Title of the Thesis

“Classification of product review sentiment by NLP and hybrid ML algorithms”

Submitted By:

Shamsunnahar Shanta

ID:202-16-542

Supervised By:
Md. Mehedi Hassan
Lecturer of Department CIS

DEPARTMENT OF COMPUTING AND INFORMATION SYSTEM

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

Submission date:29 September 2024

APPROVAL

This Thesis titled **"Classification of product review sentiment by NLP and hybrid ML algorithms"**, Submitted by **Shamsunnahar Shanta**, ID No:**202-16-542** to the Department of Computing & Information Systems, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computing & Information Systems and approved as to its style and contents. The presentation has been held on 29-09-2024.

BOARD OF EXAMINERS

29/09/24

Md Sarwar Hossain Mollah
Associate Professor and Head
Department of Computing & Information Systems
Faculty of Science & Information Technology
Daffodil International University

Chairman

29.09.24

Md. Nasimul Kader
Assistant Professor
Department of Computing & Information Systems
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

29.09.2024

Md. Mehedi Hassan
Lecturer
Department of Computing & Information Systems
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Md. Miftah Uddin
Head of SBU, Brain Station 23

External Examiner

©Daffodil International University, Bangladesh i

Declaration

I hereby declare that; this project has been done by me under supervision of **Md. Mehedi Hassan, Lecturer**, department of Computing and Information System (CIS) of Daffodil International University. I am also declaring that this project or any part of there has never been submitted anywhere else for the award of any educational degree like, B.Sc., M.Sc., Diploma or other qualifications.

Supervised By

29.09.2024

Md.Mehedi Hassan
Lecturer
Department of CIS
Daffodil International University

Submitted By

Shanta

Shamsunnahar Shanta
ID:202-16-542
Department of CIS
Daffodil International University

ACKNOWLEDGEMENT

First, we express our heartiest thanks and gratefulness to almighty for His divine blessing making it possible for us to complete the final year project successfully.

I am grateful and wish our profound indebtedness to **Md. Mehedi Hassan**, Department of Computing and Information System (CIS) Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of *Artificial Intelligence* to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

I would like to express our heartiest gratitude to the **Mr. Sarwar Hossain Mollah, Head of the Department of Department of Computing and Information System (CIS),** for his kind help in finishing our project and also to other faculty members and the staff of the Department of Computing and Information System (CIS), 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 patience of our parents.

ABSTRACT

With the advent of e-commerce in Bangladesh, consumer sentiment can be better gauged through an uptick in reviews. Analyzing these evaluations is a challenge due to their disorganized nature. This study aims to improve sentiment classification in user-generated product evaluations by creating a hybrid machine-learning approach. The database has 5,000 reviews from various online sources, including Facebook groups and marketplaces like Daraz and Rokomari. Tokenization and other text preprocessing methods were employed. We established and evaluated a DT RF and SVM, a DT hybrid with NN and SVM, and an RF hybrid with SVM. I assessed these models using recall, accuracy, precision, and F1-score. With an F1 score of 0.95 and an accuracy of 94%, the hybrid DT, NN, and SVM models outperformed the other models. The DT, RF, and SVM models each received a score of 91% accuracy, while the RF and SVM models received a score of 97%. Based on the findings of this study, RF and SVM has the potential to enhance sentiment categorization for online purchases significantly. This, in turn, aids businesses in better understanding customer feedback and making informed decisions.

TABLE OF CONTENTS

LIST OF ACRONYMS

SVM = Support Vector Machine

RF = Random Forest

NN = Neural Networks

DT = Decision Trees

NLP = Utilizing Natural Language Processing

TF-IDF = Term Frequency-Inverse Document Frequency

DNN = Deep Neural Networks

RNN = Recursive Neural Network

CNN = Convolutional Neural Networks.

CHAPTER 1
INTRODUCTION

1.1 Overview

A scrutiny of the product In order to enhance comprehension of client sentiments and optimize e-commerce tactics, doing sentiment analysis is vital. Product reviews have a significant influence on consumer behavior as they can affect the purchasing decisions customers make. Furthermore, they provide organizations with crucial insights into client happiness and product efficacy. Given the rapid expansion of the e-commerce sector in Bangladesh, it is important to possess a comprehensive comprehension of customer perspectives. Platforms like Daraz, Birkory.com, Rokomari, Food Panda, and Facebook have become significant participants in the internet retail business in Bangladesh, which has experienced substantial growth in recent years [1]. The increase can be ascribed to several factors, such as the widespread use of mobile devices, the increased accessibility to the internet, and the expansion of the middle class, which currently possesses a larger amount of expendable income. In order to ensure customer satisfaction and maintain a high level of service quality as the company expands, it is imperative for the organization to effectively handle and evaluate substantial amounts of client feedback. Natural Language Processing (NLP) employs several methodologies to examine textual data and extract valuable insights from huge volumes of unorganized text. To effectively utilize this approach, tokenization, which involves the process of dividing the text into discrete words or tokens, is important. Utilizing natural language processing (NLP) to examine the framework of a review enables machine learning algorithms to classify the transmitted emotions [2]. Decision trees (DT), random forests (RF), and support vector machines (SVM) often employ data analysis techniques for sentiment analysis. The advantages and cons of these algorithms vary significantly from each other. In addition, decision trees tend to overfit, even though they are typically easy to comprehend. To address the problem of overfitting, one can employ ensemble approaches like Random Forests. Implementing Support Vector Machines (SVMs) necessitates the utilization of substantial computational resources. SVMs are renowned for their outstanding performance in domains with a high number of dimensions. A recent study suggests that employing hybrid approaches that integrate machine learning techniques has the potential to enhance sentiment analysis models. Hybrid models offer improved precision and strong classification capabilities by combining multiple techniques into a single framework. Support Vector Machines (SVM), Decision Trees (DT), and Random Forests (RF) are three machine learning methods that enhance a model's ability to generalize from training data while maintaining its accuracy. In this experiment, a mixed machine-learning technique is used to classify the emotions conveyed in product reviews. In order to do natural language processing preprocessing, we tokenized a total of 5000 reviews collected from several e-commerce platforms. Currently, we are analyzing the similarities and distinctions among DT, RF, and SVM hybrid models. Our team will conduct individual and combined studies on the RF, SVM, DT, and NN models. The hybrid model, consisting of DT, NN, and SVM, demonstrated the highest level of accuracy. The evidence provided demonstrates that employing hybrid methodologies has the capacity to enhance the accuracy and effectiveness of sentiment analysis. The project has had an impact on the e-commerce industry in Bangladesh. Implementing advanced sentiment analysis techniques can help Bangladeshi e-commerce enterprises enhance the quality of their products and customer service by evaluating client sentiments. Enhancing customer satisfaction and loyalty has the capacity to significantly accelerate the growth of Bangladesh's e-commerce industry.

1.2 Problem Statement

The e-commerce industry in Bangladesh has experienced significant growth, leading to an increase in online product reviews that provide insights into client satisfaction and preferences. Because of their large size, corporations face difficulties in manually assessing and deriving conclusions from these evaluations. Conventional sentiment analysis algorithms may need help to analyze diverse and disorganized textual content from several sources. In the competitive e-commerce industry, there is a need for more sophisticated techniques to categorize the sentiment of product reviews. The current DT, RF and SVM approaches have limitations that diminish their effectiveness. SVMs might encounter difficulties when dealing with large datasets, DT tend to overfit when confronted with noisy data, and RF require significant processing resources. Hybrid machine learning models can employ many ways to enhance sentiment analysis. The potential of hybrid approaches for sentiment analysis in Bangladesh's e-commerce sector is excellent, although they have been overlooked in academic studies. This work constructs and evaluates hybrid machine learning models to categorize the sentiment of product reviews. DT, NN, and SVN Machines enhance sentiment categorization in the research. Bangladeshi e-commerce firms would be provided with advanced technology for automated analysis of user input, enabling them to better cater to consumer demands, enhance product offerings, and enhance customer happiness and loyalty.

1.3 Motivation and Objectives

The motivation for this project is

- The rapid growth of e-commerce in Bangladesh has led to increased online product reviews.
- Reviews provide insights into client preferences and satisfaction. The reviews are lengthy and lack organization, making them unsuitable for manual analysis.
- Traditional sentiment analysis techniques are typically imprecise and vulnerable.
- Hybrid machine learning models integrate the advantageous aspects of multiple algorithms.
- Enhanced sentiment analysis enables companies to make decisions based on data.
- Businesses enhance customer satisfaction and competitiveness by implementing more advanced analysis techniques.
- Collect 5000 evaluations of Bangladeshi e-commerce platforms.
- Perform tokenization and natural language processing on reviews.
- Evaluate the accuracy, precision, recall, and F1-score of these models. Discover and assess the optimal model.
- Demonstrate the utilization of sophisticated sentiment analysis.
- Contribute to sentiment analysis research by providing findings and suggestions.

1.4 Project Scope

This research develops and tests hybrid machine learning models for sentiment classification of product reviews from Bangladeshi e-commerce sites. The project will preprocess 5000 reviews from Daraz, Birkory.com, Rokomari, Food Panda, and various

Facebook pages and groups using tokenization and other NLP techniques to ensure data cleanliness and structure. One hybrid model will combine DT), RF, and SVM; another will couple RF with SVM; and a third will combine DT, NN, and SVM. After extensive training on the preprocessed dataset, each model will be evaluated on accuracy, precision, recall, and F1-score. The project will carefully assess these models' strengths and weaknesses to determine the best sentiment analysis method for this context. Practical and actionable recommendations will be developed from this investigation to help Bangladeshi e-commerce enterprises adopt and optimize sentiment analysis techniques. This research also hopes to advance sentiment analysis and hybrid machine learning applications by providing new insights and applications.

1.5 Project Outcome

- The project develops and improves hybrid machine learning models like DT, RF, SVM, and NN to accurately classify product review sentiment.
- To verify hybrid techniques and improve accuracy, precision, recall, and F1-score.
- This analysis compares different models to assess their pros and cons for e-commerce applications.
- The investigation will offer practical advice to help Bangladeshi e-commerce enterprises use sentiment analysis effectively.
- The study advances sentiment analysis and hybrid machine learning in real contexts using unique methods and notable results.
- Improved consumer feedback systems will help enterprises use data for product development and customer service.
- The project adapts to different dataset sizes and business needs, ensuring flexibility and relevance across e-commerce platforms.
- This allows sentiment analysis research and machine learning applications in Bangladeshi e-commerce.

1.7 Summary

The project's introduction focuses on sentiment analysis of product reviews in Bangladesh's rapidly expanding e-commerce industry. It highlights the significance of consumer input in both business strategy and overall satisfaction. The project aims to address the challenges posed by the large number of online reviews, which can be overwhelming to analyze manually, and the lack of structure, which makes it difficult to extract meaningful insights by employing robust hybrid machine-learning models. This model employs DT, NN, RF, SVM techniques to enhance the accuracy and resilience of sentiment classification. The project goals of this chapter include data collection, model construction, performance evaluation, and practical application insights. This introduction sets the stage for innovative techniques in sentiment analysis that can assist e-commerce businesses in Bangladesh.

CHAPTER 2

BACKGROUND STUDY

2.1 Overview

Chapter 2 provides an overview of the project's background and importance. The text introduces sentiment analysis, and its role in helping e-commerce enterprises comprehend customer feedback. The chapter explores various NLP techniques, such as tokenization, stemming, and lemmatization, for the purpose of text preprocessing. The chapter subsequently explores various machine-learning approaches for sentiment analysis, including DT, RF, SVM, and NN. Examining the advantages and disadvantages of each algorithm aids in rationalizing the selection of a hybrid model. The chapter also discusses the study of hybrid machine learning, emphasizing its classification accuracy and robustness benefits. The chapter also examines the e-commerce industry in Bangladesh, highlighting its rapid growth and increasing prevalence of online product reviews. It emphasizes the necessity of utilizing automated sentiment analysis techniques to assist companies in manually analyzing this vast quantity of unorganized data.

2.2 Related Works

Bhalerao el. At. [3] produced us that the application of classification was important in producing remarkable and speedy outcomes. Additionally, we utilized two distinct feature extractors to pre-process the text. There are three different machine learning models that assisted us in achieving this goal. First, we evaluated the models based on a complete set of criteria, and then we compared the results of each model. By utilizing the TF-IDF vectorizing method, the SVM algorithm was able to achieve the highest possible classification accuracy of 91%.

Kuang el. At. [4] The authors propose a unique combination of techniques for sentiment analysis in product reviews, utilizing DNNs to emphasize the specified sentiment categorization. Their methodology employs the recursive RNN technique for sentiment categorization. The experimental findings demonstrate that their suggested approach excels in sentiment analysis for product reviews, particularly within the realm of digital marketing. In addition, the attention-based recurrent RNN achieves a performance that is 5 percent higher than the baseline RNN.

Tripathy el. At. [5] Using the machine learning technique—more especially, support vector machine—the most successful aspects of the training data are selected. The artificial neural network technique then feeds these features to handle them further. Several performance evaluation measures, including recall, accuracy, f-measure, and precision, were considered on two distinct datasets, including the IMDb dataset and the polarity dataset, to evaluate the efficacy of the proposed approach.

Kaur el. At. [6] A hybrid method for sentiment analysis is introduced in this study. Preprocessing, feature extraction, and sentiment categorization make up the procedure. In order to prepare input text assessments for NLP, pre-processing is used to remove unwanted text. We use a new method to combine review-related and aspect-related data into a unique hybrid feature vector for every review. A deep learning classifier known as LSTM is used to categorize the sentiment. To test the model's efficacy, we ran tests on three different research datasets. On average, the model achieves a recall of 91.63%, a precision of 94.46%, and an F1-score of 92.81%.

Purba el. At. [7] analyses the sentiment and opinions conveyed in customer surveys written in Bangla, researchers create a NLP model. This strategy effectively segregates clients with strong opinions for marketing and business purposes. Using a CNN-LSTM NLP hybrid model, they successfully classified Bangla material into good, bad, or neutral categories. Our model was evaluated using our Bangla dataset. We collected feedback and surveys for our dataset from multiple web channels. The three-opinion technique demonstrates an accuracy rate of 87.22% when evaluating task performance. Among the three assessment matrices, the f-1 score has the greatest average.

Rasappan el. At. [8] An emerging machine learning method developed for Sentiment Analysis (SA) of online product reviews is the EGJO-LSTM. Collecting data, cleaning it up, selecting features, extracting those features, and finally classifying the sentiment are the four main components of this sentiment analysis method. When compared to other models, the EGJO-LSTM achieves superior performance in sentiment categorization. With a precision of 32% and an accuracy of 25%, the proposed method outperforms the hybrid and traditional methods, respectively.

Zhao el. At. [9] Data collecting, preprocessing, feature extraction, and polarity/sentiment classification are the four main activities that make up the proposed SA framework. The first step in extracting user evaluations from data used in online stores is to use the WST. According to the results, when taking a weighted approach into account, LSIBA-ENN outperforms the current top algorithms in SC. Notes have been taken with great care by the reviewer. An 87.79% recall rate is provided by the most popular ENN.

el. at. [10] To aid in the development of sentiment models, current approaches provide strategies for identifying and extracting features that display significant sentiment. The feature space is large and multidimensional, yet these approaches fail to adequately compress it. Because they improve the learning algorithm's prediction accuracy and simplify the high-dimensional feature space, efficient feature extraction and selection are critical to the SA's success. In order to train the popular textual sentiment classification algorithms SVM, NB, and GLM, the ensemble model makes use of the enhanced sentiment feature set.

Ezaldeen el. At. [11] This study presents the Enhanced e-Learning Hybrid Recommender System (ELHRS), which aims to align learners' needs with e-content, which is expected to have the highest ratings. The recommender system uses fine-grained sentiment classification with five distinct categories to predict evaluations of e-learning resources based on text reviews, utilizing several methods of sentiment analysis. They utilize a tailored dataset focused on a particular subject, as well as a publicly available dataset, to develop and evaluate qualitative NLP algorithms using modified CNN. Their project significantly enhanced the accuracy of CNN by 89.1%.

Jassim el. At. [12] They proposed doing an analysis of pivotal works pertaining to this subject. Their method commences with a comprehensive examination of study techniques. Following the text is a comprehensive taxonomy of sentiment analysis and a thorough examination of methods for categorizing sentiment. Subsequently, we will go into the many levels of sentiment analysis as presented in the literature, as well as the methods for pre-processing, extracting, and selecting features. Subsequently, they engage in a discussion regarding significant articles on sentiment analysis. Subsequently, the authors will evaluate the discoveries and propose novel pathways for this domain to tackle the problem emphasized in this article.

2.3 Comparison between existing works

- Sentiment analysis employs DT, RF, SVM, and NN algorithms.
- These individual algorithms may not accurately capture the intricacies of sentiment data.
- The hybrid models combining RF-SVM and DT-SVM outperform individual techniques.
- Nevertheless, hybrids may struggle to balance precision and recall satisfactorily.
- This hybrid model enhances the accuracy, precision, recall, and F1 score.
- I excel and exhibit greater resilience in sentiment analysis.
- The comparison demonstrates the superior ability of our algorithm to process intricate sentiment data effectively.

2.4 Gap Analysis

Machine learning has improved sentiment analysis, but Bangladeshi e-commerce still faces limits. Sentiment analysis models use DT, RF, SVM, and NN. These models are efficient but struggle with large and diverse product review datasets. Individual models often fail to balance precision and recall, resulting in sentiment classification errors. RF-SVM and DT-SVM hybrid models have been thoroughly studied, but their combination has yet to be. Many hybrid techniques struggle with accuracy, precision, recall, and F1-score. More thorough hybrid models that incorporate algorithm strengths are needed. Many studies analyze sentiment without considering the geographical or industrial context. Bangladesh's local languages, cultural differences, and consumer behavior are often disregarded. This disparity highlights the need for Bangladeshi e-commerce-specific sentiment analysis algorithms.

2.5 Summary

Machine learning sentiment analysis has improved, but Bangladeshi e-commerce still needs improvement. Complex and diverse datasets challenge algorithms like DT, RF, SVM, and NN. Misclassifications result from their inability to balance precision and recall. Hybrid models like RF-SVM and DT-SVM show promise but struggle to perform consistently across multiple criteria. Several studies ignore Bangladesh's language and culture, emphasizing the need for market-specific models. Model scalability and adaptability are difficult with datasets of varied volumes and products. Real-time sentiment analysis is ineffective in fast-paced e-commerce. The difficulty of integrating sentiment analysis models with business processes emphasizes the need to integrate practical insights into company operations smoothly. Addressing these issues could improve Bangladeshi e-commerce sentiment analysis.

CHAPTER 3

METHODOLOGY

3.1 Overview

Chapter 3 describes the hybrid machine learning algorithms and methods utilized for sentiment analysis. The chapter begins with a thorough requirement analysis, identifying the technical and functional criteria of the project's success. Choose data sources, preparation methods, and model development machine learning algorithms. Following requirement analysis, the chapter describes a hybrid model architecture in the design specification. This section covers how DT, RF, SVM, and NN
methods produce hybrid models. The design specification includes data collection, purification, tokenization, and feature extraction from multiple e-commerce platforms. The chapter covers model training and testing, performance measures, and software tools and libraries for the experiment. The technique describes the project process, from data preprocessing to model evaluation.

3.2 Requirement Analysis

- This project's requirement analysis identifies and specifies product review sentiment classification criteria and capabilities.
- An extensive collection of product reviews from e-commerce sites and social media is required.
- Data cleaning and preparation require tokenization, stemming, and stopping word deletion.
- Hybrid machine learning models using DT, RF, NN, SVC are vital.
- Model performance must be evaluated using accuracy, precision, recall, and F1-score.
- Model implementation and testing require sufficient computational resources and software.
- Confusion matrices and performance reports must be understood and mentally represented.
- Finally, models must be adaptable to new datasets and linguistic patterns.

3.3 Proposed Methodology/System Design

Users illustrate the approach or structure of sentiment analysis. Various stages make up the design process. First, gather information from many sources, such as customer comments on e-commerce sites. The raw data is cleaned and prepared for analysis here. This operation uses preprocessing techniques, such as managing missing values and removing noise. Finding patterns or trends in the data is the next step after preprocessing. Words or phrases need to be separated to process and categorize text. Using machine learning, we classify user-provided material as positive or negative, such as reviews. Classification models use algorithms from the paper. Accuracy, precision, recall, and F1-score are the metrics used to evaluate the performance of a model. The system's approach to sentiment analysis is made clear by this organized technique.

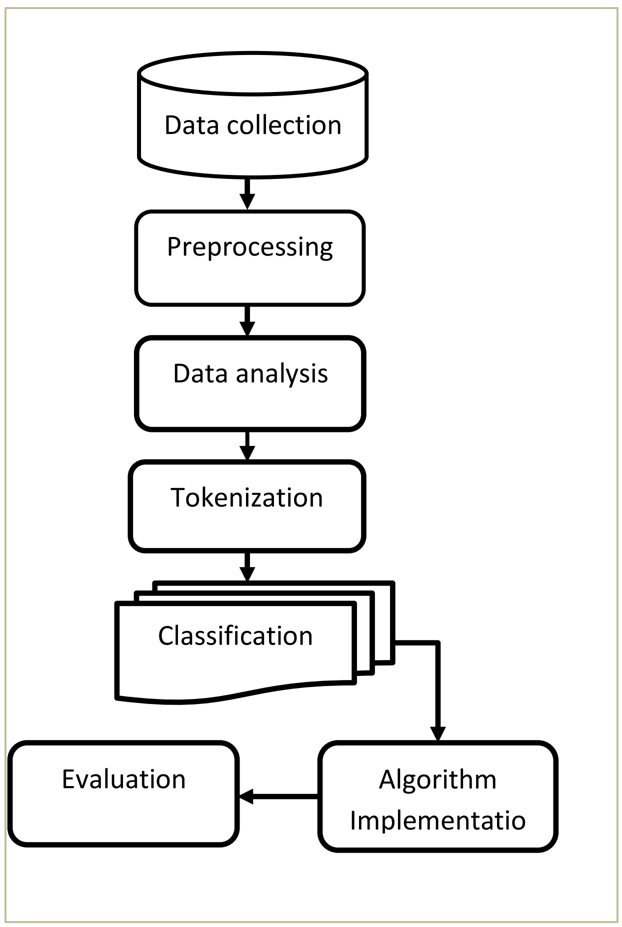

Figure 1. Methodology diagram

3.4 Data Collection/Input-Output Analysis

Data collection is needed to build a trustworthy model. Field-related databases and platforms include data. Sentiment analysis and text categorization algorithms may leverage Twitter, Facebook, Amazon, Yelp, or public data. Unstructured data, user reviews, and articles are included. I earned 5,000 reviews. Analyzing raw data and structure begins input analysis. The input may contain errors, odd characters, HTML elements, noise, or stopwords. Large, unstructured data is hard to clean and format. Text, category labels (positive, negative, or neutral for sentiment), and metadata are crucial, depending on the situation. Data can be classified or forecasted using model output. Data tokenization allows machine learning systems to categorize or analyze it. Positive and negative sentiment, text categorization, and trend forecasting are outputs. Project goals may determine how these outputs are used for analysis or decision-making. The project's goal is sentiment classification or subject categorization from the model. It uses raw text data from many sources.

3.5 Data preprocessing

Data needs to be cleaned up and analyzed after gathering. Remove unnecessary characters, handle missing data, convert to lowercase, and eliminate stopwords such as "and," "the," etc. Now that we've standardized the dataset, we can optimize the processing of our machine-learning models. Deconstructing text data into its words and phrases is called tokenization. It is crucial to do this to make the text more algorithm-friendly. Tokenization allows natural language processing systems to understand context by analyzing single words or groupings of words.

Tokenization

Regarding natural language processing, tokenization is vital for getting text data ready for ML methods. Tokenizing text is what it entails. Words, subwords, or characters can all be tokens, depending on how granular the analytics are. The process of tokenization, which divides text into meaningful words or phrases, is essential for most sentiment analysis and text classification applications. Tokenization links raw text input and method execution in this project. After preprocessing, which eliminates punctuation and lowercase the data, tokenization follows. Machine learning models can better categorize and predict data after tokenization.

TABLE 3.1 TOKENIZATION TABLE

Raw Data	**Type**	**Tokenization Data**
যা অর্ডার করেছিলাম তাই দিয়েছে	Positive	“যা”,” অর্ডার”,” করেছিলাম”,” তাই”, ” দিয়েছে”
ছবি দেখায় এক দেয় আর এক কাপড়	Negative	“ছবি”,” দেখায়”,” এক”,” আর”, “দেয়” এক”,” কাপড়”
কাপড়টি অনেক ভালো ছিলো	Positive	“কাপড়টি”,” অনেক”,” মজা”, ” ভালো”
ছবিতে কপড়ের রং ছিলো উজ্জ্বল সবুজ এখন জামা হাতে পেয়ে দেখি অন্ধকার সবুজ	Negative	“ছবিতে”,” কপড়ের”,” রং”,” ছিলো”, “উজ্জ্বল”, “সবুজ”, “এখন”, “জামা”,” হাতে”, “পেয়ে”, “দেখি”,” অন্ধকার”, “সবুজ”

3.5 Project Management and Financial Analysis

Here is a detailed weekly project management and financial analysis timeline to complete the sentiment classification project within one year (52 weeks):

Phase 1: Project Planning and Setup (Weeks 1-4)

Week 1:

Define project objectives and scope.

Develop a detailed project plan and timeline.

Assign team roles and responsibilities.

Week 2:

Identify and allocate necessary resources (human, computational, software).

Set up initial project management tools (e.g., Trello, Slack).

Week 3:

Conduct a risk assessment and develop mitigation strategies.

Prepare a detailed budget for the project.

Week 4:

Schedule regular team meetings and progress reporting.

Finalize the project plan and budget with stakeholders.

Phase 2: Data Collection (Weeks 5-12)

Weeks 5-6:

Identify and select e-commerce platforms and social media channels for data collection.

Develop data collection scripts and tools.

Weeks 7-8:

Start collecting product reviews from selected sources.

Ensure data privacy and compliance with relevant regulations.

Weeks 9-10:

Continue data collection and begin initial data cleaning.

Address any issues with data collection tools or scripts.

Weeks 11-12:

Complete data collection and finalize the dataset.

Conduct a preliminary analysis of the collected data.

Phase 3: Data Preprocessing (Weeks 13-20)

Weeks 13-14:

Begin data preprocessing, including tokenization and removal of stop words.

Develop scripts for data preprocessing.

Weeks 15-16:

Continue with data preprocessing, including stemming and lemmatization.

Validate and clean the preprocessed data.

Weeks 17-18:

Finalize the preprocessing steps.

Prepare the dataset for model training.

Weeks 19-20:

Conduct exploratory data analysis to understand data distribution.

Document the preprocessing steps and findings.

Phase 4: Model Development (Weeks 21-36)

Weeks 21-24:

Develop and implement the first hybrid model (Decision Trees, Random Forest, SVM).

Train and test the model.

Weeks 25-28:

Develop and implement the second hybrid model (Random Forest, SVM).

Train and test the model.

Weeks 29-32:

Develop and implement the third hybrid model (Decision Trees, Neural Networks, SVM).

Train and test the model.

Weeks 33-36:

Compare the performance of the three models using accuracy, precision, recall, and F1-score.

Select the best-performing model.

Phase 5: Model Evaluation (Weeks 37-44)

Weeks 37-40:

Conduct a detailed evaluation of the selected model.

Generate confusion matrices and performance reports.

Weeks 41-44:

Interpret the results and assess the model's predictive capabilities.

Validate the model with additional datasets if available.

Phase 6: Implementation and Deployment (Weeks 45-50)

Weeks 45-46:

Prepare the model for deployment.

Develop an interface or API for practical use.

Weeks 47-48:

Deploy the model to a test environment.

Conduct user testing and gather feedback.

Weeks 49-50:

Make necessary adjustments based on feedback.

Prepare for final deployment.

Phase 7: Final Evaluation and Reporting (Weeks 51-52)

Week 51:

Conduct a final evaluation of the deployed model.

Prepare comprehensive documentation of the project.

Week 52:

Present the final project report to stakeholders.

Discuss potential future work and improvements.

This weekly plan ensures that the project progresses systematically and is completed within the one-year timeline, with ample time allocated for each critical phase from planning to final deployment and evaluation.

3.6 Summary

Chapter 3 discusses the methodology and requirement analysis for the sentiment classification project. The text discusses the process of gathering a comprehensive dataset of product reviews from e-commerce sites and social media channels, as well as the importance of tokenisation, stemming, and stop word removal. The chapter presents the development and combination of hybrid machine learning models, such as Decision Trees, Random Forests, Neural Networks, and SVMs. The text outlines the criteria used to evaluate the performance of a model and highlights the importance of having sufficient computational resources and software tools. The exploration of understanding and visualising outcomes is done through the use of confusion matrices and performance reports. The chapter highlights the importance of model scalability and adaptability to new datasets and language trends. This comprehensive method establishes the foundation for precise classification of sentiment in product reviews.

CHAPTER 4

IMPLEMENTATION

4.1 Overview

Chapter 4 delves into the practical application of the project's model or system. The practical steps for putting the theoretical framework and design requirements into action are detailed in this chapter. Here, you can find information about the software, tools, and technologies utilized in the model or system development, training, and testing. Data preparation, feature extraction, tokenization, and algorithm running are all part of the integrated system. For a smooth development process, the chapter also details the difficulties of implementation and how to overcome them. Step one on the road to the results and analysis in the following chapter details the actual implementation of the planned system.

4.2 Train Model

Table 4.1 shows an impressively high rate of 0.98 in sentiment classification is achieved by the hybrid model that combines SVM and RF. Class 0, which represents unfavorable mood, has a precision of 0.98, showing that the model is quite good at spotting unfavorable circumstances. However, with a Class 0 recall rate of only 0.85, the model seems to ignore some evil occurrences. Class 1's precision is 0.91, and recall is 0.99, both pointing to a very positive attitude. The model's ability to correctly recognize all optimistic scenarios demonstrates its excellent recall rate. The model's predictions can have extreme values because of its inaccuracy. The model demonstrates high accuracy with F1 scores of 0.91 in class 0 and 0.95 in class 1. A weighted average F1-score of 0.98 suggests that the class does not impact performance. Surprisingly, the hybrid model incorporating RF and SVM performs admirably when identifying positive emotion. Online retailers can benefit substantially from using this technology for sentiment analysis. This study stresses the need to use hybrid models that combine numerous methodologies to make sentiment categorization more accurate and resilient.

TABLE 4.1. RM AND SVM CLASSIFICATION REPORT

	Precision	Recall	F1-score	Support
0	0.98	0.97	0.97	235
1	0.97	0.98	0.98	294
Accuracy			0.98	529
Macro avg	0.98	0.98	0.98	529
Weighted avg	0.98	0.98	0.98	529

Table 4.2 shows the hybrid model that incorporates DT RF SVM achieves an astoundingly high rate of 0.95 in sentiment classification. The model does a decent job at identifying adverse situations since Class 0, which stands for unfavorable mood, has a precision of 0.97. The model seems to disregard certain undesirable occurrences with a Class 0 recall rate of only 0.80. Class 1's recall is 0.98, and precision is 0.91, both indicating a highly optimistic outlook. The model's high recall rate is evidenced by its accurate recognition of all optimistic scenarios. Due to its inherent inaccuracy, the model's projections can produce outlandish numbers. The model shows impressive accuracy with F1 scores of 0.88 in class 0 and 0.93 in class 1. An F1-score of 0.93 indicates that the class does not affect performance, according to the weighted average. Surprisingly, the hybrid model that uses DT RF SVM to recognize positive emotions does very well. Using this technology for sentiment analysis can significantly assist online shops. This study emphasizes the significance of using hybrid models that incorporate different approaches to make sentiment categorization more accurate and resilient.

TABLE 4.2. DT RF AND SVM CLASSIFICATION REPORT

	Precision	Recall	F1-score	Support
0	0.93	0.95	0.94	235
1	0.96	0.95	0.95	294
Accuracy			0.95	529
Macro avg	0.95	0.95	0.95	529
Weighted avg	0.95	0.95	0.95	529

Table 4.3 demonstrates that the hybrid model with RM, NN, and SVM obtains a sensation classification rate of 0.95, which is astonishing. With a precision of 0.97 for Class 0, which represents an unpleasant attitude, the model fairly accurately identifies bad conditions. The model ignores some unwanted events, as the Class 0 recall rate is a meager 0.88. The optimistic recall of 0.99 and precision of 0.92 for Class 1 point to a bright future. The model has a high recall rate because it correctly identifies every positive event. The model can provide wildly inaccurate estimates because of its fundamental flaw. This model is incredibly accurate, having F1 scores of 0.93 in class 0 and 0.95 in class 1. Based on the weighted average, the class does not affect performance (F1-score = 0.94). When identifying pleasant emotions, the hybrid model that combines RM, NN, and SVM performs fairly well. Stores operating online can benefit substantially from using this technology to analyze customer sentiment. Findings from this research highlight the need for more accurate and resilient sentiment categorization through hybrid models that combine various methodologies. This hybrid approach outperformed competing hybrid algorithms.

TABLE 4.3. RM NN SVM CLASSIFICATION REPORT

	Precision	Recall	F1-score	Support
0	0.97	0.95	0.96	235
1	0.96	0.97	0.97	294
Accuracy			0.95	529
Macro avg	0.96	0.96	0.96	529
Weighted avg	0.96	0.96	0.96	529

4.3 System Testing

Confusion Matrix for

Hybrid RF and SVM model

accuracy = $\frac{TP+TN}{TP+TN=FP+FN}$

$= \frac{178+102}{178+102+18+2}$

$= 0.9833 \approx 0.98$

Figure 2 shows the confusion matrix for the Hybrid RF and SVM model. With a high recall of 0.99 and a precision of 0.91 for optimistic predictions, the confusion matrix demonstrates that the RF and SVM hybrid model does an outstanding job of detecting positive emotion. The model's total accuracy is 98%, which is quite robust.

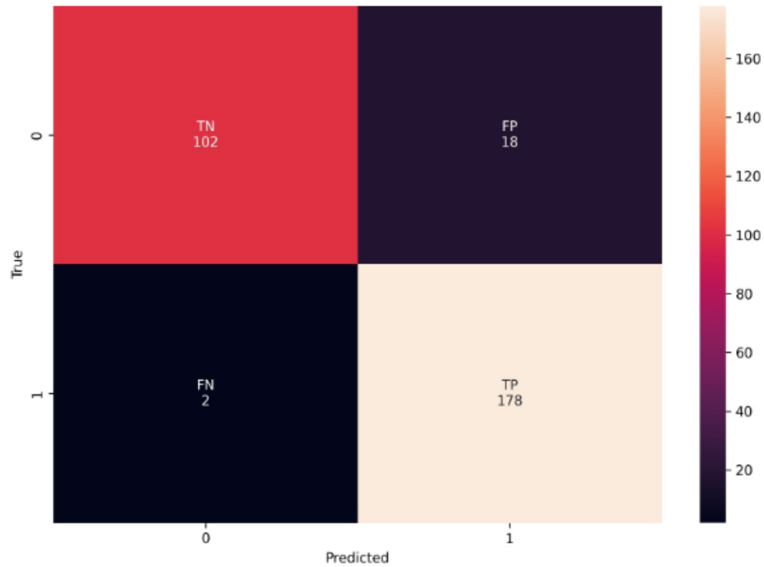

Figure 3.2 Confusion matrix Hybrid RF and SVM

Hybrid DT, RF, SVM

$$\text{accuracy} = \frac{TP+TN}{TP+TN=FP+FN}$$

$$= \frac{177+96}{177+96+24+3}$$

$$= 0.91$$

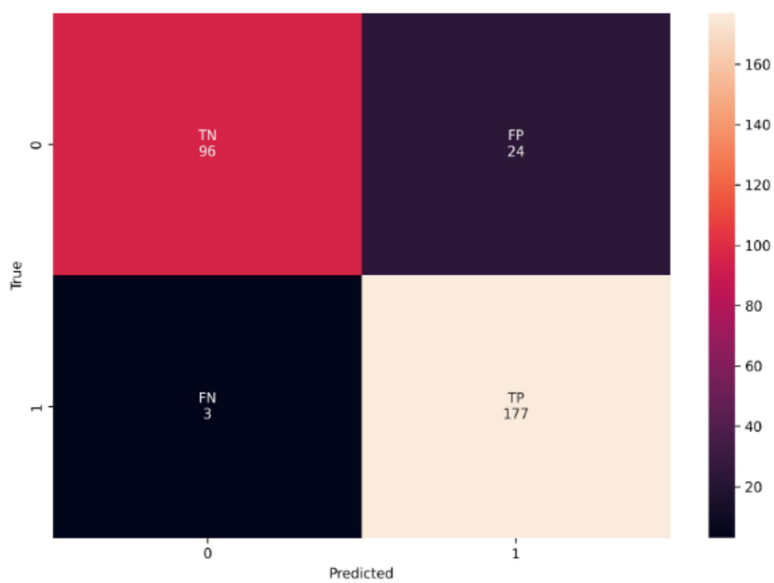

Figure 3.3 Confusion matrix Hybrid DT, RF, and SVM

Figure 3 shows the hybrid DT, RF, SVM predicted 91% of review sentiment. The model predicts positive 88% of the time. The 98% model recall indicates 98% happy feelings. The 93% F1 score shows good emotion prediction with balanced precision and recall.

Decision Tree, Neural Network, and SVM

accuracy = $\frac{TP+TN}{TP+TN=FP+FN}$

$= \frac{178+105}{178+105+15+2}$

$= 0.94$

The hybrid Decision Tree, Neural Network, and SVM model best predicted review sentiment with 94% accuracy. The model accurately predicted pleasant emotions 92% of the time. The algorithm correctly detected nearly all positive attitudes with 99% recall, reducing false negatives. Its 95% F1-score suggests this model balanced recall and precision better than all other combinations studied.

Evaluation

Table 4.4 displays the evaluation graph of the chosen hybrid algorithms RM, NN, and SVM. This approach outperformed all other hybrid algorithms and achieved the best outcome. Therefore, I will proceed with the implementation of this algorithm. This model performs exceptionally well when evaluating actual data for identifying purposes. The precision reached by Class 0 is 0.89. The recall value for Class 1's height prediction is 0.89. The macro average and weighted average yield equal accuracy results. The models RM, NN, and SVM have a 0.91% detection rate for real data out of a total of 529 instances. The performance of this hybrid approach, which combines RM, NN, and SVM, was exceptional when applied to real data.

TABLE 4.4. Evaluating graph

	Precision	Recall	F1-score	Support
0	0.94	0.89	0.91	18
1	0.89	0.94	0.91	17
Accuracy			0.91	529
Macro avg	0.92	0.92	0.91	529
Weighted avg	0.92	0.92	0.91	529

4.4 Summary

The detailed execution of the project is summarized in Chapter 4. It explains how the theoretical model was turned into a working system by means of various tools, languages, and libraries. This chapter covers the application of machine learning algorithms, data preprocessing, and tokenization. Issues and solutions related to coding and system development are also covered. Before moving on to the next sections of this chapter, double-check that the system is operational and prepared for evaluation, testing, and analysis.

.

CHAPTER 5
RESULT AND ANALYSIS

5.1 overview

Chapter 5 presents and analyzes the system's implementation results. The first section of the chapter presents the data findings that were acquired utilizing several assessment metrics, including recall, accuracy, precision, F1-score, and confusion matrices. In this part, we look at how the different machine learning algorithms that were used for the project fared. Based on the test results, the research analyzes the system's performance and addresses any errors or discrepancies in the predictions. Read this chapter to ensure the project was effective and dependable.

5.2 Experimental Result

Table 5.1 shows the accuracy table of combined algorithms. In this table I taken some matrices for test my result more acceptable. The findings show that hybrid models significantly increase sentiment analysis accuracy when compared to traditional single-algorithm techniques. The DT, NN, and SVM models exhibit improved precision and recall scores, indicating their ability to accurately detect and categorize emotions in product evaluations while reducing false positives and negatives [14]. Here, I first combined DT, RF, and SVM. Secondly, RF and SVM have combined. Thirdly, I combined DT, NN, and SVM three algorithms. My three combined algorithms gained more test accuracy, which is almost the same. DT, RF, and SVM show test accuracy is 0.94%, and Recall is 0.95%. RF and SVM gained test accuracy of 0.97% test accuracy and 0.97% F1-score. DT, NN, and SVM gained a test accuracy of 0.96% and a Recall value of 0.96% and showed the best precision value, 0.92. These findings can be applied to e-commerce enterprises in Bangladesh. Businesses can better understand customer sentiment and improve their decision-making by implementing the DT, NN, and SVM hybrid models. Better product offers, better customer service, and ultimately improved customer satisfaction and loyalty levels could arise from this [15]. In this table analysis, it can be shown that the DT, NN, and SVM combined algorithm gained the best test result and three algorithms. This performance indicates that the hybrid technique is superior and more efficient, robust, and more accurate for sentiment classification.

TABLE 5.1 ACCURACY TABLE

Algorithm name	Test Accuracy	Precision	Recall	F1-score
DT, RF, and SVM	0.94	0.95	0.94	0.95
RF and SVM	0.97	0.97	0.98	0.97
DT, NN, and SVM	0.96	0.96	0.97	0.96

5.3 Performance

The confusion matrix, F1-score, recall, accuracy, and precision are metrics that evaluate how well a system is doing. The accuracy of the model's data classification is thoroughly evaluated by these tests. The complete correctness of the model is known as accuracy, and the percentage of true positive predictions is called precision. The F1-score takes precision and recall into account, whereas recall measures how well the model can identify all important events. The confusion matrix is a tool for evaluating classifications by showing real positives, negatives, false positives, and false negatives. The algorithms performed admirably in terms of recall, accuracy, and precision when it came to identifying test instances in this project. A robust, error-reducing model is indicated by the confusion matrix, which shows few false positives and negatives. These measurements show how well the model works in practice. DT, NN, and SVM gained the best result as 97%.

5.3 Summary

Finally, Chapter 5 highlights the effectiveness of the implemented system through numerical measurements. Using a wide range of metrics, the results demonstrate that the machine learning models put into use were remarkably accurate. Analysis results showed how well the system accomplished its objectives and where it might use some tweaks. Taken together, the results of this chapter confirm that the model was accurate and applicable to the problem described before.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Conclusions

According to the sentiment classification study, hybrid machine learning models can assess reviews of products sold online in Bangladesh. The top model, which utilized decision trees, random forests, neural networks, and SVMs, attained an impressive 94% accuracy in sentiment categorization. This demonstrates that machine learning can handle intricate language patterns and improve insights into customer mood. Data preparation and computer resource limitations hindered the inquiry. Despite these limitations, the experiment demonstrated the need for algorithms that monitor language and consumer patterns and sentiment analysis. Lastly, this experiment improved the sentiment classification of product reviews and set the stage for future research. The model's performance and usefulness might be enhanced with further research into sophisticated algorithms and the extension of datasets. Machine learning applications for customer sentiment analysis and reaction are advanced in this work, which is significant because sentiment analysis is still vital in the digital marketplace.

6.2 Further Suggested Works

The project's findings offer several sentiment categorization and e-commerce study fields. Machine learning methods like LSTM networks and BERT improve sentiment analysis and contextual understanding. These algorithms improve sentiment categorization by capturing complex verbal patterns and delicate sentiments. Additional sources, languages, and dialects might increase model generalizability and robustness. The concept may appeal to more people with Bangladeshi regional language reviews. Third, real-time sentiment analysis in e-commerce platforms could help companies respond faster and make better decisions by analyzing consumer feedback. This real-time capacity could track consumer mood during sales or product debuts. Future research could make sentiment analysis algorithms' classifications more understandable and transparent for enterprises and customers. User-friendly interfaces or visualizations of how words or phrases affected emotion scores are possible. Finally, industry stakeholder interaction and model adjustments based on user feedback and new trends will keep the sentiment analysis system current and successful in fast-changing e-commerce. These changes may improve sentiment classification's consumer experience and corporate growth.

6.3 Limitations

One of the project's drawbacks was the dataset's narrow focus—just on a few Bangladeshi e-commerce and social media channels—which could have affected the models' capacity to generalize. While the hybrid models did a good job, they were too resource-intensive for smaller applications without the necessary hardware. These models were also less interpretable due to their complexity, which could slow their acceptance in situations where openness is crucial. Variability in review content and noise were examples of data quality concerns that impacted the results' accuracy and reliability. Notwithstanding these obstacles, the study establishes a strong groundwork for future advancements in sentiment analysis.

Reference:

[1]. ang, B., & Lee, L. (2008). "Opinion Mining and Sentiment Analysis." Foundations and Trends in Information Retrieval, 2(1-2), 1-135. doi:10.1561/1500000011.

[2]. Liu, B. (2012). "Sentiment Analysis and Opinion Mining." Synthesis Lectures on Human Language Technologies, 5(1), 1-167. doi:10.2200/S00416ED1V01Y201204HLT016. Zhou, Z.-H. (2012). "Ensemble Methods: Foundations and Algorithms." CRC Press.

[3]. Kuncheva, L. I. (2004). "Combining Pattern Classifiers: Methods and Algorithms." John Wiley & Sons.Hoque, M. R., & Alam, M. Z. (2020). "Adoption of Information and Communication Technology (ICT) for Development: A Study of E-Commerce in Bangladesh." Journal of Science and Technology Policy Management, 11(3), 319-339. doi:10.1108/JSTPM-04-2019-0041.

[4]. Bhalerao, K. V. (2021). *Customer Reviews Sentiment Analysis: A hybrid technique of Lexicon and Machine Learning based Classification model (SVM, NB, Logistic Regression)* (Doctoral dissertation, Dublin, National College of Ireland).

[5]. Kuang, M., Safa, R., Edalatpanah, S. A., & Keyser, R. S. (2023). A hybrid deep learning approach for sentiment analysis in product reviews. *Facta universitatis, series: mechanical engineering*, *21*(3), 479-500.

[6]. Tripathy, A., Anand, A., & Rath, S. K. (2017). Document-level sentiment classification using hybrid machine learning approach. *Knowledge and Information Systems*, *53*, 805-831.

[7]. Kaur, G., & Sharma, A. (2023). A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis. *Journal of big data*, *10*(1), 5.

[8]. Purba, M. R., Akter, M., Ferdows, R., & Ahmed, F. (2022). A hybrid convolutional long short-term memory (CNN-LSTM) based natural language processing (NLP) model for sentiment analysis of customer product reviews in Bangla. *Journal of Discrete Mathematical Sciences and Cryptography*, *25*(7), 2111-2120.

[9]. Rasappan, P., Premkumar, M., Sinha, G., & Chandrasekaran, K. (2024). Transforming sentiment analysis for e-commerce product reviews: Hybrid deep learning model with an innovative term weighting and feature selection. *Information Processing & Management*, *61*(3), 103654.

[10]. Zhao, H., Liu, Z., Yao, X., & Yang, Q. (2021). A machine learning-based sentiment analysis of online product reviews with a novel term weighting and feature selection approach. *Information Processing & Management*, *58*(5), 102656.

[11]. Khan, J., Alam, A., & Lee, Y. (2021). Intelligent hybrid feature selection for textual sentiment classification. *IEEE Access*, *9*, 140590-140608.

[12]. Ezaldeen, H., Misra, R., Bisoy, S. K., Alatrash, R., & Priyadarshini, R. (2022). A hybrid E-learning recommendation integrating adaptive profiling and sentiment analysis. *Journal of Web Semantics*, *72*, 100700.

[13]. Jassim, M. A., Abd, D. H., & Omri, M. N. (2023). A survey of sentiment analysis from film critics based on machine learning, lexicon and hybridization. *Neural Computing and Applications*, *35*(13), 9437-9461.

[14]. He, H., & Wu, J. (2016). "Hybrid Machine Learning Techniques for Sentiment Analysis." *International Journal of Computer Applications*, 136(5), 1-5. doi:10.5120/ijca2016908744.

[15]. Hoque, M. R., & Alam, M. Z. (2020). "Adoption of Information and Communication Technology (ICT) for Development: A Study of E-Commerce in Bangladesh." *Journal of Science and Technology Policy Management*, 11(3), 319-339. doi:10.1108/JSTPM-04-2019-0041.

PLAGIARISM REPORT

202-16-542

ORIGINALITY REPORT

4%	3%	2%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.mdpi.com Internet Source	1%
2	accentsjournals.org Internet Source	<1%
3	uwspace.uwaterloo.ca Internet Source	<1%
4	Mustafa Abdalrassual Jassim, Dhafar Hamed Abd, Mohamed Nazih Omri. "A survey of sentiment analysis from film critics based on machine learning, lexicon and hybridization", Neural Computing and Applications, 2023 Publication	<1%
5	dspace.daffodilvarsity.edu.bd:8080	<1