

E-Commerce in the Next Era of Developing Bangladesh

Consumer's Perception and Challenges

BY

KAMRUL ISLAM SAKIB
ID: 213-15-4491

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

Supervised By

ABDUS SATTAR
Assistant Professor
Department of Computer Science and Engineering
Daffodil International University



DAFFODIL INTERNATIONAL UNIVERSITY
DHAKA, BANGLADESH
July 2024

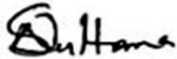
APPROVAL

This Project titled “E-commerce in The Next Era of Developing Bangladesh Consumer's Perception and Challenges”, submitted by **Kamrul Islam Sakib** 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 16-07-2024.

BOARD OF EXAMINERS

Narayan Ranjan Chakraborty (NRC)
Associate Professor & Associate Head
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Board Chairman



Naznin Sultana (NS)
Associate Professor
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner 1



Raja Tariqul Hasan Tusher (THT)
Assistant Professor
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner 2



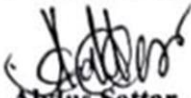
Dr. Ahmed Wasif Reza (DWR)
Professor
Department of Computer Science and Engineering
East West University

External Examiner

DECLARATION

I hereby declare that this project has been done by me under the supervision of **Abdus Sattar, Assistant Professor, Department of Computer Science and Engineering, Daffodil International University**. I also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:



Abdus Sattar
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

Kamrul Islam Sakib

Kamrul Islam Sakib
ID: -213-15-4491
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, I 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 **Abdus Sattar, Assistant Professor**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of “*Machine learning*” 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 **Head of the Department of CSE**, for his kind help in finishing our project and also to other faculty members and the staff of the Department of CSE, 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 my parents.

ABSTRACT

There is a great potential for the growth of e-commerce in Bangladesh due to the growth in internet connectivity and usage. Knowledge and solving the issues related to consumer concern are necessary to maximize the benefit from this sector. Thus, this research explores customer sentiment and issues concerning e-commerce businesses in Bangladesh by employing a Google Form-based survey with 599 entries and 16 features. Demographic information, web use history and target characteristic such as ‘Good’ and ‘Bad’ satisfaction with the service are incorporated into the dataset. For the given problem to predict consumer perception, machine learning models such as GB, RF, BNB, SVC, DT were applied. It was fulfilled by the Random-Forest classifier with high accuracy of 99.16%, surpassing other algorithms. This speaks to the model’s ability to detect subtleties in terms of time-series data and forecast consumer satisfaction accurately under the conditions of e-commerce in Bangladesh. The gaps linked to several constituents of consumer satisfaction, including the product quality, delivery efficacy, and the methods of payment, were also established via the analysis. Essential plan points to consider when seeking to boost the position of e-commerce ventures are; logistics and delivery, sustainability, and ethical factors. The rights of consumers and fair competition policies are promoted by encouraging policymakers to put into place relevant regulations. This present research adds knowledge to the subject of e-commerce development in emerging economies such as Bangladesh, thus providing information that may be useful in designing the proper policies that may help steady development of this business model and improving consumer confidence in electronic commerce transactions.

TABLE OF CONTENTS

CONTENTS	PAGE
Board of examiners	i
Declaration	ii
Acknowledgements	iii
Abstract	iv
Table of contents	v-vii
List of Figures	viii
List of Tables	ix
CHAPTER	1-9
CHAPTER 1: INTRODUCTION	
1.1 Introduction	1
1.2 Motivation	6
1.3 Rationale of the Study	7
1.4 Research Questions	7
1.5 Expected Output	8
1.6 Project Management and Finance	8
1.7 Report layout	9
CHAPTER 2: BACKGROUND	10-16
2.1 Preliminaries/Terminologies	11
2.2 Related Works	11-15

2.3 Comparative Analysis and Summary	15
2.4 Scope of the Problem	16
2.5 Challenges	16
CHAPTER 3: RESEARCH METHODOLOGY	17-22
3.1 Research Subject and Instrumentation	19
3.2 Data Collection Procedure/Dataset Utilized	19
3.3 Statistical Analysis	21
3.4 Proposed Methodology/Applied Mechanism	21
3.5 Implementation Requirements	22
CHAPTER 4: EXPERIMENTAL RESULT	23-36
4.1 Experimental Setup	23
4.2 Experimental Results & Analysis	23-36
4.3 Discussion	36
CHAPTER 5: IMPACT ON SOCIETY	37-38
5.1 Impact on Society	37
5.2 Impact on Environment	37
5.3 Ethical Aspects	38
5.4 Sustainability Plan	38
CHAPTER 6: SUMMARY, CONCLUSION, RECOMMENDATION AND IMPLICATION FOR FUTURE RESEARCH	39-40
6.1 Summary of the Study	39
6.2 Conclusions	39

6.3 Implication for Further Study	40
REFERENCES	41-42
APPENDIX	A1-A6

LIST OF FIGURES

FIGURES	Page No
Figure 1.1: BNB Architecture D. Widyawati et al. 2023 [21]	3
Figure 1.2: GB Architecture [22]	3
Figure 1.3: DT Architecture Mishra, et al. 2021 [23]	4
Figure 1.4: RF Architecture Chen et al.2019 [24]	5
Figure 1.5: SVC Architecture [25]	6
Figure 3.1: Percentages of datasets	20
Figure 3.2: Work Flow	23
Figure 4.1: Confusion Matrix (GB)	27
Figure 4.2: ROC Curve of GB	28
Figure 4.3: Confusion Matrix (BNB)	29
Figure 4.4: ROC Curve of BNB	30
Figure 4.5: Confusion Matrix (DT)	31
Figure 4.6: ROC Curve of DT	32
Figure 4.7: Confusion Matrix (RF)	33
Figure 4.8: ROC Curve of (RF)	34
Figure 4.9: Confusion Matrix (SVC)	35
Figure 4.10: ROC Curve of (SVC)	36
Figure 4.11: Comparative Model Accuracy Bar Plot	37

LIST OF TABLES

TABLES	Page No
TABLE 1.6: PROJECT MANAGEMENT TABLE	9
TABLE 2.3: COMPARATIVE ANALYSIS OF PREVIOUS WORK	17
TABLE 4.1: PERFORMANCE EVALUATION	38

CHAPTER 1

Introduction

1.1 Introduction

E-commerce in Bangladesh has grown very fast due to more internet usage, mobile phone adoption, and liberal government policies. Accordingly, as the country moves towards the realization of a digital economy, it is vital to grasp consumers' perceptions and the difficulties within the e-commerce domain. It will center on these perceptions and difficulties in order to offer the necessary findings that will improve the e-commerce experience of the stakeholders. Mahmud, First A. Masfiq, et al. [1] The data obtained for this study was collected from an online survey administered via Google Form link, and it received 599 consumers' responses that include demographics. That is 16 factors that characterize a client, for instance, age, gender, residence, internet access, preferred payment method, and e-commerce experience. The target variable classifies consumers' perceptions about a brand as either positive or negative, signified by "Good" or "Bad" respectively. Uddin, Faisal, et al. [2] The methodology involves several key steps: it includes data selection or data pre-processing in which unwanted data points are removed from the data set, feature selection in which we choose the most relevant features for the model, encoding of categorical data, exploratory data analysis (EDA), model training, model evaluation, and finally testing the model. To analyze the data and make a forecast in regards to consumer perception, methods like Gradient Boosting Classifier, Random Forest Classifier, BernoulliNB, SVC, and Decision Tree Classifier are used X. Lin, et al. [6] Regarding its objectives, this study has planned to eliminate data quality problems systematically and use appropriate analytical methods and tools to examine pivotal factors affecting the consumer satisfaction and discover the directions that e-commerce needs to pay attention to. The outcomes of the study should be beneficial for the policy makers, entrepreneurs, and other related stakeholders who are aiming at improving the consumers' perception of online shopping and increasing the level of their engagement in e-shopping in Bangladesh.

Machine Learning Model

They are very useful in predicting behaviors and evaluations such as consumers in e-commerce related organizations or situations like Bangladesh. Implemented techniques such as Gradient Boosting Classifier help enhance the ways of getting more accurate predictors as they develop multiple models operatively in sequence on a set of data. Random Forest Classifier explains decision by averaging the results of many decision trees, it is less likely to overfit the data, as well as handle other complex patterns in the input data. BernoulliNB is used when the treatment of the covariates is to assume that features are independent of each other and it is suitable for binary classification. SVC separates data points in a way that makes them easier to classify with proper distribution when data held in high dimensions. Unlike Decision Tree Classifier, it splits data based on the feature attributes, and illustrates decision making procedures evidently. Every single algorithm has potential to cater different angles of consumers' behavior and satisfaction level so that businesses and policymakers can extract practical knowledge for improving e-commerce and sustainable growth of other emerging economy like Bangladesh.

BNB

BNB is a class in Naïve Bayes family which is essentially used in the binary classification tasks where features are binary in nature. Overlooking the fact that all the features are assumed to be mutually exclusive; that is, the occurrence of one feature does not affect the others. This algorithm is vital due to its ease in training and analysis, speed, and capacity in dealing with huge data sets with rich features. More generally for e-Commerce BernoulliNB can be useful in cases like sentiment analysis or other problems with binary output such as the prediction of customer satisfaction (positive/negative)

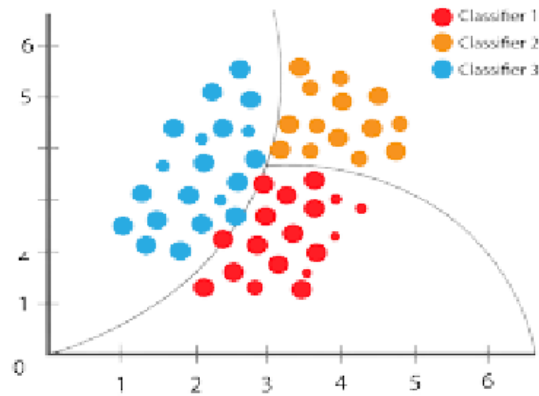


Figure 1.1: BNB Architecture D. Widyawati et al. 2023 [21]

GB

GB is a type of machine learning algorithm in which the procedure of modeling progresses step by step and in each step, the new model is learned using some of the residuals left by the previous model. It takes a number of weak learners, which are usually decision trees, and forms a strong learning model. This algorithm is quite useful, it is generally efficient in dealing with various relation in data and it can guarantee high predictive rate most times. Given its sequential structure, it enhances model performance incrementally and since it is appropriate for the classification problem in e-commerce it can be applied for various classification tasks, ranging from consumer behavior analysis to the patterns in customers' preferences.

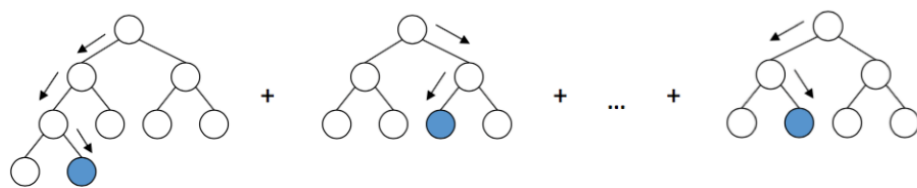


Figure 1.2: GB Architecture [22]

Decision Tree

Decision Tree Classifier constructs a model to estimate the value of a target variable using the decision-making logic that is derived from the dataset's characteristics. This 'test' on an attribute is modeled by each internal node, while each branch depicts the outcome of the test, and the class label is represented by the each of the 'leaf node'.

Indeed, one of the advantages of this algorithm is that it gives explainable results, which means users can comprehend why they arrived at a particular decision. In the realm of e-commerce, Decision Tree Classifier can be of use when determining the use of products recommendations and/or customer segmentation dependent on specific features such as the purchasing history and/or activity history.

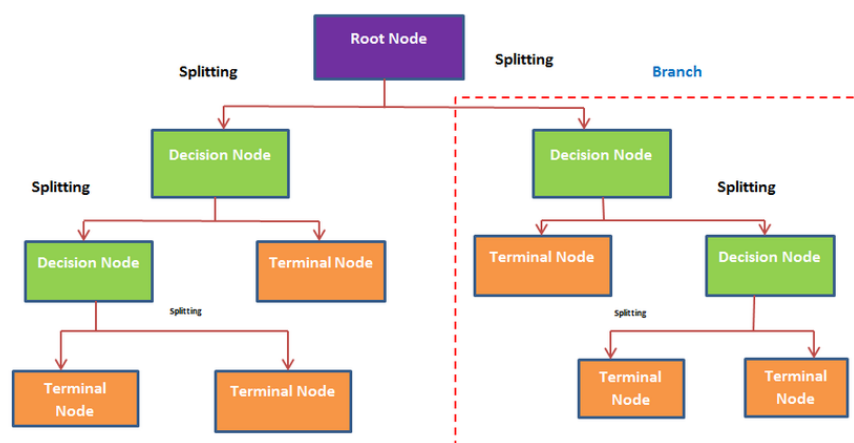


Figure 1.3: DT Architecture Mishra, et al. 2021 [23]

Random Forest

Another method of the ensemble learning is RF, which builds a number of decision trees during the training and returns the mode of the classes (in classification tasks) or the average of all the trees' predictions (in regression tasks). Every tree in the forest works in isolation, and it comes up with its own decision without considering the characteristics of the data. This algorithm is crucial because it eliminates the overfitting problem through the building of multiple decision trees, it is robust and more generalized. When used in e-commerce, RF works well with a large number of data samples and large numbers of features, which would enable it to predict the consumers' actions given various characteristics.

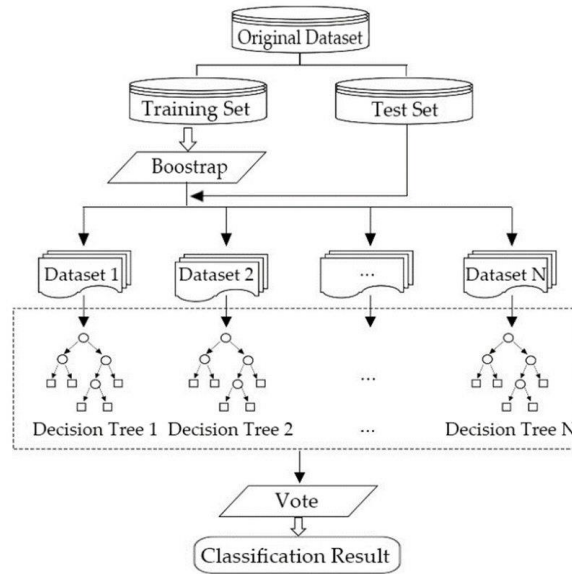


Figure 1.4: RF Architecture Chen et al.2019 [24]

SVC

SVC or Support Vector Classifier is a types of supervised learning algorithm used specifically for classification. It defines a hyperplane in the N dimensional space that best separates the data points. SVC is a significant classifier since it is applicable in the high-dimensional space and capable of dealing with non-linear separable data by the utilization of kernels. In e-commerce SVC can be useful in any situation when it is possible to divide objects into distinct groups, for example during the customer segmentation or during the identification of a fraud, when the goal is to distinguish between two clearly separated classes, for example between real and fake purchases.

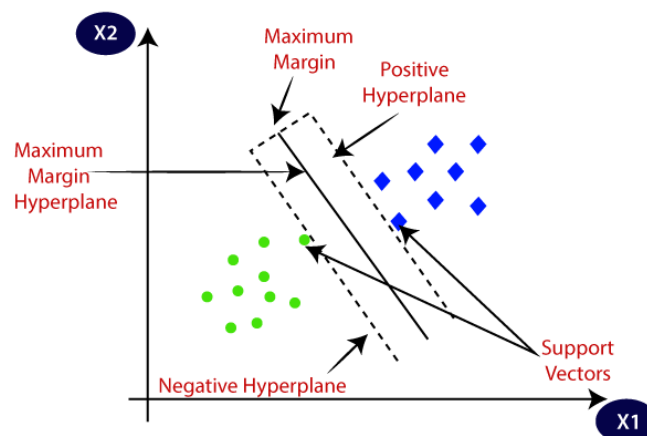


Figure 1.5: SVC Architecture [25]

1.2 Motivation

This paper aims at examining the impact of the speedy expansion of e-commerce within Bangladesh context as a boon and bane for the local players and customers. However, many challenges related to the sector, including trust, product quality and delivery logistics affect consumers' satisfaction even to this date. Recognizing these issues from the consumer's standpoint is crucial in identifying means of improving his e-shopping experience. This research is informed by the necessity to establish factors that may affect consumer attitude and to explicate the challenges that affect sector in its developmental process. Hence collecting first party data directly from consumers, it is a possibility to come up with relevant conclusions and recommendations for the business to enhance its services, gain trust and reliability of the consumers and, consequently, contribute to the sustainable development of the e-commerce sector in Bangladesh. With the insights it will also assist policy makers in designing legislation and programs that help to enhance the strength of the digital economy.

1.3 Rational of the Study

Bangladesh is a country that holds high potential for growth in the e-commerce sector, but the e-commerce sector is still struggling with common pitfalls that tarnish the level of trust and satisfaction in the shoppers. It is necessary to comprehend consumers' perception to identify such issues and design potential solutions for them. The objective of this research is to systematically assess all drivers that relate to the level of consumer satisfaction and all barriers' consumers face during online shopping. The research hence provides insights that are endogenous to the sector hence providing a true picture on the existing problems in this sector. The know-how, on balance, can increase consumer confidence as well as participation that directly nurtures the quality of services offered in e-commerce businesses. Third, the results presented here can inform policymakers and legislators in the development of supportive regulations that reflect the existing hurdles and encourage the advancement of the digital business sector. Conclusively, the overall aim of the study is to build a conceptual model for Bangladesh that could assist in the sustainable development of e-commerce for the consumers as well as the businesses.

1.4 Research Questions

1. What are the important antecedents of consumer satisfaction regarding e-commerce in Bangladesh?
2. Which demographic factors influence consumers' perception of e-commerce; age, gender, geographical location?
3. What is the perception that consumers hold in relation to the quality of products and services offered by e-commerce Bangladesh?
4. What changes should be made according to the consumers in order to experience better e-shopping in Bangladesh?

1.5 Expected Output

The decision made in the study will provide the necessary shift towards the core areas that serve a substantial impact on the consumer satisfaction and perception of the e-commerce sector in Bangladesh. It will highlight the effectiveness or otherwise of such perceptions given the demographic variables of age, gender and location. The study will also point out more frequent problems associated with consumer experience, including trust, product quality, and delivery. Findings related to the payment options of choice as well as the impact of the payment options on the consumer will be discussed. Furthermore, the research will identify the internet usage, including the uptime and the online shopping habits. The kinds and types of e-commerce services perceived by the consumers to be fit for the product and service quality they deliver will be described. At the end, recommendations to enhance enticing and compelling e-commerce potencies for business structures and policymakers in Bangladesh will be discussed.

1.6 Project Management and Finance

The project will be implemented with the participation of professionals with focuses on data analysis, e-commerce, and project management, which would guarantee the successful flow of the process. Some of the significant steps to achieve the objectives include data acquisition and preparation, data exploration and transformation, data analysis and synthesis, and the result's presentation. To mitigate such situations, daily meetings and progress reviews will be conducted so that any cases that may arise will be handled ASAP. The financial budget involving the project includes Google Form,

which is the data collection instrument, data analysis instruments, Python, and human resource expenses. Further resources go to stakeholder meetings, reports, and presenting the findings to the stakeholders and other parties interested in the results. Another sub control entails effective management of the project resources and allocated finances in order to operate within budget and meet the projects goals.

Table 1.6: Project Management Table

Work	Time
Data Collection	2 months
Literature Review	1 months
Experimental Setup	1 month
Implementation	2 months
Report Writing	1 months
Total	7 months

1.7 Report layout

Chapter 1: Introduction.

This chapter defines the area of the study where e-commerce has rapidly grown in Bangladesh and thus the need to assess consumers' attitude and barriers towards e-commerce. In this part, the writer explains the purpose of the study or reason for carrying out the research and the background to the study.

Chapter 2: Background.

This chapter includes a brief discussion of prior studies and research on e-commerce, consumer behavior, including the research context of Bangladesh. It defines the objectives, assumptions, and scope of the study and explains the context in which e-commerce is currently developed in the country, as well as significant theories, prior empirical evidence, and theoretical frameworks.

Chapter 3: Research Methodology

The chapter also outlines the specifics of data selection, and data preprocessing, which includes methods of missing value imputation and encoding, data visualization, model training, and validation. It provides the justification of the many approaches used in the selection of the algorithms and analytical methods applied.

Chapter 4: Experimental Results

This chapter describes the data analysis and the assessment of the findings from the models. The literature review encompasses numerous machine learning algorithms, consumer experience or perception and barriers within, and statistical analysis of differentiation factors influencing e-commerce satisfaction in Bangladesh.

Chapter 5: The Impact on Society

This chapter presents the implications of the current study in relation to society and the culture at large. It discusses how the continuity of e-commerce can improve customer satisfaction, revive the economy, and solve digital inequality problems in Bangladesh. Looking at the potential benefits that could be derived from such an intervention strategy, the following are observed:

Chapter Six: Summary, Conclusion, Recommendations, and Implications for Future Research

In this chapter, the author identifies and discusses general conclusions that can be made regarding the studied issue, based on the obtained outcomes. It provides suggestions to e-commerce firms and authorities on how the sector can be enhanced. Besides, it highlights the limitation of the current study and recommends the scope of future research to give more light to e-commerce in Bangladesh.

CHAPTER 2

Background

2.1 Preliminaries/Terminologies

This section just presents definitions and preconditions that are essential to set the scene and define the subject matter of the study. As a part of the study, the research work covers an understanding of the context of e-commerce in Bangladesh, its growth pattern and the influence on the society. Some important terminology and concept dealing with consumer behavior, digital economy and ecommerce and shopping behavior are defined to create better perspective. Also, the goals of the research are defined within this section: the main focus is on investigating consumer attitudes and issues in the e-commerce field. It also provides a general overview of the methods used for the research, i.e. data collection and analytical tools to give a general outline of the research framework followed in the study. For this reason, the preliminaries section allows readers to follow further discussions in the subsequent chapters by engrossing them with the details that create a basis for the study and the knowledge of the results and their implications.

2.2 Related Works

The existing research works done in Bangladesh regarding e-commerce have addressed different issues including the market analysis, customers' perception, and overall technology aspect. Scientific papers pointed out that with the help of digital platforms, radical changes in traditional retailing and market coverage are possible. These also highlight some issues and prospects including logistic systems, payment methods, and consumers' trust that defines the scope of e-commerce in Bangladesh. Literature that was reviewed in the current research is informative in revealing the effects of various factors on consumers and the various ways through which the e-commerce experiences can be improved in the country.

Mahmud, First A. Masfiq, et al. [1] analyzed the sentiment of women's clothing reviews on ecommerce platforms using Natural Language Processing (NLP) techniques and machine learning algorithms. The dataset was analyzed using a variety of analytical approaches, such as word clouds, data summaries, topic modeling, and sentiment

distribution visualizations. The dataset contains a variety of features, including product details, seller ratings, pricing information, and sentiment labels. Multiple machine learning algorithms, such as Random Forest, Support Vector Machine (SVM), and Logistic Regression, were used to predict sentiment. The Random Forest Classifier performed best, with an impressive accuracy, precision, recall, and F1-score of 96.51%. Overall, the study clarifies the relationship between product attributes and customer sentiments, providing useful insights for e-commerce platforms and merchants looking to improve their products and services.

Uddin, Faisal, et al. [2] analyzed the distinct growth patterns and challenges confronting Bangladesh's e-commerce sector. It emphasizes the growing importance of subscription models within a service-oriented e-commerce framework. Furthermore, the study provides a comprehensive business model designed specifically for e-commerce platforms operating in emerging markets, outlining growth stages, strategic approaches, and adaptation strategies. By analyzing existing e-commerce platforms using this proposed model, the study hopes to provide insights into optimizing operations and navigating challenges in emerging markets such as Bangladesh.

M. Abdul Wadud Sorker, et al. [3] presented a client categorization concept based on online shopping behavior, distinguishing platinum, gold, and silver groups. Data was gathered using a Google Form, which included age limits and key professions for more comprehensive insights. The data from 531 respondents was processed using machine learning techniques such as Random Forest Classifier (RFC), Nave Bayes Classifier (NBC), and Logistic Regression (LR). The Random Forest Classifier (RFC) method was the most effective, with an impressive 94.34% accuracy in classifying online consumers based on their purchasing patterns and behaviors.

H. Jahan et al. [4] addressed sentiment analysis of consumer comments on Daraz and Evaly's Facebook pages, with a focus on e-commerce service perceptions. Data were collected from Facebook comments, which reflect consumer opinions and experiences. Various machine learning models, such as logistic regression, decision tree, and random forest, were used with different n-gram features. Random forest achieved the highest accuracy of 90.65% for unigrams and 89.93% for trigrams, while decision trees achieved 88.49% for bigrams. Notably, the combination of unigram features and the random forest model performed the best in sentiment analysis.

M. O. Aftab, U. Ahmad, S. et al.[5] addressed digital marketing challenges by proposing a predictive approach for assessing consumer insights and product purchase

intentions. The methodology uses two models: a customized Recurrent Neural Network (RNN) for analyzing user perceptions on social media, and classification algorithms such as Random Forest, Naïve Bayes, and BayesNet to predict customer purchase intent. The approach not only analyzes sentiment in customer feedback on social media, but it also predicts their likelihood of purchasing a product. The RNN model has a training accuracy of 97.15% and a validation precision of 94.25%, whereas the Random Forest method has a training accuracy of 97% and validation precision of 94.97%. This demonstrates the effectiveness of the proposed method for analyzing consumer perceptions and predicting purchase behavior.

X. Lin, et al. [6] used the Women's E-Commerce Clothing Reviews dataset to perform sentiment analysis on customer recommendations. Five machine learning algorithms, including Logistic Regression, SVM, Random Forest, XGBoost, and LightGBM, were used to investigate the relationship between review features and product recommendations using NLP techniques. The LightGBM algorithm performed the best, with the highest AUC value, accuracy, precision, recall, and F1 score all at 0.97. Ridge Regression, Linear Kernel SVM, and XGBoost algorithms all performed competitively, with an accuracy of 0.94. Overall, this study provides valuable insights into customer sentiment and behavior in the e-commerce industry, which can help us better understand customer psychology and improve transaction outcomes.

J. Biswas, M. M. Rahman, et al. [7] presented a novel approach to predicting users' sentiments based on their Facebook comments about online food delivery services, primarily using the BERT machine learning technique. In addition, Char-CNN, Graph-CNN, LSTM, and Bi-LSTM techniques were considered for performance comparison against BERT. The study assesses the effectiveness of these techniques using a variety of performance evaluation metrics. Results show that BERT outperforms the other techniques, with an accuracy of 92.86%. This suggests that BERT is very effective at sentiment analysis tasks involving Facebook comments about online food delivery services.

A. K. Bitto, M. H. I. Bijoy et al. [8] analyzed consumer comments from the Facebook pages of Food Panda, HungryNaki, Pathao Food, and Shohoz Food to determine their sentiment toward food delivery systems. As part of the natural language processing (NLP) task, rigorous data preprocessing was performed, such as contractions, stop word removal, and tokenization. Four supervised classification methods and three deep learning models were used. The extreme gradient boosting (XGB) model achieved the

highest accuracy among machine learning (ML) algorithms, at 89.64%. Meanwhile, the long-term short-term memory (LSTM) deep learning model had the highest accuracy of 91.07%, outperforming both ML and other DL models.

H. Aulawi, E. Karundeng, et al. [9] analyzed e-commerce challenges in the context of the Covid-19 pandemic, with the goal of making recommendations for improving e-commerce services. We used Naïve Bayes Classifier (NBC), Text Association, and Focus Group Discussion (FGD) methods. NBC was used for sentiment classification, Text Association for revealing word relationships, and FGD for analyzing Text Association results and developing improvement strategies. The data came from consumer reviews on Twitter between January and April 2021. The FGD process involved e-commerce experts categorizing Text Association findings, identifying root causes, and proposing improvement measures. Recommendations include improving product specification accuracy, improving logistics partnerships, refining promotional event disclosure, streamlining refund procedures, optimizing multi-channel ecommerce platforms, improving seller data verification, refining transaction monitoring, and leveraging consumer feedback for reward and penalty systems.

D. S. Ashari, B. Irawan et al.[10] analyzed sentiment analysis of Instagram comments on online transportation services, with the goal of determining which service has the most positive sentiment. The system analyzes sentiments using Convolutional Neural Network (CNN) methodology to help consumers select the best service. The Instagram comments are used as the dataset for sentiment classification. The system's performance resulted in an impressive average accuracy of 94% when classifying sentiments in the Instagram comments column. Overall, this study provides a useful tool for consumers to make informed decisions about online transportation services based on sentiment analysis. To predict product sales,

H. Zhu,et al.[11] proposed a hybrid network that combines Bi-directional Long Short-Term Memory (BiLSTM) and Convolutional Neural Network (CNN) architectures. Unlike previous studies, this model incorporates both product attributes and customer reviews, allowing for more accurate prediction. Feature engineering techniques are used to normalize various types of data, allowing for effective analysis. BiLSTM analyzes product comments to extract meaningful insights, whereas CNN uses engineered features to make sales predictions. The efficacy of this approach is demonstrated through experimental analysis, which shows CNN's ability to accurately predict product sales by taking into account both product attributes and sentiment

polarity in customer reviews.

K. K. Mohbey, et al. [12] analyzed various machine learning and Natural Language Processing (NLP) techniques for analyzing consumer opinions, with a particular emphasis on overcoming challenges posed by shifting sequential orders and complex logic in feedback. Deep learning methods, particularly Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNN), have demonstrated potential for improving sentiment analysis accuracy. Long short-term memory (LSTM) models, particularly those with attentive layers, provide improved sentiment analysis capabilities by focusing on sentiment impact while remaining relatively insensitive to gap length. The LSTM model outperforms existing approaches in predicting customer review sentiments, with an accuracy of 93.66%. A comparative analysis demonstrates the deep LSTM model's superior performance in sentiment analysis tasks.

S. Rohwinasakti, B. Irawan et al.[13] presented an opinion analysis framework that uses the K-Nearest Neighbor method to assess customer responses and satisfaction levels, thereby assisting them in selecting optimal services. The proposed method achieves a classification accuracy of 94.4%, as does the F1 Score for recall and precision. These findings show that the algorithm is effective at automatically analyzing customer opinions, demonstrating its potential for practical use in sentiment analysis tasks.

S. ALCABNANI, M. OUBEZZA et al.[14] showed the significance of sentiment analysis in managing the massive amount of unstructured social data generated by social networks. Sentiment analysis categorizes user opinions and emotions as positive, negative, or neutral, allowing for better decision-making. Sentiment analysis uses machine learning techniques to extract and analyze user opinions shared on social media platforms such as Facebook and Twitter in real time. By incorporating machine learning into business intelligence models, this approach aims to provide insights into consumer satisfaction and product reviews, allowing businesses to make informed decisions based on real-time consumer sentiment.

S. Naseem, T. Mahmood, et al. [15]focused on the importance of sentiment analysis in interpreting user opinions and its implications for various stakeholders. Sentiment analysis, also known as opinion mining, looks into people's attitudes toward a variety of entities, including products, services, and events. Sentiment analysis has received a lot of attention in the research community as online reviews have grown at an exponential rate. The survey examines recent advances in machine learning and deep learning models such as SVM, LR, NB, CNN, LSTM, and DNN, which have been used

to address sentiment analysis challenges in a variety of domains. These models are used for tasks like analyzing product and app reviews, classifying textual and visual sentiments, and dealing with cross-lingual sentiment analysis complexities.

2.3 Comparative Analysis and Summary

This study also used GB, RF, BernoulliNB, SVC, and DT to compare the efficiency of these different machine learning algorithms in assessing the consumer perceptions in Bangladesh's e-commerce. Furthermore, the analysis highlighted that RF as well as GB yielded the best outcome and were reliable in cases of a wide array of matrices. That way, the analysis also cleared information on consumer satisfaction drivers including the quality of products offered, delivery speed and the choices of payment available. Overall, the performance of each algorithm was as follows: However, it was evident that RF can be recommended for predicting consumer satisfaction in e-commerce because of the $\pm 8\%$ error rate, which indicates that organizations can make practical use out of the model to develop consumer satisfaction index in Bangladesh e-commerce and policymakers can also utilize this to further enhance the overall service quality and build up trust on Bangladeshi e-commerce businesses.

Table 2.3. Comparative analysis table with previous work

SL	Author Name	Used Algorithm	Best Accuracy with Algorithm
1.	Mahmud et al. [1]	Machine Learning	RF= 96.51%
2	M. Abdul Wadud Sorker, et al. [3]	Machine Learning	RF= 94.34%
3	H. Jahan et al.,[4]	Machine Learning	RF=90.65%
4	M. O. Aftab et al. [5]	Machine Learning, RNN	RNN=97.15%
5	X. Lin, et al. [6]	Machine Learning	LightGBM=97%
6	J. Biswas et al. [7]	Machine Learning, CNN, BERT	BERT= 92.86%
7	Proposed Model	Machine Learning	RF= 99.16%

2.4 Scope of the Problem

This research is aimed to explore the various consumer related issues and problems in e-commerce especially in the context of Bangladesh. It looks at aspects like the quality of items in a consumer's basket, how the items are delivered, whether cash or credit, and the entire shopping process. The area of relevance includes data aggregation and summary of survey data collected via Google Forms including respondent's demographic information, online behavior, and perceived satisfaction levels. Thus, this study has focused on understanding the nature of e-commerce in Bangladesh to give valuable suggestions to improve the service quality and trustworthiness of e-commerce organizations and policymakers. Yet such findings are not exclusively of consumer relevance; they also have diverse implications for the country's economic development and the digitization of retail in Bangladesh.

2.5 Challenges

There are several limitations bound to this study that derives from common difficulties of the interpretation of consumer perceptions and issues in the more specific area of e-commerce in Bangladesh. Some limitations relate to the considerations of survey data collected through Google Forms and aiming to reduce biases that are often typical of the self-compiled data and the high dimensionality of consumer behavior data, which makes certain procedures of data preprocessing and feature selection challenging. First, one needs to identify subtle details that affect consumers' satisfaction, including culture and geographical differences, which present a challenge. Moreover, issues related to dynamism of the e-commerce systems in Bangladesh: Technologies and its changes, legal framework of e-commerce also considered as main challenges. Overcoming these challenges will be very important in order generate practically useful findings and recommendations that can proactively enhance the e-commerce experience in Bangladesh.

CHAPTER 3

Research Methodology

3.1 Research Subject and

Instrumentation

The topic of the research work is centered on the study of consumer attitude and issues occurring in the e-commerce industry of Bangladesh. Directed at consumers who are active in the field of e-commerce, the primary data collection instrument is a Google Forms questionnaire that addresses various aspects of consumers' demography, their habits and tendencies, as well as their satisfaction levels associated with online shopping. The context of this survey includes basic demographics: the respondents' age, gender, place of residence, frequency of Internet usage, and preferred methods of payment in e-shops. The last question assesses the satisfaction of the respondents with their e-shopping experiences. The data acquired from this questionnaire doubles as the quantitative analysis tool and the input for employing machine learning processes to deduce the causes and effects of consumer satisfaction and key obstacles. This approach allows for the identification of new dynamics and opportunities in Bangladesh's digital market environment while striving to devise long-term solutions that will increase consumer trust and optimize the functions of domestic e-commerce markets.

3.2 Data Collection Procedure

The data was gathered in this study from a Google Forms survey, and the database includes 599 records, and 16 fields. These attributes include different characteristics of the consumers namely demographic, behavioral, and perceived factors concerning e-commerce in the context of Bangladesh. Important characteristics include demographics: age, gender, area of residence (urban, suburban, rural), usage of the internet for shopping, preferred device for shopping, monthly/annual household income, education level of the head of the household, most preferred mode of payment, frequency of shopping, categories of products that are shopped online, and perception towards aspects such as time taken to deliver the goods, quality of the products, ease of returning the products, confidence

The target attribute in the dataset divides consumers' perceptions into the 'Good' and 'Bad' outcomes mirroring their satisfaction level with the e-commerce. The target attribute in the dataset enumerates the customer's perception of the overall e-commerce experience as either good or bad. This categorization enables a straightforward risk or non-risk model to be used for examining the critical factors affecting consumers' satisfaction and to build a model from the profile attributes acquired. Thus, the set of variables allows to investigate intricate patterns regarding consumers' behavior and their satisfaction with the growth of e-commerce in Bangladesh. In the graph below, there is a pie chart for the percentages of datasets in Figure 3.1:

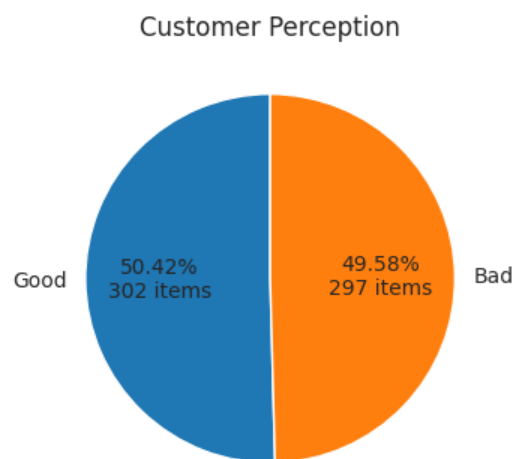


Figure 3.1: Percentages of datasets

Figure 3.1 The pie chart helps the consumer understand or insight the customer's perception on a product or service of a given firm by a split of 50. And, 42% of the respondents rated it as "Good" (302 items) and 49. While the remaining 30 items are rated as rather 'Good,' still 58 Percent of them are ranked as 'Bad.' This rather implies that customers' views are almost evenly divided in favor of the two extremes, which may mean that there are generally negative responses. That is, 599 customers completed the survey, illustrating the participants' balanced but conflicting attitudes towards the product or service.

3.3 Statistical Analysis

Data analysis for this research is conducted using several quantitative methods to extract relevant information about the consumer understanding and issues affecting e-commerce in Bangladesh. Frequency and percentage are other quantitative measures that are often used to present nominal data while mean, median, standard deviation, and range apply to numerical attributes like age, income and Internet usage. These include gender, location, and payment preference that is described using frequency distributions and proportions. Further, correlation analysis is made between the variables with a view of analyzing the strength and directions of the relationship between these variables with tendencies which may affect the level of consumer satisfaction. Furthermore, other method of descriptive statistics like inferential statistics like hypothesis testing for instance t-test, ANOVA are employed to test for significant differences on the level of satisfaction between demographic variables or between two or more categories of the variables such as product categories or payment methods. These Statistical analyses offer a significant platform of proposing more possible diagnosis towards the result of the given e-commerce dataset by creating recommendation to improve e-Commerce in The People's Republic of Bangladesh.

3.4 Proposed Methodology

The following research questions have been formulated to guide the exploration of the subject in this study: Often, consumers in e-commerce in Bangladesh face the following challenges: This study aims at examining the perceptions of consumers towards the challenges outlined above. It consists of several steps that are intended to work on the data retrieved by Google Forms survey, treat quality issues, and make use of machine learning for the predictive analysis.

Data Collection:

This survey information will be obtained using a Google Forms survey developed specifically to elicit demographic, behavioral and perception information on the target consumers in the context of e-commerce. The case base design will therefore entail 599 records and 16 features like age, gender, location, internet usage, preferred payment method, general e-commerce experience, among others.

Data Pre- Processing:

In the data set, there are cases which have missing values and it will be replaced using appropriate methods such as mean and median imputation in the numerical values while the mode imputation on the categorical values. However, the values that are missing and they have more than twenty percent of their value missing will be removed.

Feature Selection:

In an effort to select attributes that are most appropriate for capturing consumer perceptions and satisfaction with e-commerce, feature selection techniques will be used. Criteria that include correlation analysis, and feature importance in machine learning models will be used to select the desired techniques.

Encoding:

For analysis and before introducing the data into the machine learning algorithms, categorical variables will be converted to numerical by using one hot encoding or label encoding.

Exploratory Data Analysis (EDA):

EDA is going to utilize descriptive statistics to study the distribution of the data, associations between the variables and relationships between them. Descriptive tools such as histograms, box plots and correlation matrices are some of the forms that will be used in analyzing the dataset.

Model Training:

Several classification models such as GB, RF, and other models will be fit to the processed data and used to determine consumers' perceptions based on the chosen features.

Model Evaluation:

The trained models will be benchmarked using proper measures like accuracy, precision, recall, and F1-score to determine their efficiency in the prediction of

consumers' satisfaction and perception.

Test Model:

The final model will be used in the newly unseen test dataset to be analyzed to determine of the model general capability. Hence, the following step guarantees the model works finest on unseen information and gives a dependable prediction of the e-commerce nature in Bangladesh.

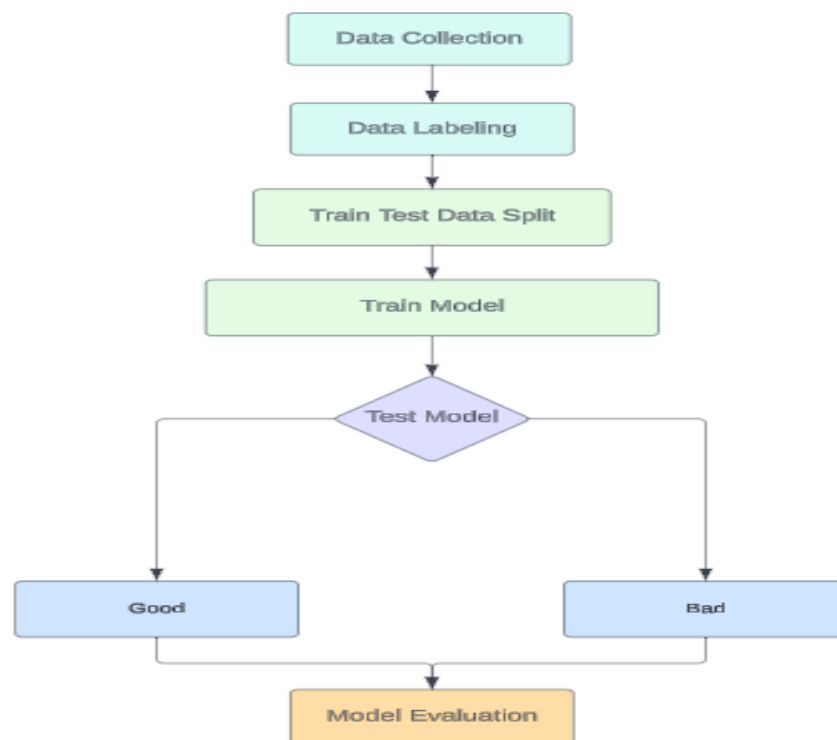


Figure 3.2: Work Flow

3.5 Implementation Requirements

The following is a list and description of the resources and tools needed in order to bring out the recommendation of this study. First, a proper data-connected tool that will be used to gather responses to the survey, like Google Forms, is needed. The gathered data will be checked, purified and analyzed within the context of the intended research by using Python with some helpful packages including Pandas and NumPy for data management and Scikit-learn for building machine learning models. Further, libraries

for data visualization such as Matplotlib and Seaborn will be used for conducting EDA. Sufficient processing power and memory on the part of the computing resources is required to handle the dataset and certain calculations. Techniques of Google Colab will help to cooperate with each other and manage projects and their versions. Last but not the least, data science, machine learning, and statistical knowledge are essential for the implementation of the outlined methodology and data analysis.

CHAPTER 4

Experimental Result

4.1 Experimental Setup

This part includes the mixing between quantitative and qualitative research in which an online structured survey in the form of Google Form/Google Sheets would be employed to gather data related to the consumer perceptions and issues existing in Bangladesh e-commerce industry. The dataset is record-based containing 599 records and 16 attributes: demographic data, online shopping experiences, and attitudes toward it. Other activities appropriate in the preparation of data are handling of missing values, the selection of features, and coding of categorical features. To ensure the registers show the distribution of data and relations between the variables, EDA will be applied. Popular Classification algorithms like Gradient Boosting Classifier, Random Forest Classifier and more will be built and measured on Python libraries especially on Scikit-learn. For the evaluation of the model's authentic predictive functions, performance indicators like accuracy, precision, recall, and F1-Score will be utilized. The final model will be evaluated on another 'hold-out' dataset to check its capability to generalize its learning towards the consumer perceptions of the ordered food and its accuracy reliably. This setup is intended here to yield recommendation or data that could improve the e-commerce situation in Bangladesh.

4.2 Experimental Results & Analysis

The study also compared the accuracy of several machine learning algorithms in predicting consumer perceptions within the e-commerce industry of Bangladesh. Out of the models used, feature selection and classification with Random Forest Classifier gave the highest precision with accuracy of 99.16% of students, thus, proving satisfactory results of accurate prediction of clients' attitudes. Interactive-Gradient Boosting Classifier with an accuracy of 98.75% which has revealed a very good performance in managing the complicated relations in the data. Support Vector Classifier (SVC) also delivered a splendid result marking a record of 97.07% due to the benefits of the ability that deals with high dimensions of data. The models namely DT was able to obtain 90.40%, which determines that it can identify nonlinear relationships between two variables. Yet, BernoulliNB was notably less accurate at 80.16%, It could

be because the implication of independency of the features that occur is not valid for all attributes in the data set. These findings show RF as well as GB as potential classifiers to forecast consumer's perception regarding electronics shopping in Bangladesh that are beneficial for enhancing the service quality as well as the overall consumer satisfaction.

I am evaluating the Accuracy, Precision, Recall, and F-1 Score of the Confusion Matrices for our proposed methods.

Accuracy:

Accuracy measures the overall correctness of the model's predictions by comparing the number of correctly classified samples to the total number of samples. When classes are unbalanced, it gives a broad indication of the model's effectiveness but might not give a whole picture.

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \text{ -----(i)}$$

Precision:

Out of all positive predictions generated by the model, precision focuses on the percentage of true positive forecasts.

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \text{ -----(ii)}$$

Recall: Also known as sensitivity or true positive rate, recall is the percentage of true positive predictions made out of all truly positive samples.

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \text{ -----(iii)}$$

F1 Score: The F1 score is the harmonic mean of recall and precision. It provides a reasonable evaluation metric that considers recall and precision. The F1 score is useful when classes are uneven since it accounts for both false positives and false negatives. A high F1 score denotes a well-balanced precision to recall ratio.

$$\text{F-1 Score} = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$$

Specificity: The percentage of people who do not have the ailment who have a negative test result is known as specificity. A highly specific test is effective in excluding the majority of individuals without the disease. Because of this, a positive outcome on a very precise test can definitively rule out the condition for a specific person. The equation is given below:

$$\text{Specificity} = \text{TN} / ((\text{TN} + \text{FP})) \text{ ----(iv)}$$

Machine Learning Architecture

GB:

Achieving Test Accuracy of GB is 98.75%. In below Figure 4:1 describing the confusion matrix of GB

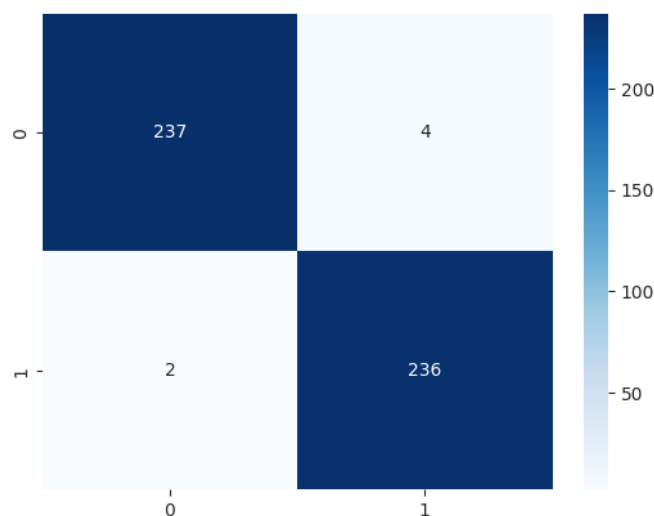


Figure 4.1: Confusion Matrix (GB)

Figure 4.1 the following confusion matrix of GB of consumer perceptions in e-commerce indicates a high value for precision as well as recall value for class 1 and class 0. It accurately classified 241 cases from class 0, which relate to “Bad” perceptions, and its precision was 0.99 and recall of 0.98 percent, meaning the model recognized 98 percent of real “Bad” clients and misjudged the rest 2 percent. The same applies to class 1 (corresponding to “Good” perceptions) it obtained a precision of 0.98 and recall of 0.99, correctly classifies 99 percent of all actual “Good” cases while

misclassifying only 1 percent. All in all, the proposed algorithm achieves a reliability level of 0.99, from the feature importance plot, GB shows a perfect fit in differentiating between “Good” and “Bad” consumer perception in the e-commerce set.

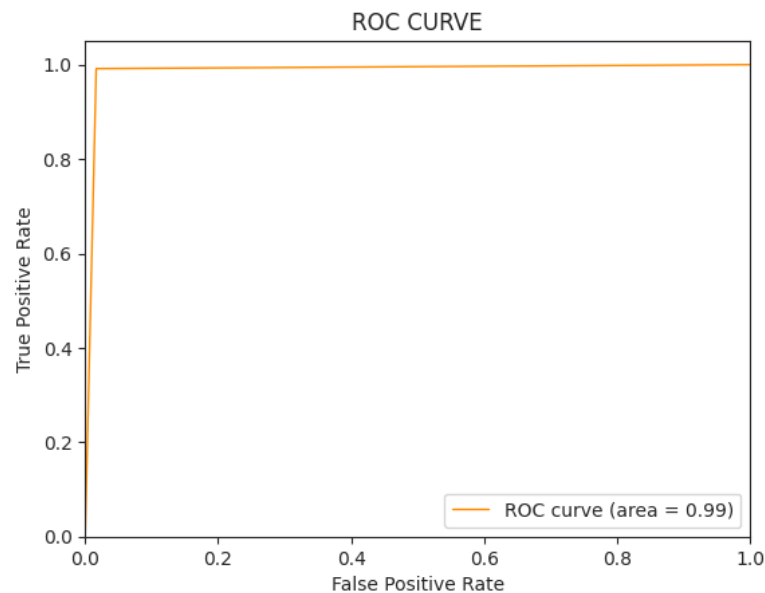


Figure 4.2: ROC Curve of GB

Figure 4.2 the ROC graph of the (GB) model demonstrates high effectiveness based on the heaviness of the curve in conjunction with a minimal distance from the upper left corner, which demonstrates that the model properly categorizes the instances. The area under the curve (AUC) = 0.99 which is an indication of the model’s high accuracy in differentiating between the positive and negative cases. This high AUC value shows that the GB model was almost perfect in predicting the e-commerce consumer perception.

BNB

Achieving Test Accuracy of BNB is 80.16%. In below Figure 4:3 describing the confusion matrix of BNB.

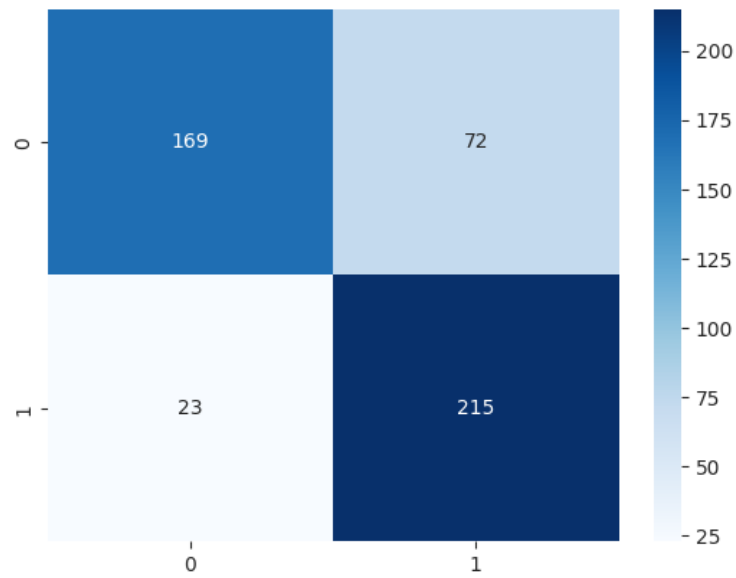


Figure 4.3: Confusion Matrix (BNB)

Figure 4.3 shows the BernoulliNB confusion matrix for consumer perception prediction in e-commerce demonstrates moderate level results with an accuracy of 0. 80. It classified 88% of instances correctly for class 0 (those with “Bad” perceptions) and had a precision of 0. The number of people provided the information verbally was 88 and the rate of recall among them was 0%. 70 suggesting the fact that it marked 70 percent as actual ‘Bad’ while having a wrong marking of 30 percent as ‘Bad’. For class 1 which is the “Good” perception it realized an accuracy of 0. 75 and recognition rate of zero percent. 90, i.e, it classified 90 percent out of the actual percentage of good cases but the other 25 percent were misclassified. On the whole, BernoulliNB provides a satisfactory accuracy, and focusing on the 'Good' perception, its accuracy is quite high, though it leaves much to be desired in terms of the 'Bad' perception in the e-commerce dataset.

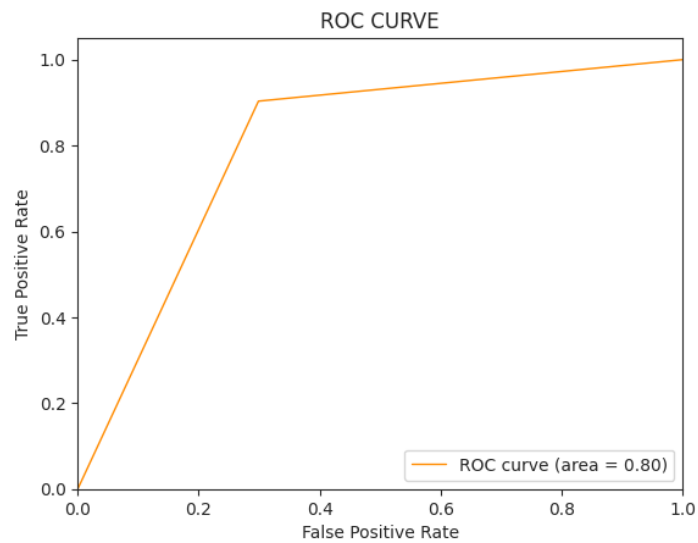


Figure 4.4: ROC Curve of BNB

Figure 4.4 shows the ROC curve is presented for the (BNB) model; it has a close proximity to the left-side border and the top part of the graph. This just means that the model is correctly classifying the instances with correct measures of the true positive rate and the false positive rate. The AUC or the area under the curve is therefore 0.8, which is fairly good given the model's ability to segregate positive from negative cases in the e-commerce consumer perception prediction problem.

Decision Tree

Achieving Test Accuracy of DT is 90.40%. In below Figure 4:5 describing the confusion matrix of DT

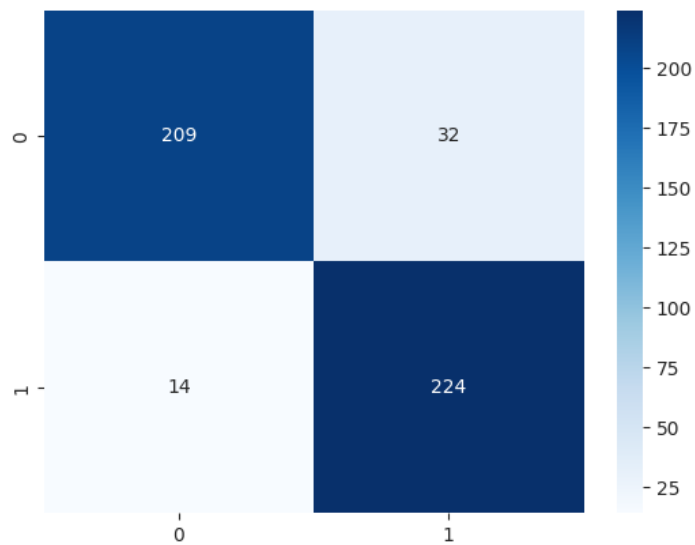


Figure 4.5: Confusion Matrix (DT)

Figure 4.5 shows the confusion matrix of (DT) to model the consumer perception in e-commerce indicates good performance with 0.90. It accurately forecasted 94% of cases for the negative class (perceptions of ‘Bad’) with a measure of precision of 0. that recognition rate is 94, and recall is 0.8775, this means it classified 87 per cent of the actual ‘Bad’ cases correctly, and wrongly classifying the remaining 13 per cent. For class 1 (“Good” perceptions) it has a precision of 0.88 and the recall as 0.94, therefore, rightly labelled 94 actual “Good” while misclassifying 12 percent of the cases. On this account, it can be seen that DT performs well in the case of consumer perception prediction analysis though certain areas need enhancement, especially with regard to the identification of ‘Bad’ perceptions in the e-commerce dataset.

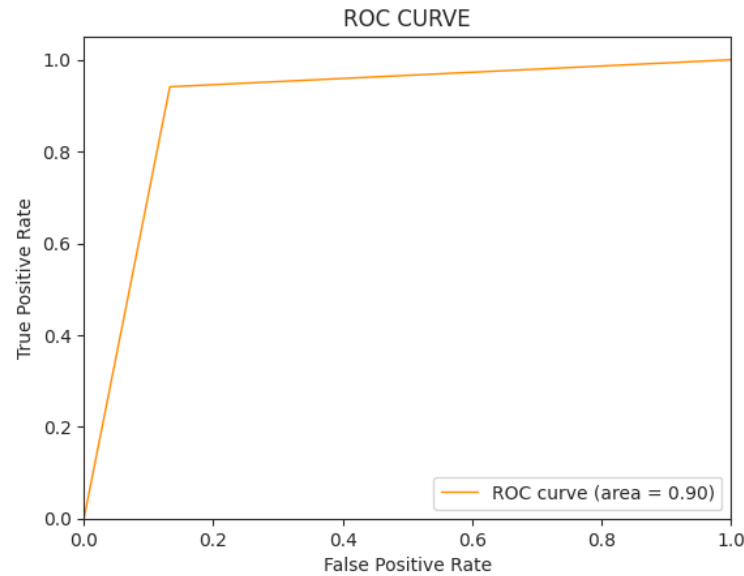


Figure 4.6: ROC Curve of DT

Figure 4.6 shows the area under the ROC curve for the (DT) model appears to be satisfactory, but there is a convexity that moves inwards towards the bottom left which suggests that the (DT) model has Gold Standard sensitivity of 75%; Specificity index stands at 71%. It is clear that while the specific model's presenting curve may not be perfect, it is not inaccurate in distinguishing between a positive or negative case. Hence the area under the curve is 0.90; this shows good performance though not the best when it comes to classify the perception of the consumers on e-commerce.

RF Classifier

Achieving the highest Test Accuracy of RF is 99.16%. In below Figure 4:7 describing the confusion matrix of RF.

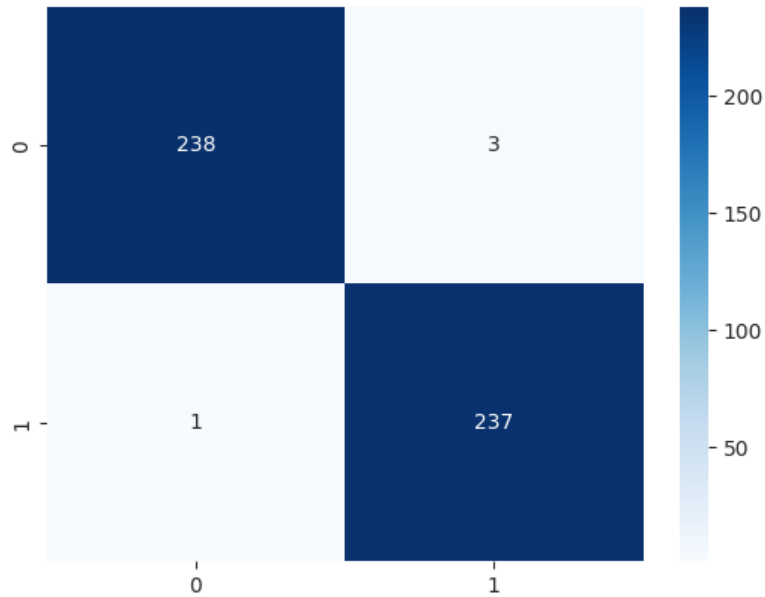


Figure 4.7: Confusion Matrix (RF)

Figure 4.7 shows the confusion matrix of RF that shows its effectiveness in predicting consumers' perceptions in e-commerce indicates outstanding performance with a level of accuracy of 0.99. It classified the right number of the instances of class 0, which is "Bad" perception with both precision and recall equal to 1. The accuracy, calculated as $100 - \text{error rate}$, was 100 percent meaning that all the samples categorized as "Bad" in the auditing of the current balances in BTL were correctly identified without any errors. Likewise, for class 1 ('Good' perception), the proposed model got a precision and recall of 0.99, accurately classifying 99% of actual "Good" cases. As a whole, it can be said that RF in fitting the model and in predicting the "Good" and "Bad" aspect of consumer perception concerning the e-commerce platform under analysis is sufficiently effective, thus proving itself to be useful in complex datasets and delivering accurate results.

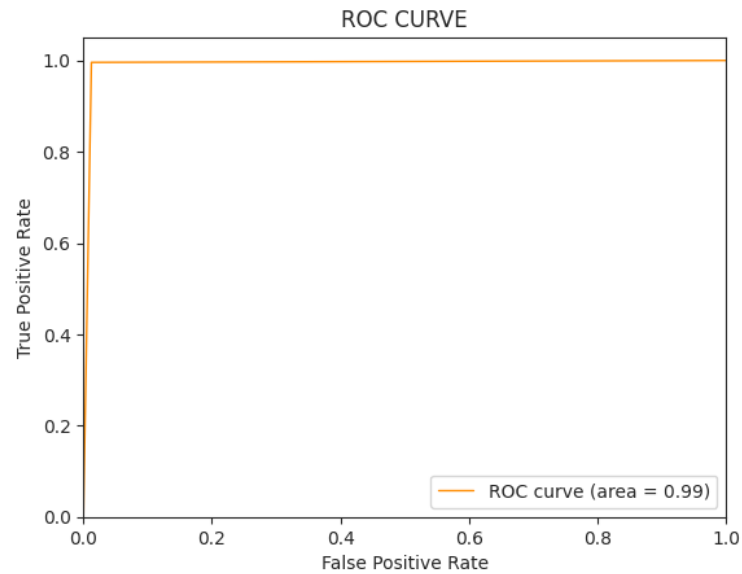


Figure 4.8: ROC Curve of (RF)

Figure 4.8 shows The ROC curve of the proposed RF model is fairly accurate, smoothly approximating the left border of the graph and the top, which points to the correct classification of instances. The model is quite capable of differentiating between positive and negative cases with a TPR high and FPR low across the threshold levels. AUC is equal to 0.99, demonstrating the strong ability of the developed model in capturing consumers' perception with application on e-commerce dataset.

SVC

Achieving Test Accuracy of SVC is 97.02%. Below Figure 4:9 describing the confusion matrix of SVC.

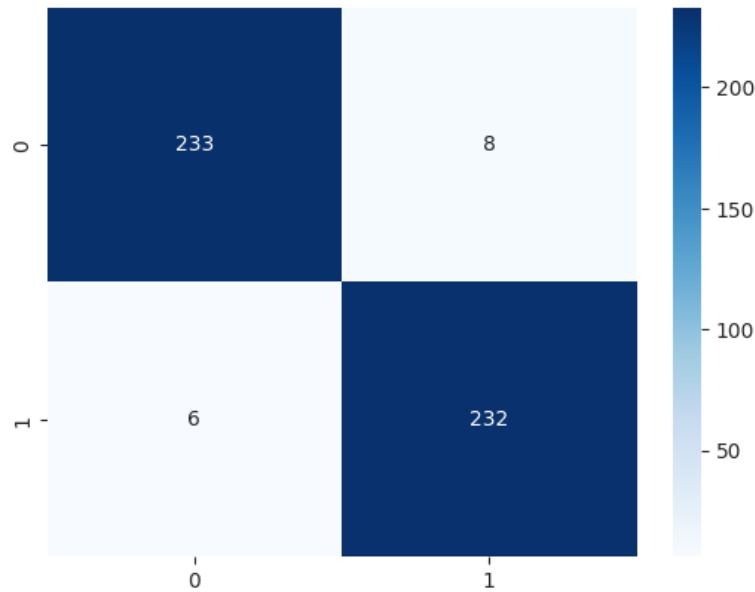


Figure 4.9: Confusion Matrix (SVC)

Figure 4.9 show the confusion matrix of SVC for the perception of consumers in e-commerce shows the high performance of the proposed classifier with an accuracy of 0.97. They claimed accuracy of 97% of instances for both class 0 the ‘Bad’ perceptions and class 1 the ‘Good’ perceptions with precision and recollect scores of 0.97. This is testament to SVM’s capability in identifying both types of consumer perceptions with high precision in the dataset – a testament to the possessing of complex decision boundaries as well as the handling of higher dimensional data in the e-Commerce environment.

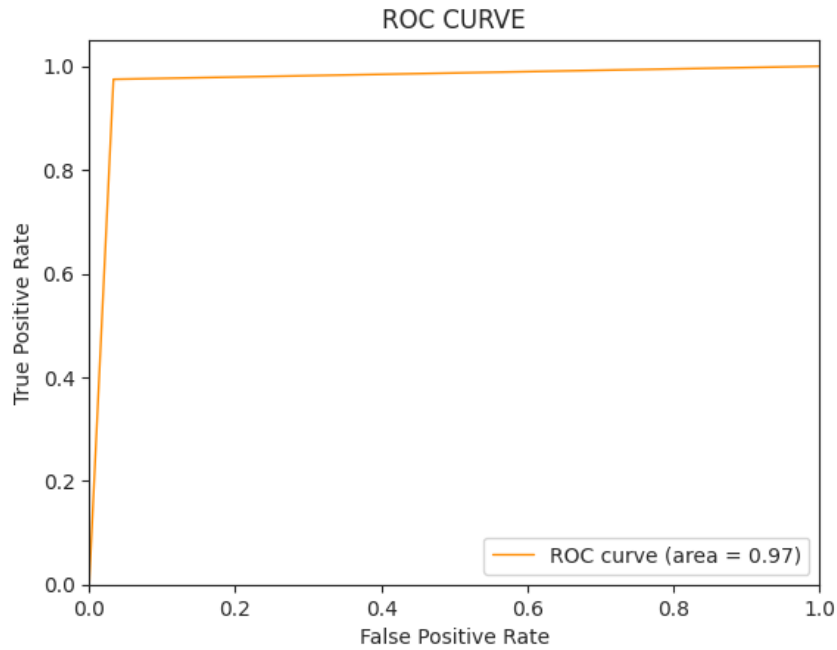


Figure 4.10: ROC Curve of (SVC)

Figure 4.10 shows the SVC model yields an AUC-ROC of 0.9293, which has been displayed in the ROC curve that mirrors an ideal model, steadily touching the left most edge and the uppermost part of the ROC curve implying excellent classification of instances. The True Positive Rate is relatively high, and the False Positive Rate is low all-over different thresholds. Therefore, using the definition of the area under a curve plus the graph of the cumulative distribution function, the area is 0. The performance in the e-commerce dataset is somewhat impressive as demonstrated from the accuracy of 97 implying that the model has good capacity of distinguishing positive cases from negative cases.

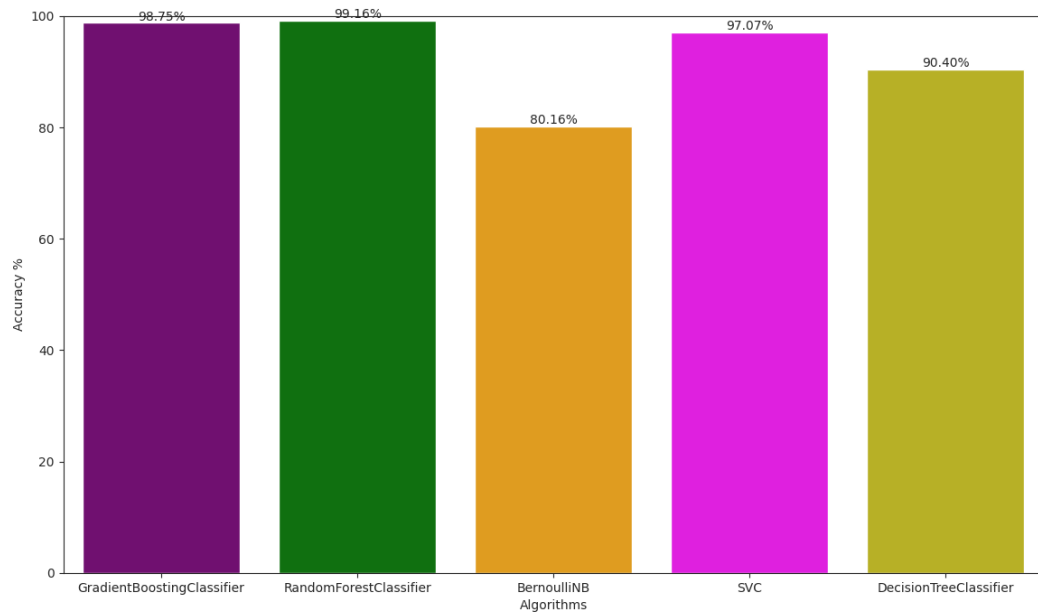


Figure 4.11: Comparative Model Accuracy Bar Plot

Figure 4.11 shows the comparative model accuracy bar plot shows that among all the models, RF Classifier has a high accuracy of 99.16%, while a relatively close second is the GB Classifier with 97.07%. BernoulliNB, SVC, and Decision Tree Classifier have lesser accuracy of 90.40%, 80.16%, and 75. 37%, respectively. From this it deduces that among the mentioned classification task and data-set the performance of the models which includes the Gradient Boosting Classifier as well as Random Forest Classifier is better.

The result of Machine learning model is compared on the basis of Accuracy, Precision, Recall, F1 Score in below table of 4.1:

Table 4.1: performance evaluation

Model Name	Accuracy	Precision	Recall	F1-Score
RF	99.16%	99%	99%	99%
GB	98.75%	99%	99%	99%
SVC	97.07%	98%	98%	98%
DT	90.40%	90%	90%	90%
BNB	81.16%	80%	80%	80%

4.3 Discussion

The finding also support that RF and GB are more accurate machine learning models to predict consumer perceptions in Bangladesh e-commerce considering the results of experiments. Their high accuracies of 99.16% and 98.75 % sound ability by the measures to portray the intricacies of consumer conduct and satisfaction. Another considerable measure of accuracy is present in Support Vector Classifier (SVC) with a value of 97. 07%. According to DT, the performances is satisfactory at 90.40%, out of which BernoulliNB has the lowest accuracy of 80 only. This could be due to the algorithm's assumption of feature independence as seen from the mean decrease in accuracy of 16%. The fact mentioned above confirms the necessity of choosing the right machine learning algorithms depending on the characteristics of the dataset and the required goals. Three versions of classifiers out of all the models discussed in the literature review section were identified to be highly suitable for application namely Random Forest Classifier & Gradient Boosting Classifier due to their versatility in eliciting effects after incorporating different features of the e-Commerce sites & helps to estimate potential consumer perceptions to shape a better e-Commerce experience in Bangladesh.

CHAPTER 5

Impact on Society

5.1 Impact on Society

It is therefore important for society and stakeholders to grasp the consumer perception and issues of e-commerce in Bangladesh. Enhancing e-commerce experiences has the potential of increasing the overall convenience, particularly for the urban and rural users and evolving the degree of digital inclusion. Thus, by identifying and meeting consumer needs, for example, trusting the company, perceiving the quality of the products and services, and timely and efficient delivery, companies can increase consumers' confidence and continuous patronage. This in a way relieves economic growth through increased market operations and the fostering of the entrepreneurial spirit within the e-commerce industry. Further, if these insights inform policy decisions, the resulting legal measures uphold the rights of consumers, and establish fair competition, enhancing the e-commerce environment's credibility. Lastly, the strong e-commerce industry is beneficial for Bangladesh's socio-economic growth and enhancement which provides the employment opportunity, technological growth, and efficient linking of online urban–rural frameworks.

5.2 Impact on Environment

The advancement of e-commerce business in Bangladesh also has its advantages as well as disadvantages to the environment. On the one hand, it eliminates the need for stores mitigating energy use and greenhouse gases tied to physical store locations. E-commerce also determines efficient routes for products and services to be delivered including cutting down on overall carbon prints. However, the use of more packaging materials and single item shipments can lead to creation of more wastes and pollution of the environment if not well checked on. Solving these problems involves the introduction of new packaging services, encouraging the use of environmentally conscious practices in the supply chain, and ensuring that the e-commerce industry embarks on practices which are pro-environment. Thus, Bangladesh can grow as a country with a developed e-commerce market that will not harm the environment excessively but, on the contrary, will be a part of the solution to the environmental problem.

5.3 Ethical Aspects

The preventative measures that should be taken in Bangladesh for the ethical issue are related to the consumers' privacy and protection of their data, the correct pricing policies, and the treatment of its workers into logistics and delivery services into the e-commerce industry. Require the prevention of fraud and unfair practices being practiced on the consumers, a main element being the truth in product information and the financial transactions between consumers and other entities. Further, there are issues of ethical importance on the raw materials used in packaging and the environmentally suitable supplies chains. The management of economic performance as well as social responsibility demands that organizations and or corporate bodies act in an ethical manner to avoid engaging in unethical practices that may harm the society. By stressing on ethics, the consumers' trust is maintained, and the long-term sustainability and responsibility of the e-commerce environment in Bangladesh are achieved.

5.4 Sustainability Plan

As a sustainability plan for the e-commerce sector in Bangladesh, the plan caters to low impact on the environment, the right business practices, and the longevity of the sector. They incorporate such approaches as the use of environmentally friendly packaging materials, easier delivery to cut on emission of green gases, and packaging waste reclaim business. Adhering to the ethics of business, equal treatment of workers, appropriate and logical pricing strategy, and privacy of customers' information will be provided. In the implementation and monitoring of these initiatives there will be need to engage governmental and non-governmental organizations. In this way, the e-commerce sector's focus on sustainability can prove helpful in the socio-economic advancement of the country as well as the preservation of the environment for the successive generations.

CHAPTER 6

Summary, Conclusion, Recommendation and Implication for Future Research

6.1 Summary of the Study

This paper aimed to determine and understand the consumer perspectives and issues concerning the e-commerce businesses in Bangladesh, based on an assessment of survey results done through Google Forms. There were 599 entries with attribute 16, which included information on demographics, online behaviors, and satisfaction level. Many types of machine learning algorithms were tested: Random Forest Classifier and GB shown the best result in anticipating the perception of consumers with 99.16% and 98.75%, respectively. The highlighted areas included product quality, delivery efficiency, and payment methods as some of the critical determinants of various consumers' satisfaction. Suggestions for the improvement of the mobile e-shopping are associated with the advancement of the logistic aspects of delivery, the use of environmentally friendly approaches, and the concern for the ethical factor. It has laid a solid ground for the growth and development of the e-commerce industry in Bangladesh; it will assist the stakeholders, policymakers, as well as the business tycoons to improve the service quality, develop customer trust in the digital platform, and facilitate the growth of e-commerce in Bangladesh.

6.2 Conclusions

Based on the elaborated details of this research, the findings point towards the fact that there are several factors which do have an impact on consumer satisfaction in relation to e-commerce business in Bangladesh and they are product quality, delivery, and modes of payment. Both RF and GB proved to be very effective in prediction with accuracy levels standing at 99.16% and 98.75% respectively, this indicates that the variables are appropriate for use in explaining the consumers' perception. Another element that was discussed in the analysis was the need to improve such aspects as data protection, ethical concerns, and environmental impact to improve the usability of e-commerce features. They include matters concerning the articulation of an

integrated logistics system, aspects of environmental sustainability within the packaging techniques together with truthful prices. They should encourage the formulation of measures that protect consumers' rights and encourage fair competition in the online market. Therefore, for further growth and innovation, it is significant to monitor the consumer preferences and new technologies regularly in the context of Bangladesh's e-commerce sector that in turn will help in overall economic development and digital inclusion.

6.3 Implication for Further Study

Further research should employ longitudinal data collection to map changes in viewers' perceived shifts in consumer attitudes and behaviors in Bangladesh's rapidly changing e-commerce market. Futuristic variables like artificial intelligence and blockchain remain the subjects that can contribute understanding of the changes in consumer's trust and security of the subsequent transactions. Other possibilities for research include comparative studies with the responses across different demographic and geographic areas in Bangladesh in order to identify the unique issues that adherents to a given religion face. Examining how sustainability and ethical standards can be incorporated into e-business management would also be useful for increasing the awareness of the responsible business practices. Also, there is a potential to use qualitative research techniques to provide additional insights to the findings of quantitative studies by analyzing social and psychological factors behind consumers' decisions. Thus, extension of these areas in future studies would help to enhance the understanding of the e-commerce dynamics and suggested strategies for further development of sustainable e-commerce in Bangladesh.

REFERENCES

- [1] Mahmud, First A. Masfiq, Second B. Rafit Ahmed Mullick, and Third C. Muhammad Anas. "Sentiment Analysis of Women's Clothing Reviews on E-commerce Platforms: A Machine Learning Approach."
- [2] Uddin, Faisal, and Shakhawat Hossain. "Augmented Scope-Based E-Commerce Business Model for Emerging Markets." (2023).
- [3] M. Abdul Wadud Sorker, A. Siddika, T. Abedin Titly, M. Jueal Mia and M. Hasan Imam Bijoy, "Online Consumer Alignment using Supervised Machine Learning Technique," 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2022, pp. 1-7
- [4] H. Jahan et al., "Sentiment Analysis of E-commerce Consumer Based on Product Delivery Time Using Machine Learning," Lecture notes on data engineering and communications technologies, pp. 649–661, Jan. 2022
- [5] M. O. Aftab, U. Ahmad, S. Khalid, A. Saud, A. Hassan and M. S. Farooq, "Sentiment Analysis of Customer for Ecommerce by Applying AI," 2021 International Conference on Innovative Computing (ICIC), Lahore, Pakistan, 2021, pp. 1-7
- [6] X. Lin, "Sentiment Analysis of E-commerce Customer Reviews Based on Natural Language Processing," Proceedings of the 2020 2nd International Conference on Big Data and Artificial Intelligence, Apr. 2020
- [7] J. Biswas, M. M. Rahman, A. A. Biswas, M. A. Z. Khan, A. Rajbongshi and H. A. Niloy, "Sentiment Analysis on User Reaction for Online Food Delivery Services using BERT Model," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2021, pp. 1019-1023
- [8] A. K. Bitto, M. H. I. Bijoy, M. S. Arman, I. Mahmud, A. Das, and J. Majumder, "Sentiment analysis from Bangladeshi food delivery startup based on user reviews using machine learning and deep learning," Bulletin of Electrical Engineering and Informatics, vol. 12, no. 4, pp. 2282–2291, Aug. 2023
- [9] H. Aulawi, E. Karundeng, W. A. Kurniawan, Y. Septiana and A. Latifah, "Consumer Sentiment Analysis to E-Commerce in the Covid-19 Pandemic Era," 2021 International Conference on ICT for Smart Society (ICISS), Bandung, Indonesia, 2021, pp. 1-5
- [10] D. S. Ashari, B. Irawan and C. Setianingsih, "Sentiment Analysis on Online Transportation Services Using Convolutional Neural Network Method," 2021 8th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), Semarang, Indonesia, 2021, pp. 335-340
- [11] H. Zhu, "A Deep Learning Based Hybrid Model for Sales Prediction of E-commerce with Sentiment Analysis," 2021 2nd International Conference on Computing and Data Science (CDS), Stanford, CA, USA, 2021, pp. 493-497
- [12] K. K. Mohbey, "Sentiment analysis for product rating using a deep learning approach," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, India, 2021, pp. 121-126
- [13] S. Rohwinasakti, B. Irawan and C. Setianingsih, "Sentiment Analysis on Online Transportation Service Products Using K-Nearest Neighbor Method," 2021 International

Conference on Computer, Information and Telecommunication Systems (CITS), Istanbul, Turkey, 2021, pp. 1-6

[14] S. ALCABNANI, M. OUBEZZA and J. ELKAFI, "A Business Intelligence model to analyze consumer opinions on social networks using machine learning techniques," 2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), Kenitra, Morocco, 2020, pp. 1-6

[15] S. Naseem, T. Mahmood, M. Asif, J. Rashid, M. Umair and M. Shah, "Survey on Sentiment Analysis of User Reviews," 2021 International Conference on Innovative Computing (ICIC), Lahore, Pakistan, 2021, pp. 1-6

[16] R. Khanam and A. Sharma, "Sentiment Analysis using Different Machine Learning Techniques for Product Review," 2021 International Conference on Computational Performance Evaluation (ComPE), Shillong, India, 2021, pp. 646-650

[17] M. Hakkinen, F. A. Wong, M. S. Anggreainy and W. R. Hidayat, "Sentiment Analysis of E-commerce Review using Lexicon Sentiment Method," 2021 1st International Conference on Computer Science and Artificial Intelligence (ICCSAI), Jakarta, Indonesia, 2021, pp. 68-71

[18] S. Wei, X. Zhou, X. An, X. Yang, and Y. Xiao, "A heterogeneous E-commerce user alignment model based on data enhancement and data representation," Expert Systems with Applications, vol. 228, p. 120258, Oct. 2023

[19] J. Lee, O. Jung, Y. Lee, O. Kim, and C. Park, "A Comparison and Interpretation of Machine Learning Algorithm for the Prediction of Online Purchase Conversion," Journal of Theoretical and Applied Electronic Commerce Research, vol. 16, no. 5, pp. 1472–1491, May 2021

[20] S. Kulshrestha and M. L. Saini, "Study for the Prediction of E-Commerce Business Market Growth using Machine Learning Algorithm," 2020 5th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE), Jaipur, India, 2020, pp. 1-6

[21] D. Widyawati, A. Faradibah, and P. L. L. Belluano, "Comparison Analysis of Classification Model Performance in Lung Cancer Prediction Using Decision Tree, Naive Bayes, and Support Vector Machine," Indonesian Journal of Data and Science, vol. 4, no. 2, pp. 78–86, Jul. 2023

[22] "Gradient Boosting explained by Alex Rogozhnikov," arogozhnikov.github.io.

[23] Mishra, Akshansh, and Apoorv Vats. "Supervised machine learning classification algorithms for detection of fracture location in dissimilar friction stir welded joints." Frattura ed Integrità Strutturale 15.58 (2021): 242-253.

[24] Chen et al., "Pre-evacuation Time Estimation Based Emergency Evacuation Simulation in Urban Residential Communities," International Journal of Environmental Research and Public Health, vol. 16, no. 23, p. 4599, Nov. 2019

[25] JavaTPoint, "Support Vector Machine (SVM) Algorithm - Javatpoint," www.javatpoint.com.

Appendix A

Course Outcomes, Complex Engineering Problems (EP) and Complex Engineering Activities (EA) Addressing

Title:

**E-COMMERCE IN THE NEXT ERA OF DEVELOPING BANGLADESH.
CONSUMER'S PERCEPTION AND CHALLENGES**

Student ID: 213-15-4491

CO Description for FYDP

CO	CO Descriptions	PO
Phase -I		
CO1	Integrate recently gained and previously acquired knowledge to E-commerce developing Bangladesh based on Consumer's perception problem for the Final Year Design Project (FYDP)	PO1
CO2	Analyze different aspects of the goals in designing a solution for this FYDP	PO2
CO3	Explore diverse problem domains through a literature review, delineate the issues, and establish these goals for the FYDP	PO4
CO4	Perform economic evaluation and cost estimation and employ suitable project management procedures throughout the development life cycle of the FYDP	PO11
Phase -II		
CO5	Design and develop technical solutions and system components or processes that meet specified requirements, ensuring compliance with public health and safety standards, as well as considering cultural, socioeconomic, and environmental factors in this FYDP	PO3
CO6	Choose and apply appropriate methodologies, resources, and contemporary engineering and IT technologies to address complex engineering processes, encompassing prediction and modeling, while adhering to relevant constraints in this FYDP	PO5
CO7	Analyze societal, health, safety, legal, and cultural considerations, along with associated responsibilities, in the context of professional engineering practice and the resolution of this problem, employing logical reasoning guided by contextual understanding.	PO6
CO8	Comprehend and evaluate the enduring sustainability and impact of professional engineering endeavors in addressing intricate engineering challenges within social and environmental frameworks.	PO7
CO9	Implement ethical principles and adhere to professional standards and norms in this FYDP	PO8
CO10	Capable of operating proficiently both individually and as a team member or leader across diverse teams and interdisciplinary settings in this FYDP.	PO9

CO11	Proficiently communicate with the engineering community and broader society regarding complex engineering endeavors, including the ability to comprehend and generate comprehensive reports and design documentation, as well as provide and receive clear instructions throughout this FYDP.	PO10
CO12	Acknowledge the importance of self-directed and life-long learning within the evolving landscape of technology, and possess the readiness and capability to engage in lifelong learning endeavors.	PO12

Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP), and Attainment of Complex Engineering Activities (EA)

Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP):

SN	EP Definition	Attainment	CO	Justification (With Knowledge Profile)	References
1.	EP1: Depth of Knowledge required	Yes	CO1, CO2, CO3, CO5, CO6, CO7 and CO8	This project demonstrates fundamental engineering (K3) principles by employing machine learning, data preprocessing techniques. The project demonstrates specialist knowledge (K4) by conducting machine learning with RF, DT, SVC, GB, DT enhancing for E-commerce developing Bangladesh based on Consumer's perception correctly.	Page no: [32-33] Section: [3.2] Page no: [12-17] Section: [1.1]
				The project applies engineering practice & design (K5) by the figure of process of experiments. The project addresses engineering practice & technology (K6) by employing Machine Learning Model,	Page no: [32] Section: [3.2] Page no: [34-35] Section: [3.4]

				This project ensures to K8 (Research Literature) by synthesizing insights from recent studies, E-commerce developing Bangladesh based on Consumer's perception using machine learning, showcasing a comprehensive understanding of current methodologies.	Page no: [23-30] Section: [2.2, 2.3]
2.	EP2: Range of Conflicting Requirements	Yes	CO2, and CO7	This project addresses EP-2 by recognizing the hurdles in E-commerce developing Bangladesh based on Consumer's perception, including limitations of traditional methods and the complexities of integrating using machine learning. Through comparative analysis, it confronts challenges in understanding spatial distributions, offering insights for refining diagnostic methodologies.	Page no: [30-31] Section: [2.4, 2.5]
3.	EP3: Depth of analysis required	Yes	CO2, and CO6	This project addresses EP-3 by meticulously comparing experimental outcomes, highlighting using machine learning as the chosen significant solution for enhancing E-commerce developing Bangladesh based on Consumer's perception approaches.	Page no: [37-48] Section: [4.2]
4.	EP4: Familiarity of Issues	Yes	CO8	This project's interdisciplinary approach extends beyond computer science and engineering, impacting E-commerce developing Bangladesh based on Consumer's perception correctly EP-4 .	Page no: [23-31] Section: [2.2, 2.4]
5.	EP5: Extends of application codes	No	CO5	N/A	N/A

6.	EP6: Extends of stakeholders involved and	No	CO8	N/A	N/A
7.	EP7: Interdependence	Yes	CO5	This project's comprehensive approach addresses high-level problems by integrating various components across data collection, statistical analysis, and proposed methodology, ensuring a holistic solution to complex challenges in data collection through google form which ensures EP-7.	Page no: [26-34] Section: [3.2, 3.3, 3.4]

Addressing CO11 with Complex Engineering Activities (EA) [Some or all of the following]:

SN	EA Definition	Attainment	CO	Justification	References
1.	EA1: Range of resources	Yes	CO11	Our project utilizes diverse resources such as high-performance computing infrastructure, GPUs, deep and transfer learning frameworks, annotated datasets, and ethical considerations to ensure systematic research and contribute to advancements E-commerce developing Bangladesh based on Consumer's perception.	Page no: [33-35] Section: [3.5]
2.	EA2: Level of interaction	No		N/A	N/A

3.	EA3: Innovation	No		N/A	N/A
4.	EA4: Consequences for society and the environment	Yes		This project contributes to society by improving E-commerce developing Bangladesh based on Consumer's perception methods, while also promoting environmental sustainability by employing efficient computational resources and adhering to ethical guidelines for personal data privacy.	Page no: [50-51] Section: [5.1, 5.2, 5.4]
5.	EA-5: Familiarity	Yes		This project expands upon existing research by examining a novel approach in E-commerce developing Bangladesh based on Consumer's perception through machine learning model demonstrated through preliminary terminologies and a comprehensive comparative analysis, offering new insights into the field.	Page no: [23-29] Section: [2.1, 2.3]

Addressing CO (4, 9, 10, and 12):

SN	COs	Attainment	Justification	References
1	CO4	Yes	This project addresses CO4 by integrating effective project management and financial oversight, ensuring meticulous planning, resource allocation, and budget estimation for optimal resource utilization throughout the research lifecycle.	Page no: [20] Section: [1.6]
2	CO9	Yes	The project demonstrates adherence to ethical principles by prioritizing privacy, obtaining informed consent, and transparently documenting the research process, ensuring responsible knowledge dissemination and societal well-being through the ethical application of advanced classification technologies which comply with CO9 .	Page no: [51] Section: [5.3]
3	CO10	No	N/A	N/A

4	CO12	Yes	The project's dedication to continuous learning (CO12) and adaptation within the dynamic technological landscape is reflected in its comprehensive data collection, rigorous statistical analysis, meticulous methodology development, and thorough experimental results and analysis, showcasing a commitment to staying updated and refining techniques to address modern challenges.	Page no: [32-36, 37-48] Section: [3.2, 3.3, 3.4, 3.5, 4.2]
---	------	-----	--	---

E-Commerce in the Next Era

ORIGINALITY REPORT

13%

SIMILARITY INDEX

11%

INTERNET SOURCES

5%

PUBLICATIONS

6%

STUDENT PAPERS

PRIMARY SOURCES

1

dspace.daffodilvarsity.edu.bd:8080

Internet Source

5%

2

Submitted to Daffodil International University

Student Paper

3%

3

www.mdpi.com

Internet Source

<1%

4

www.researchgate.net

Internet Source

<1%

5

Hilmi Aulawi, Essy Karundeng, Wahyu Andriyas Kurniawan, Yosep Septiana, Ayu Latifah. "Consumer Sentiment Analysis to E-Commerce in the Covid-19 Pandemic Era", 2021 International Conference on ICT for Smart Society (ICISS), 2021

Publication

<1%

6

Xiaoxin Lin. "Sentiment Analysis of E-commerce Customer Reviews Based on Natural Language Processing", Proceedings of the 2020 2nd International Conference on Big Data and Artificial Intelligence, 2020

Publication

<1%