POPULAR COSMETIC PRODUCT ON ONLINE PLATFORM ANALYZING BENGALI REVIEW - A MACHINE LEARNING APPROACH

STUDY BY

Suzana Zannet Tumpa

ID: 172-15-10055

Rifa Shanzida Khanam

ID: 172-15-9780

AND

Plaboni Roy

ID: 172-15-9735

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

Supervised By

Tapasy Rabeya

Lecturer

Department of CSE

Daffodil International University

Co Supervised By

Zerin Nasrin Tumpa

Lecturer

Department of CSE

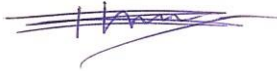Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

MAY 2021

APPROVAL

This Project/internship titled **"Predict popular cosmetic product on online platform analyzing Bangali review: A machine learning approach"**, submitted by Suzana Zannet Tumpa (172-15-10055) and Rifa Shanzida Khanam (172-15-9780) and Plaboni Roy (172-15-9735) to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 31-05-2021.
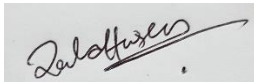
<u>BOARD OF EXAMINERS</u>

Chairman

_____

Dr. Touhid Bhuiyan

Professor and Head

Department of Computer Science and Engineering

Faculty of Science & Information Technology Daffodil

International University

Internal Examiner

_____

Zahid Hasan

Assistant Professor

Department of Computer Science and Engineering

Faculty of Science & Information Technology

Daffodil International University

Internal Examiner

Md. Riazur Rahman

Assistant Professor

Department of Computer Science and Engineering

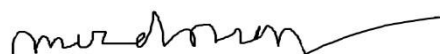Faculty of Science & Information Technology

Daffodil International University

External Examiner

Dr. Md Arshad Ali

Associate Professor

Department of Computer Science and Engineering

Hajee Mohammad Danesh Science and Technology University

DECLARATION

We hereby declare that this project has been done by us under the supervision of **Tapasy Rabeya ,** **Lecturer, Department of CSE,** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.
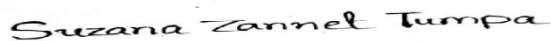
Supervised by:

_____

Tapasy Rabeya

Lecturer

Department of CSE

Daffodil International University

Submitted by:

_____

Suzana Zannet Tumpa

ID: 172-15-10055

Department of CSE

Daffodil International University

Rifa Shanzida Khanam.

_____

Rifa Shanzida khanam

ID: 172-15-9780

Department of CSE

Daffodil International University

Plaboni Roy

_____

Plaboni Roy

ID: 172-15-9735

Department of CSE

Daffodil International University

ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes it possible to complete the final year project/internship successfully.

We are really grateful and wish our profound indebtedness to **Tapasy Rabeya, Lecture,** Department of CSE, Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "Sentimental Analysis" 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.

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

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

Finally, we must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

Nowadays people of different ages are getting used to internet service in their daily life. It has become the most valuable source of getting information, generating ideas, reviews for a product and choosing the right one. Millions of reviews are generated on the internet about products, so it is difficult to understand which product is actually useful. Sentiment analysis is such a research area which can understand and extracts the opinion from the reviews. Our research report reflects an experimental study on product review. It has been noticed that various research has been done on sentiment analysis fields from English text or other languages such as Chinese. But detecting sentiment from Bengali texts still needs a lot of attention. In this research, we have worked with Bengali product review comments or sentences and get out the most popular products. We have considered them either positive or negative. To bring a better result, we have used four methods and found a satisfactory accuracy.

TABLE OF CONTENTS

CONTENTS	PAGES

CHAPTERS

Chapter 1: Introduction

Chapter 2: Background Study

Chapter 3: Research Methodology

Chapter 4: Experimental Result and Discussion

Chapter 5: Impact on Society

Chapter 6: Summary, Conclusions, Recommendation and Implication for Future Research

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Introduction

Machine learning approach defines sentiment analysis, the most ongoing field of research in text mining has been successfully applied to analyzing and taking decisions on huge amounts of data which are collected from different sources. Text doesn't always contain only informative news, it can also determine the emotion or feelings on different topics. Here we demonstrate the prediction of popular cosmetic products on some online platforms analyzing Bangla reviews or comments of the users who have already used the product.

Lots of work has been done on sentiment analysis where the sentence or text determines the sentiment expression either positive or negative and sometimes it can be neutral. We have worked with Bangla platform. As this platform still requires lots of attention.

1.2 Motivation:

In the natural language processing technique, we use sentiment analysis to settle whether data is positive, negative, or neutral. Nowadays sentiment analysis is being used in many sectors like social media monitoring customer support, customer feedback, brand monitoring and reputation management, the voice of the customer, the voice of the employee, market research, competitive research, and product review or product analysis. We want to use sentiment analysis for predicting popular cosmetics products on an online platform analyzing Bangla reviews. Because there has been a lot of such work in different languages outside of the world but the number of such works in the Bengali language is very low or not. Moreover, the Bengali language is the fourth most spoken language in the world.

As we know the Bangladesh government has taken some steps in this regard and those are speech synthesis, recognition of speech, information retrieval and machine translation, etc. Natural

language processing is the technology which is used to deteriorate the communication between machine and human. In addition, several areas of NLP offer a better future and solutions. Notable among them manufacturing, healthcare, agriculture/food processing, education, retail/customer engagement, aid for differently-abled/accessibility technology, human and robot interaction/intelligent automation, public utility services, general/other issues related to AI, big data, national security, enablers for AI technology development, enablers for AI entrepreneurship.

So we think this work will help people to choose their right product.

1.3 Objectives:

Our main objective is to authorize Bengali speaking people to buy popular cosmetic products online with less hassle based on reviews.

We are trying to do a computation with sentiment analysis using Machine Learning Approach so that people can easily understand which cosmetic company's product is good and effective. Through this, Bengali speaking people will be able to buy good quality cosmetics without any inconvenience.

The biggest problem with buying cosmetics online is people do not understand which product is compatible with their skin or complexion. Because there are no skilled workers online like the store that will help people to understand which product is right for them. We often hear around us that many people have succumbed to cancer by using a bad or substandard cosmetic product.

By adopting this proposal, we can help people to buy the right cosmetic products they need through the reviews. By the establishment of our system we want to give a useful message to those people who say buying products online means just cheating. Moreover, our main goal is to make it easy for people to understand which company's products are good and effective.

1.4 Research Question

Research question 1: How do people understand product reviews are positive or negative?

Research question 2: How reviews help people when buying cosmetic products?
Research question 3: Do the comments contain any accurate information about the most popular products?

1.5 Expected outcome

We have chosen to predict popular cosmetic products based on Bengali reviews because there are many such works in many languages but the number of such works in Bengali is negligible. Our expected outcome will be that Bengali speaking people can see the Bengali reviews and buy the effective product for them.

1.6 Report Layout

Chapter 1: Introduction

We explain here about the introduction, Objective, motivation of the work, expected outcome and also the layout of the report of our project.

Chapter 2: Background Study

Background part of our research is discussed in chapter 2. We also explain literature review, related works, argumentation of the problems and claims about our project in this portion.

Chapter3: Methodology

In this chapter we have included overall procedures that we have used to build this proposed system. Methods are explained here step by step.

Chapter 4: Design Specification

Here the section is for discussing the graphical interface and display to the users.

Chapter 5: Implementation and Testing

Our empirical results are shown along with the performance analysis that is achieved by the proposed system. We have also covered the results summary in this chapter.

Chapter 6: Conclusions and Future plans

In this part of the report we have discussed the conclusion and future aims of our proposed system for general people.

CHAPTER 2

BACKGROUND STUDY

2.1 Introduction

Sentiment analysis is an on-going process to understand a human's mind through text mining. This process mainly helps to identify the state of human psychology. It also determines and sorts out subjective information from given source material. The most important task before detecting an emotion is to detect the sentiment of the specific sentence. So it would be easier to sort out the possibility of how much it is positive or negative the sentence is. After analyzing the sentiment in positive or negative the sentence could be categorized as what it actually meant. We have mainly worked on how to compute sentiment from Bengali Text of different cosmetic products from online platforms and predict the most popular cosmetic product.

Because of the lack of Bengali resources and its complexity, there are very few researchers who have worked on Bengali language. Though Bengali Language is considered as the sixth most popular language in the world and spoken by a population of over 250 million people.

From a Bengali Text it is very difficult and complicated to carry out the sentiment. In this research, we have tried to identify the positivity and negativity of a given Bengali sentence. We have trained our data and used it to sort out the possibility of a given sentence and give an output whether the sentence is positive or negative. If positive it's a good review otherwise it's a negative review.

Depending on the review, the most popular cosmetic product can be detected.

2.2 Literature Review

Xing Fang* and Justin Zhan [1] believe that there are some flaws in the online review data. The first of these errors is that people feel free to express their opinions but are skeptical about the quality of their opinions. The second error is that most of these data grounds are false and here they also show two ways to fix the error. As a solution to fix the error, they say that they should be checked before posting every review or comment. To fix the second error, they said, reviews of each product should

be included in the rating system. The highest value will be 5 and the lowest value will be 1. They collect over 5.1 million data on product review for their work from amazon.com. From February 2014 to April 2014, they collect information on reviewer id, product id, rating, review time, benefits and review text from the reviews.

Ankur Goyel, Jyoti Gautam and Sitesh Kumar[2] admitted that on twitter people tweet their precious views, opinion and experiences on trade and marketing. They have shown that use of SentiWordNet along with Naive Bayes can improve accuracy of classification of tweets providing a positive, negative and objective score of words present in tweets. For actual implementation of this system python with NLTK and python-twitter APIs are used.

Rudy Prabowo, Mike Thelwall[3] worked with rule-based classification,supervised learning and machine learning into a new combined method to do sentiment analysis . Their basic target was to bring sentiment analysis in a new approach. They tested the method on movie reviews, product reviews and Myspace comments. After testing the method the results showed that a hybrid classification can improve the effectiveness in terms of micro and macro averaged that takes both the precision and recall of a classifier's effectiveness into account.

Apoorv Agarwal, Boyi Xie, Ilia Vovsha, Owen Rambow, Rebecca Passonneau [4] said that the world's most popular microblogging website Twitter categorizes their tweets with three types of sentiment. They are neutral sentiment, positive sentiment and negative sentiment along with they have created two models for these classification tasks by demonstrating on tree kernel based model, feature based model and unigram model. The two new models are to classify sentiment into positive and negative classes which is a binary task as well as a multiple task of classifying into neutral, positive and negative.

Rajesh Bose, Raktim Kumar Dey, Sandip Roy and Debabrata Sarddar [5] suggested how to provide customer satisfaction and identify customer behaviors with the popular initial eight emotions through sentiment analysis. They also show that in many cases customer reviews or comments differ from ratings as well as how these things affect digital marketing.

Doa Mohey El-Din Mohamed Hussein [6] presents a survey on the sentiment analysis challanges relevent to their approaches and techniques. In the evolution of the internet as websites, social networks, blogs, online portals, reviews, opinions, recommendations, ratings and feedback are generated by users. So this sentiment content can be about anything personal or proffesional requirements where the author of the paper generated content requires using the text mining techniques and sentiment analysis.

2.3 Comparative Studies

The variable results have a significant impact on the methods used in the studies on the prediction of popular cosmetics and on some of these methodological factors on the results of the studies. The results of the studies suggest different research methods and definitions of key variables. Further research would help to clarify a more consistent methodology (ideally experimental schemes) to determine the prediction of popular cosmetics on online platforms.

2.4 Challenges

Our research "predict popular cosmetic products on online platforms analyzing Bengali review- a machine learning approach" is quite challenging work for us. In our system we have to collect a lot of Bengali reviews on popular cosmetic products which is very difficult for us. Because nowadays people are more comfortable giving reviews in English. So in the midst of so many English reviews, we have had to hurry to find Bengali reviews. And also showing exact proper information is so challenging that we had needed the proper reviews for collecting.

Chapter 3

Research Methodology

3.1 Introduction

In the previous chapter we have discussed the related works of sentiment analysis. We have also explained about some different types of procedures that were used in previous. In this chapter we will bring out the outline of research methods that were carried out to detect emotion which would be positive or negative from Bengali Text. It provides information about how data can be collected by applying some techniques to sort out the emotion of the experience of using products. Our research model and the reasons behind choosing it, also discussed in detail. It also provides the methods which were used to analyze the data. Finally, the implementation issues and requirements that were followed in the process are also discussed.

3.2. Research and Instrumentation

3.2.1 Research Subject

The main goal of this research is to detect sentiment from a given Bengali text from different types of product review and be able to predict the most popular product by using different types of methods.

Sentiment analysis is the very specific way in case of finding the emotion of a sentence. A sentence with no sentiment may not have any emotion. Our research subject is to find the sentiment of a sentence to see how people generally express their emotion while using a product they need and get experience from it.

3.2.2 Instrument

For our research purposes, we have collected around 400 Bengali sentences from different online sources. Our work is to detect and predict popular cosmetic products by applying sentiment analysis. Some well performed algorithms like Multinomial Naive Bayes, Random forest, Decision tree, Support Vector Machine etc are used in case of sentimental analysis. Therefore these algorithms give a very high accuracy of almost 100%.

3.3 Methodologies

Here we are working with four algorithms:

A. Multinomial Naive Bayes

B. Random Forest

C. Decision Tree
D. SVM

A. Multinomial Naive Bayes: Multinomial Naive Bayes classification algorithm tends to be a baseline solution for sentiment analysis work. We have used this technique in our research work. This method is highly sophisticated with its classification method. This is easy to build and particularly useful for long datasets. This method is very easy and fast to predict the dataset. It is simple and can be used for predicting real-time applications and highly scalable and can easily handle large datasets.

B. Random Forest: The random forest is a classification algorithm consisting of many decision trees. It uses bagging and features randomness when building each individual tree. It uses bagging and features randomness when building each individual tree to try to create an uncorrelated forest of trees whose prediction by committee is more accurate than that of any individual tree.

C. Decision Tree: We have used this method in our work. Decision trees always provide an effective method of decision making. Decision tree is used to solve the classification problem. Common usages of decision trees are valuable selection, prediction, data manipulation.

D. SVM: Support vector machine (SVM) is a learning technique that performs well on sentiment classification. It is a supervised machine learning algorithm that can be used for both classification or regression challenges, but most of the time it is used in classification problems. Here in our work, we have used this algorithm.

3.4 Data collection procedure

A good research mainly relied on both primary and secondary data in order to come up with accurate and objective findings. Primary data is the raw data which is mainly used for the original purpose which is directly taken from the field by interviews and questionnaires. Whereas secondary data is collected for purposes other than the original use. The research has been carried out using secondary data. The main intention was to create a proper trained data set consisting of Bengali review keywords. The main data source of data is Bengali cosmetic review from many online websites. As we know, sharing and discussing reviews freely in any online place could be counted as a source. Figure 3.1 shows the collection of our raw data which we have collected from different websites.

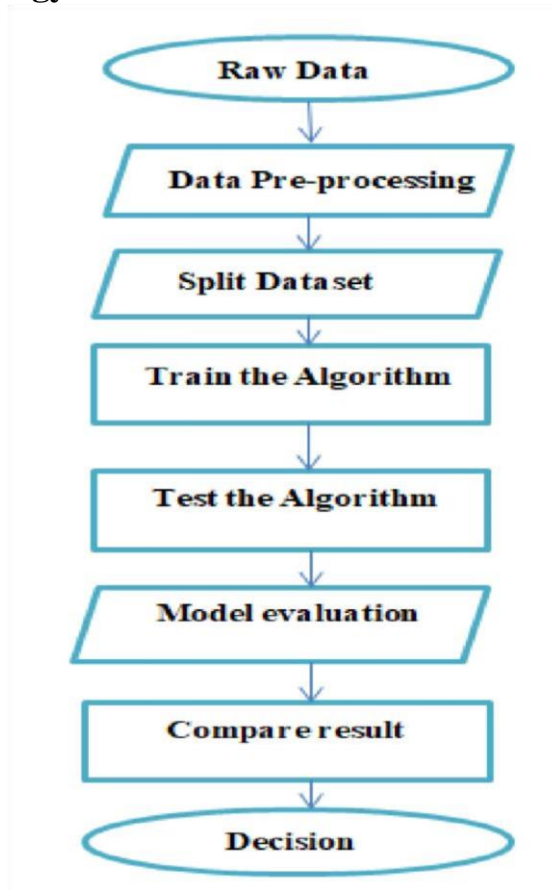| ইউনিলিভার	ক্লিয়ার শ্যাম্পু মেন কুল স্পোর্ট মেনথল এন্টিড্রান্ড্রফ	1.মানসম্মত। 2.অথেনটিক 😄 3.ভালো পণ্য। 4.অরিজিনাল প্রোডাক্ট। ভালো লাগছে। 6.অসাধারন 😃 7.সত্যিই ভালো পণ্য। 8.প্রোডাক্ট ভালো ছিল। 9.অরিজিনাল প্রোডাক্ট হাতে পেয়েছি। 10.ভালো পণ্য।
	ক্লিয়ার শ্যাম্পু কম্পলিট একন্টিড কেয়ার এন্টি ড্যানড্রফ	1.একদম অরিজিনাল। 2.খুবই ভাল প্রোডাক্ট 👍 3.অসাধারণ। 4.অরিজিনাল। 5.ভালো ছিল। 6.ভালো। 7.সন্তুষ্ট। 8.ভালো না। 9.একদম নিম্নমানের পণ্য। 10.রেকমেন্ডেড।

.

3.5 Research Methodology of Process

Figure 3.5.1 This diagram is describing the whole procedure of our work

3.6 Word Corpus

We have divided the review into two categories stored in word corpus; Positive and negative. Expressions of different reviews for analysis are also stored in the database.

Table 3.6.1 Comprises the listing of the categorization made with different lexicons.

Table 3.6.1 Lexicon Categories

হ্যাঁ বোধক	১. পণ্যটি ভালো।

	২.আসাধারণ পণ্য। ৩.অরিজিনাল প্রোডাক্ট
না বোধক	১.পণ্যটি ভালো না। ২.এর মান ভালো না। ৩. একদম ভালো লাগেনি

3.7 Implementation Requirement

When we started to implement this research, it was very new to us. That was hard to cover up this challenge. To implement the research, we had to maintain a big survey and need a computer to implement the code. Besides, it requires some soft skills.

       a.   Multinomial Naive Bayes

       b.   Random forest

       c.   Decision tree

       d.   Support Vector Machine (SVM)

CHAPTER 4

EXPERIMENTAL RESULT AND DISCUSSION

4.1 Introduction

In this chapter, the results of sentiment analysis of Bengali comments of product review are presented. The result of this research we have carried out is totally experimental based. We have collected around 400 comments of product reviews from websites like daraz, shajgoj, stylin. The experiment has been carried by using Multinomial Naive Bayes, Random Forest, Decision tree and SVM method. The comments were identified as positive or negative and the results have been discussed with our experiment in detail. Then the results have finally come out with a satisfactory conclusion.

4.2 Experimental Results & Analysis

The reviews that we accumulated for analysis purposes have been stored as two kinds of review in our corpus and our corpus; we have a total 400 reviews present in our corpus which are shown in the table 4.2.1 in detail.

Table 4.2.1 Review Distribution

হ্যাঁ বোধক	না বোধক
P	N
318	82

Recalling the research question part, we have started our research with four questions. For getting the answer to our research questions, we have applied multinomial naive bayes, random forest, decision tree and SVM. After analysis reviews system work is to say the product is good or bad. We have performed 4 algorithms for each brand of products. The result after performing the review has been picturised of predicted popular cosmetic product of assembled Bengali reviews.   To finalize

the total accuracy of our experiment, we sorted out individual accuracy for each algorithm. After getting the individual accuracy then we can say which cosmetic product is more popular than others.

Accuracy= (TP+TN)/(TP+FP+TN+FN)

Here, TP= True Positive

TN= True Negative

FP= False Positive

FN= False Negative

At the very beginning of our research we faced some difficulties about which product reviews are positive and which are negative. Reviews are like:

অস্থির প্রোডাক্ট
চরম পণ্য
দেখে যতটা ভালো মনে করছিলাম, ততটা ভালো না
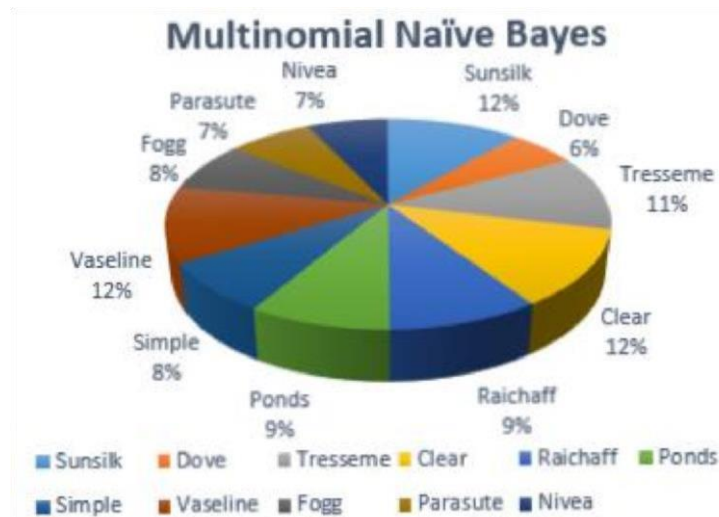
।

The accuracy of Multinomial Naive Bayes:

Figure 4.2.1: Accuracy of Multinomial Naive BayesHere, Sunsilk= 12%, Dove= 6%, Tresemme=11%, Clear= 12%, Rainchaff= 9%, Ponds= 9%, Simple= 8%, Vaseline= 12%, Fogg= 8%, Parachute= 7%, Nivea= 7%
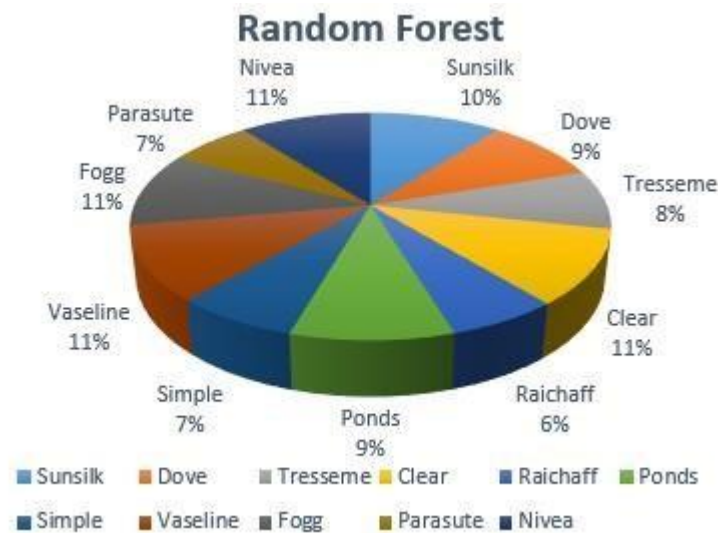
The Accuracy of Random Forest:

Figure 4.2.2 Accuracy of Random Forest

Here, Sunsilk= 10%, Dove= 9%, Tresemme=8%, Clear= 11%, Rainchaff= 6%, Ponds= 9%, Simple= 7%, Vaseline= 11%, Fogg= 11%, Parachute= 7%, Nivea= 11%
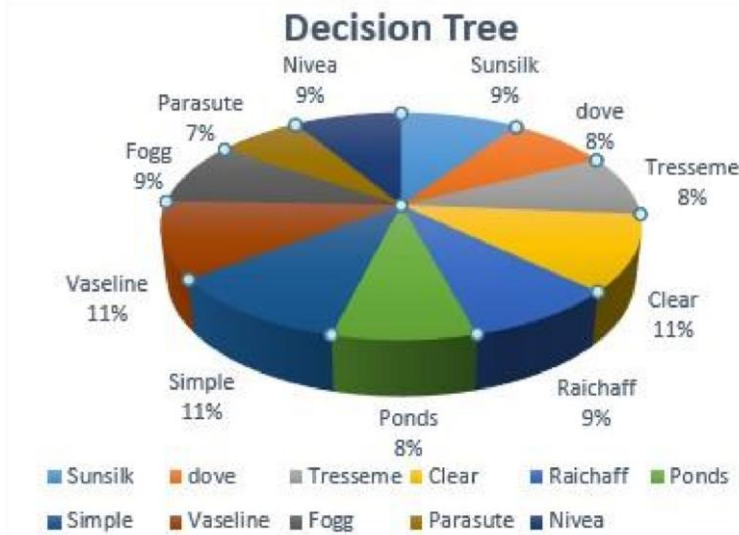
The Accuracy of Decision Tree:

Figure of 4.2.3 Accuracy of Decision Tree

Here, Sunsilk= 9%, Dove= 8%, Tresemme=8%, Clear= 11%, Rainchaff= 9%, Ponds= 8%, Simple= 11%, Vaseline= 11%, Fogg= 9%, Parachute= 7%, Nivea= 9%

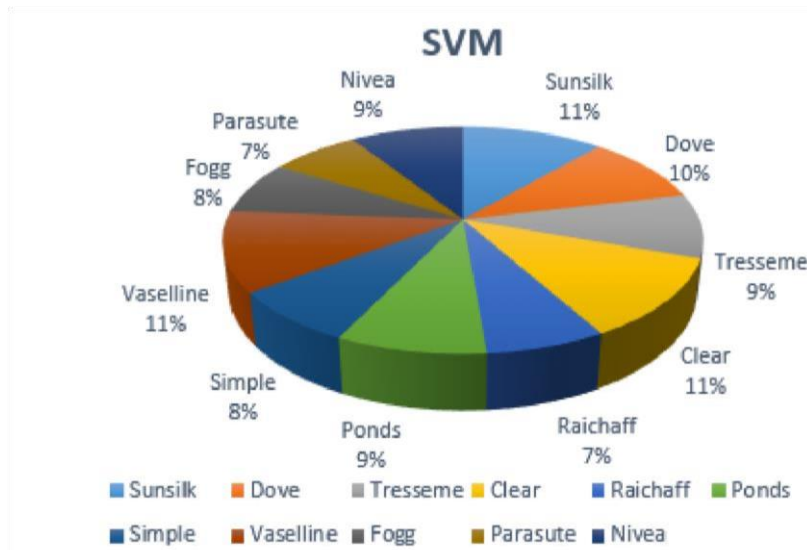## The Accuracy of Support Vector Machine (SVM):

Figure of 4.2.4 Accuracy of SVM

Here, Sunsilk= 11%, Dove= 10%, Tresemme= 9%, Clear= 11%, Rainchaff= 7%, Ponds= 9%, Simple= 8%, Vaseline= 11%, Fogg= 8%, Parachute= 7%, Nivea= 9%

Then we have gone through a small survey which includes a questionnaire having 10 complicated controversy Bengali Reviews. Our main purpose was to know if the product is positive or negative.

At the end, we come up with almost the same probability.

In this experimental study to see how bangla reviews are affecting people to buy their cosmetic products. To solve this problem we have applied four algorithms and at the end we look at the accuracy of those products to determine which cosmetic product is better.

Table 4.2.2 Accuracy Table

Brand	Multinomial Naive Bayes	Random Forest	Decision tree	SVM
Sunsilk	100.0%	100.0%	87.5%	100.0%
Dove	50.0%	87.5%	75.0%	87.5%
Tresemme	100.0%	80.0%	80.0%	80.0%
Clear	100.0%	100.0%	100.0%	100.0%
Raichaff	80.0%	60.0%	80.0%	60.0%
Ponds	75.0%	87.5%	75.0%	75.0%
Simple	66.66%	66.67%	100.0%	66.66%
Vaseline	100.0%	100.0%	100.0%	100.0%
Fogg	66.66%	100.0%	83.3%	66.66%
Parachute	62.5%	62.5%	62.5%	62.5%
Naivea	60.0%	100.0%	80.0%	80.0%

4.3 Discussion

This chapter shows the clear result of our research question. After experimenting we have found that a review is positive or negative. To do so we have used four algorithms. After that the experimental result we have come up with the answer of our research question with the accuracy of the reviews. Therefore the entire accuracy refers to the more popular cosmetic products than other products.

CHAPTER 5

IMPACT ON SOCIETY

5.1 Impact on society

Customer reviews of online products are very important for a company or a business because when buyers review a product online, people from all over the country and the world will see it. Two things can happen to that trader or company. The first is that if the review is negative, the sales rate of that trader or company may go down and many people who work in that organization may lose their jobs. As a result, many people will lose their jobs and become unemployed and may get involved in various anti-social activities, which is not desirable for our society at all. The second is that if the customer's review is positive, then the business of that trader or that company will be profitable and their business will need more manpower and many people will be able to work there and eliminate their unemployment. Unemployment is a major obstacle in our society today and if we overcome it, it will be a significant achievement in our society. Finally, I would like to say that online based product reviews of customers can have a huge impact on our society.

5.2 Ethical Aspects

Just as every work has some ethical aspects, so people should comply with ethical compliance when it comes to online product reviews. For example, other companies are exploiting the negative reviews of the products of different big companies to satisfy their interests, which is outside the law of business. Nowadays technology becomes more powerful and it can be used for nefarious reasons as well as good, so cyber security should be given more priority. Therefore everyone should adhere to the ethical aspects.

CHAPTER 6

CONCLUSION

6.1 Summary of the Study

In this research, we tried to predict popular cosmetics products from Bengali reviews using sentiment analysis. Classifying the reviews into either positive or negative using the random forest, decision tree, and then predicting popular cosmetic products and rating those products based on the reviews was our main goal. It's a lexicon-based approach where the positive and negative product reviews are stored in the database as a unique hash value. For each the review will be generated

whether it's a positive one or a negative one. Then rating will be detected. Our trained data and our test data both showed our expected results. After this experimental result finally, we have come up with a decision on which popular cosmetic brand is well-liked to people.

6.2 Conclusion

Predicting popular cosmetics from Bengali reviews was not that easy as Bengali people, in particular, are not interested in expressing their reviews on the cosmetic product. Our method helps us to predict the exact popular cosmetic brand that the majority of people think about.

Among different approaches we have used that helps us to gain almost 100% accuracy.

6.3 Implications for Further Research

Nowadays, the demand for data mining analysts is highly appreciated. This is due to the presence of abundant amounts of data in our surroundings. To be more precise, it is the right time to work with these sorts of complex data, so that a new pattern can be introduced to resolve different critical problems. Sentimental analysis is one of the fundamental branches of data mining. The experimental study which we have carried out on emotion detection with a satisfactory outcome is leaving a strong footprint behind our work. It has been observed that works on sentiment analysis in English have lots of valuable impacts in everyday life. As we are food lovers and besides in this modernized world, people are seen very active in social media like Facebook, it has been observed that people post Bengali food reviews in groups like foodbank, food bloggers etc. So it can be a very good start

of mining different Bengali comments or reviews more on other products. Research can be done on how our project can be utilized to build something for disables those who are especially blind.

6.4 Recommendations

As at present, our corpus doesn't have neutral lexicons. So generating accuracy from given input Bengali comments will be failed if it does not contain specific lexicons. So before going for a test add the necessary keywords to the database. While giving input keep the focus on the spelling of the lexicons. In case of a spelling mistake, the program will fail to generate the right expressions.

So users may get the wrong answer.

6.5 Future Scope

Some of our work we have done but some of them we have planned for our future work to develop this system more understandable and reliable. To achieve the goal we have to do some work and those are-

a. We have done our preliminary experiment with a small amount of data. We will work with a large amount of data to gain the result more accurately.

b. We have worked on only two results, either positive or negative. In future we will work on neutral results also.

c. We also want to work on the basic emotions (e.g, happy, sad, angry, disgust) of a user to make reviews more specific.

References

[1] Xing Fang* and Justin Zhan, Sentiment analysis using product review data, Fang and Zhan Journal of Big Data (2015) 2:5

[2] Ankur Goel, Jyoti Gautam, Sitesh Kumar, Real Time Sentiment Analysis of Tweets Using Naive Bayes, 2016 2nd International Conference on Next Generation Computing Technologies (NGCT-2016)

[3] Rudy Prabowo 1, Mike Thelwall , Sentiment analysis: A combined approach, Journal of Informetrics 3 (2009) 143–157

[4] Doaa Mohey El-Din Mohamed Hussein,A survey on sentiment analysis challenges, Journal of king university- Engineering science(2016)XXX, xxx-xxx

[5] Apoorv Agarwal, Boyi Xie, Ilia Vovsha, Owen Rambow, Rebecca Passonneau, Sentiment Analysis of Twitter Data

[6] Rafeeque Pandarachalil, Selvaraju Sendhilkumar, G. S. Mahalakshmi, Twitter Sentiment Analysis for Large-Scale Data: An Unsupervised Approach, Cogn Computer

[7] Rajesh Bose, Raktim Kumar Dey, Sandip Roy and Debabrata Sarddar, Sentiment Analysis on Online Product Reviews

Paper