internet. it may in like manner damagingly effect society in light of the entirety. on-line lengthy-reach relational correspondence has made it easy to unfold records unexpectedly. seeing that Facebook and website plans move so swiftly, watchers don't periodically check what they have got visible. plenty of substance is moreover spread thru images a lot, which might be set up on full-size records. unique snap shots are made to be intriguing, beguiling, or over the top. Others, regardless, are wished to persuade our reasoning. really, even relationships with affirmed articles may be misdirecting. it is made sure to talk that a huge quantity of folks that see a thing and association by no means examine the entire article. unique areas pass on phony news usually as misleading materials. They need people to tap on lovely functions to get clicks so that they can promote all of the additional progress. There are further targets with a stable philosophical inclination. these is probably insane preservationist, left-wing, or planned areas. A primary quantity of these targets isn't seeing about dispersing "news" stories that lower back their angle. digital memoirs, irrespective of whether proper or no longer, reliably come to be a web sensation. The more ridiculous and newsworthy some thing is, the essentially certain that specific people will share it. all through an altogether hypnotized political choice, people are moved to percentage something strengthens their competitor or, essentially actually, ambushes the renouncing competitor. this will accomplish faux tales commonly streaming. irrespective of whether counterfeit testimonies are pulverized and retracted, the naughtiness is beginning at now executed. proper when a withdrawal is dispersed, limitless individuals may also have as of past due regarded the tale.

5.2 Ethical Aspects

Digital lifestyles are possibly the most ideal approach to control find out and speak with new individuals who've doubtful pursuits from you, as on-line frameworks company awards you to watch out social affairs which are secured towards your high-quality conditions and aspect pastimes. this may be unfathomable for meeting new amigos but additionally for companionship interests and internet relationship which are becoming greater brilliant than a widespread inverse social affair, through the distinction of online life and any closeness to Tinder. digital existence is an outstanding process to get out the phrase hastily over the globe, with "breaking information" tweets getting incalculable retweets in minutes. this could be altogether large while restoring people on primary records, as an instance, climate empowerment and missing youngsters. As mentioned above net coordinating has while the entirety is said in completed modified society in extraordinary effective manners anyhow for glaringly no additional cost as all simple online lifestyle's degrees are unfastened. try and bear in mind something else or association that has ever changed yourself as tons as online life and from that point don't forget its cost.

5.3 Impact on Environment

it is no matter what you seem like at it sharing phony tales that reason hurt via electronic frameworks company media. in the event that you're a character or commercial enterprise that offers thousands of substances, perhaps with the manual of online existence programming, you must be more cautious. It just takes a quick comfort or to test something you spot via techniques for electronic frameworks company media. think about the wellspring of the story. at the occasion which you've in no way considered the large picture, Google it and check whether it's first-rate. in case you don't have that lots time, it is best to overlook it, especially if it takes subsequent to something that could be a parody, beguiling substance, or conscious openness. by way of now not sharing blemished fabric, you can help with taking out the growth of guile and phony information. As cited multiple digital life's basic perfect situations is the way that facts may be unfold to endless individuals in an unimaginably quick quantity of time. regardless of the way that this may be regarded as a tremendous high-quality in an occasion in a disaster, it may comparatively be a monster impediment as information with no credibility can take part in a quick second. this may motive certified trickery and out-of-manipulate situation. An event of this will be seen when speak unfold that the Queen has kicked the basin, down to the Queen missing a Christmas affiliation right down to a everyday contamination. This in spite of distinctive hoaxers making counterfeit reviews precipitated extraordinary to recognize that the Queen has kicked the box.

5.4 Sustainability

- There are typically 2.3 Billion generally speaking unique web-based life customers.
- 95% of tremendous corporate brands have in any event two web-based life stages.
- 75% surprisingly feel abnormal and awkward when they can't get to their online life profiles.

CHAPTER 6

CONCLUSION & FUTURE WORK

CHAPTER 6

CONCLUSION & FUTURE WORK

6.1 Conclusion

Test mining was developed by limiting information to non-trivial guesses to understand subjects' normal thoughts about these thoughts. The method I propose captures both the user's conclusions and the original content for these metrics. Therefore, this strategy removes all guesswork by making them a priority and providing clear insights into who will buy how much. The purpose of the plan is to answer questions about the drivers of each strategy registered in the database and to examine the generality of that strategy. Here I used the ML group technique to find a hypothetical research on data from customer reviews. Logistic regression provides high precision to predict controls I tried my idea with a set of controls and it worked. I write regular reviews for 5-star carriers and will test our strategy with more brands in the future. I always use my approach to check various fields to check the suggested method. It should be noted that while the instruction-based techniques used in hypothetical events can identify underlying assumptions, the interpretation can sometimes be ambiguous, such as inconsistency or generalization in understanding. Future testing should try to answer these problems.

6.2 Recommendation

The trial evaluation is the modern-day example to realize the prerequisites of the mass open; it's a much less complicated and sharp way to deal with perceiving how the humans are feeling approximately a particular the issue of difficulty and the logo effect of extra constrained length contributing to a weblog.

- Remove predisposition from dataset.
- Need the proportion of information classes is similarly appropriated.
- Use more AI characterization calculations.
- Create neural organization for better outcome.
- Need boundary tuning for characterization.

6.3 Future Work

The sensitive improvement of facts on the web and web-based totally existence districts license associations to use quit assessment to get a knowledge of customers' estimations approximately their things or companies. evaluation examinations in present composing are frequently accomplished reliant at the constrained Twitter information the use of or 3 days' really worth of information. Twitter data is regularly open simplest for as much as every week and with a restriction of round 1500 Tweets specially. however, within the occasion that Twitter data is when doubtful usually downloaded this obstruction thwarts verifiably significant and extensive consequences being procured. A vast examination of tweets to decide honestly huge customer appraisal needs to deal with multiple measures. these consolidate, at the base, 1) a satisfactorily large time c program language period over which tweets are collected to make sure representativeness as opposed to human being's quick response following a touch of statistics about the occasion, 2) a quality quantity of tweets that satisfactory cope with each geographic region, three) a take a look at of viable tendency if the tweets start from a given geographic area and four) if the closures drawn fit the unmistakable disposition got from different marketplace resources.

Reference

- [1] Nguyen, Bang, Van-Ho Nguyen, and Thanh Ho. "Sentiment analysis of customer feedbacks in online food ordering services." *Business Systems Research: International journal of the Society for Advancing Innovation and Research in Economy* 12.2 (2021): 46-59.
- [2] Singla, Zeenia, Sukhchandan Randhawa, and Sushma Jain. "Statistical and sentiment analysis of consumer product reviews." 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE, 2017.
- [3] Tang, Yachun. "Research on Word Vector Training Method Based on Improved Skip-Gram Algorithm." *Advances in Multimedia* 2022 (2022).
- [4] Asgari-Chenaghlu, Meysam. *Word Vector Representation, Word2vec, Glove, and Many More Explained*. Diss. University of Tabriz, 2017.
- [5] Nguyen, Bang, Van-Ho Nguyen, and Thanh Ho. "Sentiment analysis of customer feedbacks in online food ordering services." *Business Systems Research: International journal of the Society for Advancing Innovation and Research in Economy* 12.2 (2021): 46-59.
- [6] Zhang, Shaozhong, et al. "Research on Logistics Service Assessment for Smart City: A Users' Review Sentiment Analysis Approach." *Electronics* 11.23 (2022): 4018.
- [7] Bhatt, Aashutosh, et al. "Amazon review classification and sentiment analysis." *International Journal of Computer Science and Information Technologies* 6.6 (2015): 5107-5110..
- [8] Laksono, Rachmawan Adi, et al. "Sentiment analysis of restaurant customer reviews on tripadvisor using naïve bayes." 2019 12th international conference on information & communication technology and system (ICTS). IEEE, 2019..
- [9] Sharma, Tamanna, Anu Bajaj, and Om Prakash Sangwan. "Deep learning approaches for textual sentiment analysis." *Research Anthology on Implementing Sentiment Analysis Across Multiple Disciplines*. IGI Global, 2022. 256-267.
- [10] BİLEN, Burhan, and Fahrettin HORASAN. "LSTM network based sentiment analysis for customer reviews." *Politeknik Dergisi* 25.3 (2021): 959-966.
- [11] Nandwani, Pansy, and Rupali Verma. "A review on sentiment analysis and emotion detection from text." *Social Network Analysis and Mining* 11.1 (2021): 81.
- [12] Singla, Zeenia, Sukhchandan Randhawa, and Sushma Jain. "Statistical and sentiment analysis of consumer product reviews." 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE, 2017.
- [13] Alzahrani, Mohammad Eid, et al. "Developing an intelligent system with deep learning algorithms for sentiment analysis of e-commerce product reviews." *Computational Intelligence and Neuroscience* 2022 (2022).

- [14] Alsubari, Saleh Nagi, et al. "Development of integrated neural network model for identification of fake reviews in E-commerce using multidomain datasets." *Applied Bionics and Biomechanics* 2021 (2021).
- [15] Elmurngi, E. I., & Gherbi, A. (2018). Unfair reviews detection on amazon reviews using sentiment analysis with supervised learning techniques. *J. Comput. Sci.*, 14(5), 714-726.
- [16] Dharma, Arie Satia, and Yosua Giat Raja Saragih. "Comparison of Feature Extraction Methods on Sentiment Analysis in Hotel Reviews." *Sinkron: jurnal dan penelitian teknik informatika* 7.4 (2022): 2349-2354.
- [17] Vaishanvi, S., et al. "Product Recommendation Using Sentiment Analysis." 2022 *International Conference on Engineering and Emerging Technologies (ICEET)*. IEEE, 2022.

Turnitin Originality Report

Processed on: 27-Jul-2023 15:18 +06

ID: 2137503933

Word Count: 8658

Submitted: 1

193-35-487 By Md Rabiul Awal Shuvo

Similarity Index	Similarity by Source
25%	Internet Sources: 18% Publications: 14% Student Papers: 12%

4% match (Nourin Islam, Nasrin Akter, Abdus Sattar. "Sentiment Analysis on Food Review using Machine Learning Approach", 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 2021)

[Nourin Islam, Nasrin Akter, Abdus Sattar. "Sentiment Analysis on Food Review using Machine Learning Approach", 2021 International Conference on Artificial Intelligence and Smart Systems \(ICAIS\), 2021](#)

2% match (Internet from 18-Jun-2022)

<https://www.coursehero.com/file/148762637/Sentiment-Analysis-of-Customer-Feedback-in-Online-pdf/>

2% match (student papers from 02-Jan-2023)

[Submitted to Daffodil International University on 2023-01-02](#)

2% match (Internet from 06-Dec-2022)

https://mdpi-res.com/d_attachment/electronics/electronics-11-04018/article_deploy/electronics-11-04018.pdf?version=1670145386

2% match (Internet from 03-Jun-2023)

<https://www.jurnal.iaii.or.id/index.php/RESTI/article/download/4751/767/>

1% match (Internet from 21-Jul-2023)

<http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/10005/22638.pdf?isAllowed=y&sequence=1>

1% match (Internet from 29-Jun-2023)

<http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/9624/22177.pdf?isAllowed=y&sequence=1>

1% match (Internet from 24-Dec-2022)

<http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/3488/P13592%20%2826%25%29.pdf?isAllowed=y&sequence=1>

1% match (Internet from 15-Jan-2023)

http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/7365/172-35-2120%20%2814_%29%20clearance.pdf?isAllowed=y&sequence=1

1% match (Shaozhong Zhang, Haidong Zhong, Chao Wei, Ding kai Zhang. "Research on Logistics Service Assessment for Smart City: A Users' Review Sentiment Analysis Approach", Electronics, 2022)

[Shaozhong Zhang, Haidong Zhong, Chao Wei, Ding kai Zhang. "Research on Logistics Service Assessment for Smart City: A Users' Review Sentiment Analysis Approach", Electronics, 2022](#)

1% match (student papers from 17-Jan-2023)

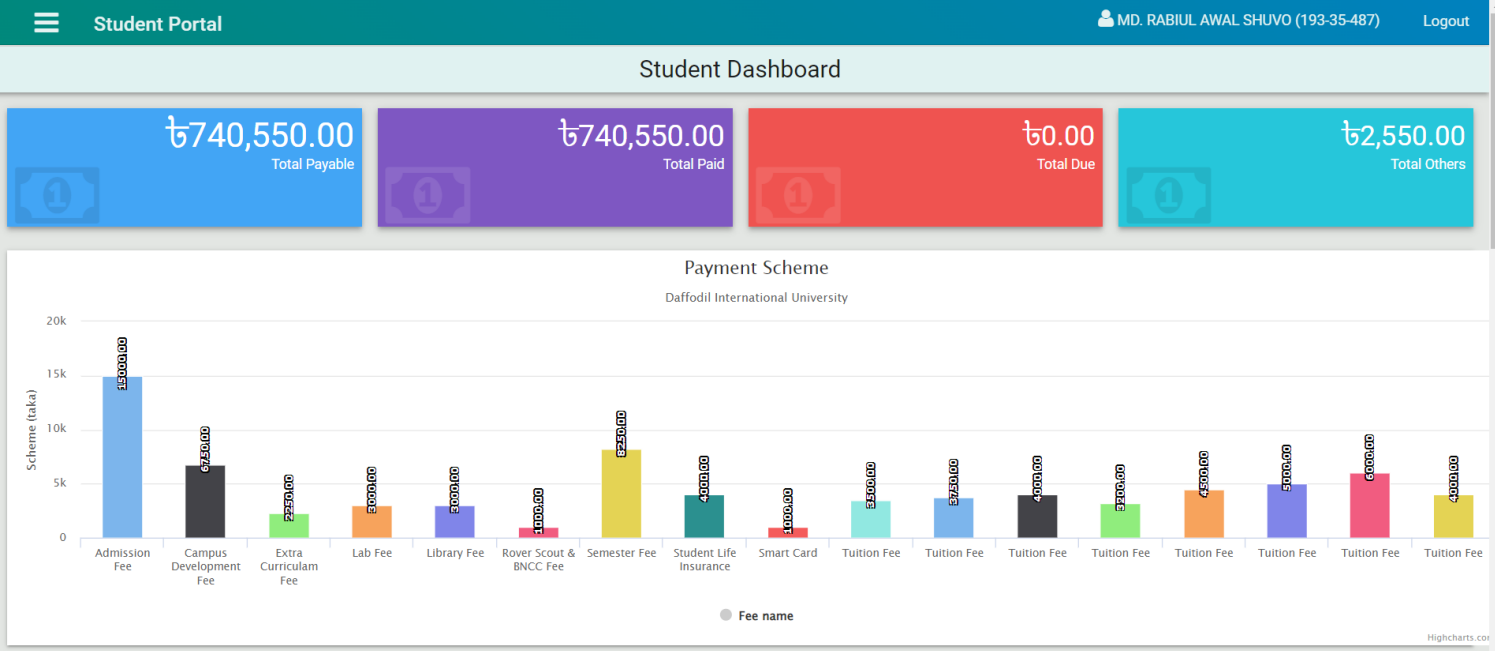
[Submitted to Univerza v Ljubljani on 2023-01-17](#)

1% match (student papers from 08-Nov-2022)

[Submitted to ESCP-EAP on 2022-11-08](#)

< 1% match (Internet from 25-Oct-2022)

Accounts Clearance



Student Portal Clearance

Student Portal


MD. RABIUL AWAL SHUVO (193-35-487)

Logout

Semester	Registration	Mid Term Exam	Final Exam/Assessment
Fall, 2019	✓	✓	✓ (Final Exam)
Spring, 2020	✓	✓	✗ (Final Exam)
Summer, 2020	✓	✗	✗ (Final Exam)
Fall, 2020	✗	✗	✗ (Final Exam)
Spring, 2021	✓	✗	✗ (Assessment)
Summer, 2021	✗	✗	✗ (Final Exam)
Fall, 2021	✓	✗	✓ (Final Exam)
Spring, 2022	✓	✗	✓ (Final Exam)
Summer, 2022	✓	✗	✓ (Final Exam)
Fall, 2022	✓	✗	✓ (Final Exam)
Spring, 2023	✓	✓	✓ (Final Exam)


All Rights Reserved @ Daffodil International University

Library Clearance

 Gmail


Search in mail


Active





8 of 3,015


Final Thesis Report External Inbox x

**MD. RABIUL AWAL SHUVO** Thu, Jul 27, 12:46 PM (7 days ago) ☆
Please check the plagiarism of this thesis paper Name: MD RABIUL AWAL SHUVO ID: 193-35-487 B.Sc in Software Engineering

**Md. Rittique Alam** Thu, Jul 27, 1:56 PM (7 days ago) ☆
Please proceed.

**Project Report Library** Thu, Jul 27, 3:20 PM (7 days ago) ☆
Dear Student, Your Plagiarism Result is 25% for details Please see the attachment file. Please read the instruction: - Report to be arranged according to the

**MD. RABIUL AWAL SHUVO** Wed, Aug 2, 4:02 PM (1 day ago) ☆
I have Completed all the instructions, Attache the supervisor & my signature on the approval and declaration page, And I have filled up all my information in In

**Project Report Library** 10:01AM (9 hours ago) ☆
to Software, me, Rittique ▾

Dear Student,
Your Defence clearance is done.

Thanks a lot. Thank you so much for the great news! Thank you for informing me.

Reply Reply all Forward