Exploring the Psychological Factors of Consumer Purchase Behavior in Bangladesh on the Adoption of Digital Payment Apps

by

Samanta Karmaker
20101436
Mst. Elma Fardouse Oishi
20101483
Abrar Qasem
20101505
Shad Been Sharif Sami
20101492

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science and Engineering

Department of Computer Science and Engineering
School of Data and Sciences
Brac University
January 2024

Declaration

It is hereby declared that

1. The thesis submitted is our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

সামন্ত কর্মকার

Samanta Karmaker

ID: 20101436

Elma

Mst. Elma Fardouse Oishi

ID : 20101483

Abrar

Abrar Qasem

D : 20101505

ShadSami

Shad Been Sharif Sami

ID : 20101492

Approval

The thesis titled "Exploring the Psychological Factors of Consumer Purchase Behavior in Bangladesh on the Adoption of Digital Payment Apps" submitted by

1. Samanta Karmaker (20101436)
2. Mst. Elma Fardouse Oishi (20101483)
3. Abrar Qasem (20101505)
4. Shad Been Sharif Sami (20101492)

of Fall, 2023 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science and Engineering on January 18, 2024.

Examining Committee:

Supervisor:

Jannatun Noor

Assistant Professor
Department of Computer Science and Engineering
School of Data and Sciences
Brac University

Program Coordinator:
(Member)

Dr. Golam Rabiul Alam

Professor
Department of Computer Science and Engineering
School of Data and Sciences
Brac University

Head of Department:

Dr. Sadia Hamid Kazi

Associate Professor
Department of Computer Science and Engineering
School of Data and Sciences
Brac University

Acknowledgement

Firstly, we offer our utmost gratitude to the Almighty for enabling us to conduct this study effectively.
Additionally, we would like to thank our supervisor, Dr. Jannatun Noor Mukta ma'am, and the entire C2SG research group for their kind support and invaluable counsel throughout our work. Their expert advice and mentoring made this study a reality.
Lastly, we cannot express our gratitude to our families for their constant support in this complex roller-coaster journey. Without their encouragement, sacrifices and prayers, we might not have been able to be in this position.

Abstract

Consumer purchase behavior and its relation with human psychology is a fascinating and important field of study. Understanding consumer behavior and factors that can drive their decision on adopting new products and services are some of the fundamental topics to study for any startup to succeed. In this research, we will dive deep into the world of consumer purchase behavior and how the complex motivation, attitude, preference and perception along with demographic features drive a consumer's behavior in the context of the digital payment app industry that has taken Bangladesh by a storm. Moreover, we will explore how the adoption of digital payment apps have influenced the consumer purchase behavior of the Bangladeshi market. To fulfill our objectives, we have taken help from a diverse pool of methods such as survey design to obtain user generated data, a combination of statistical and state of the art Machine Learning based methods for quantitative analysis and incorporating basic Natural Language Processing tools along with thematic analysis for a comprehensive qualitative analysis. In the end, our research will serve as a bridge between consumers and fintech service providers, facilitating user - friendly services and innovative marketing strategies.

Keywords: Consumer Purchase Behavior, Digital Payment Apps, Human Computer Interaction, Machine Learning.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Consumer purchase behavior is one of the crucial subjects to study in the field of modern day marketing. It is the study of consumer behavior before, during and after purchasing a product or service [32]. There are many factors that influence consumer purchase behavior such as personal, psychological, social and cultural factors [32]. In today's competitive market, understanding consumer purchase behavior is very crucial to achieve any business goal. According to Zhong et al. [14], analyzing the behavior of consumers is both a prerequisite for engaging in marketing activities and their foundation. It is also a significant aspect that plays a vital role in the decision-making process of businesses [14]. Companies can understand consumers' psychology through the analysis of consumer purchasing behavior, which is vital to the formulation of marketing strategies, product development, and product pricing. As a result, businesses that are able to effectively manage the data collected from their customers' feedback have, on average, a 6% higher profit margin and a 5% higher rate of productivity than their competitors [28]. Furthermore, user behavior analysis is essential for service providers to understand customers' preferences, interests, and needs. It can help them make data-driven decisions, optimize services, drive innovation, retain and acquire customers, and gain a competitive edge.

Digital payment apps are applications that enable users to make online transactions using their smartphones or other devices. They are a form of electronic money that can be used to pay for goods and services, transfer funds, or store value. Digital payment apps have become a transformative force in the economic landscape, as they offer convenience, speed, security, and accessibility to both consumers and businesses. According to Statista [36], the total transaction value in the digital payments market in Bangladesh is predicted to reach 14.67 billion USD in 2024, and 26.72 billion USD in 2027, with a compound annual growth rate (CAGR) of 19.07%. They have also changed the way consumers make purchasing decisions, as they provide them with more information, options, and feedback.

However, despite the growing popularity and adoption of digital payment apps, there is a lack of research on how they affect consumer purchase behavior, especially in developing countries like Bangladesh. Moreover, it is important to understand the influence of personal factors such as age, gender, occupation and economic condition and psychological factors such as consumer motivation, attitude perception and concerns on consumer behavior in the context of the digital payment industry of Bangladesh. Thus, this study aims to fill this gap by exploring the influence of personal and psychological factors on consumer behavior in the context of the digital payment industry and the effects of usage of these services on the Bangladesh market. To accomplish our objectives, we will use primary data collected through a well thought out survey questionnaire. Then using a combination of traditional methods like chi square test and thematic analysis along with state of the art machine learning algorithms, we will try to find the factors that determine consumer behavior in a comprehensive way which is also not present in the current literature. In the end, this study will mark a new beginning in the field of consumer behavior analysis and have far reaching implications in the field of data analytics, marketing strategy development and product development.

1.1 Gaps in Existing Literature

Although, the current literature provides valuable insights into the field of consumer purchase behavior and digital payment apps, it is also important to consider the limitations of the studies.

First of all, In the study by Afroj, Hasan and Fuad [27] has the limitations, including the lack of access to key user information due to privacy concerns as Uber declined to share such data. Furthermore, the study does not incorporate information regarding surge pricing. The study prioritizes understanding user motivations but overlooks other crucial psychological influences on user decision-making. Furthermore, its geographic limitations make results less generalizable. Employing advanced data analysis methods could enhance the study's insights. Secondly, in the research of Mahmud et al. [30] has limitations like using only one data source, not considering reviewer information, and not conducting a more in-depth analysis of the data to gain a better understanding of user perception. Also, in case of Balakrishna, Lok and Rahim [20], the accuracy of all models used were found to be unsatisfactory, possibly due to the noisy nature of the social media data obtained from Google Play Store reviews which could be solvable by obtaining data through proper survey. Moreover, The previous studies do not include studies from diverse geographical locations or cultural contexts, which may limit the applicability of the findings to other settings. Also Lewis and Perry [9], in there study, only focuses on individuals' daily financial behaviors, encompassing expenditure, savings, sharing, and budgeting practices but it does not show the comprehensive portray of personal, social and psychological factors which influence the of consumer purchase behavior the most. Additionally, the aim of Himel et al. [22] was to gain insights into the users' attitude and intention regarding the adoption of digital financial services in Bangladesh. The study does not show the influential factors like psychological,personal and demographic changes of consumers. It only explored how Bangladeshis feel and plan to use digital financial services. Additionally, some studies focus solely on behavioral intention and do not examine actual usage behavior [22]. Furthermore, Zhou et al. [37] only used Multinomial Logistic Regression (MLR) as their ML analysis technique is too specific and doesn't fully capture the limitation of relying solely on one algorithm. Relying solely on Multinomial Logistic Regression might limit the study's ability to capture the full complexity of the data. The use of a single ML technique could overlook potential alternative explanations for the findings. Without exploring other models, the study's conclusions might be specific to the chosen algorithm and less generalizable.

We can see that there are several failings like exploring the influence of personal and psychological factors, such as consumer motivation, attitude, perception, and concerns, on consumer purchase behavior in the context of the digital payment industry from the existing studies. Additionally, the existing works have the lackness of exploring the effects of the usage of digital payment apps on the Bangladesh market, such as consumer satisfaction, adoption, spending habits, perceived security, financial management etc.

1.2 Motivation and Implication

The motivation of our study is to provide a comprehensive and in-depth understanding of the factors that influence consumer purchase behavior in the context of the digital payment industry in Bangladesh. Although this industry is a fairly recent development in the country, it has become an integral part of the daily life of many consumers. Additionally, it also has a vital impact on consumer purchase behavior, as it offers convenience, speed, security, and accessibility to financial transactions. However, the existing literature on this topic is limited and insufficient, as it only focuses on the psychological factors themselves such as consumer motivation or tries to understand consumer purchase behavior on its own. Thus, this study aims to fill this gap by exploring the influence of personal and psychological factors, such as consumer motivation, attitude, perception, and concerns, on consumer purchase behavior in the context of the digital payment industry. Moreover, our study also examines the effects of the usage of digital payment apps on the Bangladesh market, such as consumer satisfaction, adoption, spending habits, perceived security, financial management etc. Lastly, we aim to provide a comprehensive framework that combines traditional methods with more sophisticated methods that can extract hidden patterns and correlations within complex primary data to figure out the most vital factors that influence consumer buying decisions.

The implications of this study are manifold and far-reaching, as it contributes to the fields of consumer behavior analysis, data analytics, marketing strategy development, and product development. Firstly, this study provides valuable insights and recommendations for businesses and policymakers to design and improve digital payment services and strategies that can enhance consumer satisfaction and loyalty, and foster the development and adoption of digital payment apps in Bangladesh. Secondly, this study employs a combination of traditional and state-of-the-art methods, such as chi-square test, thematic analysis, machine learning, and natural language processing, to analyze the complex user generated data collected from a well-designed survey questionnaire. This will demonstrate the potential and usefulness of these methods for extracting hidden patterns and correlations within the data, and provide valuable insights. Thirdly, this study marks a new beginning in the field of consumer behavior analysis, as it explores a novel and emerging topic that has not been adequately studied before, especially in developing countries like Bangladesh.

In a nutshell, this study will provide a framework and a model that can be applied to other contexts and industries, and opens up new avenues and directions for future research.

1.3 Research Objectives

In this study, we aim to conduct a comprehensive investigation into the dynamics of consumer purchase behavior within the context of the digital payment industry in Bangladesh. Our primary objectives are to provide a comprehensive understanding of the relationship between personal and psychological factors, while also shedding light on how the usage of digital payment apps influences consumer purchase behavior across different groups of consumers with diverse demographics and psychological features.

1. **Holistic Data Collection:** Employ structured survey questionnaires for accurate and targeted insights into consumer purchase behavior.

2. **Advanced Data Analysis Techniques:** Leverage machine learning, data visualization, and natural language processing (NLP) alongside human computer interaction (HCI) based qualitative and quantitative analysis to uncover hidden patterns, correlations, and predictive models in consumer data.

3. **Understanding Influential Factors:** Explore the complex relationship of personal, psychological factors shaping consumer choices in general of digital payment app users.

4. **Demographic and Psychological Variation:** Investigate how consumer purchase behavior varies across demographics, considering concerns, attitude, perception, and motivations.

5. **Enhanced User-Provider Relations:** Bridge the gap between the producers and the market, leading to innovative marketing strategies, improved user experiences, and increased business revenue.

Apart from all this, our primary objective is to answer the following questions that are not being answered yet in present literature:

1. How personal and psychological factors influence consumer purchase behavior in the context of the digital payment industry in Bangladesh?

2. How the usage of digital payment apps influence the consumer purchase behavior of the Bangladeshi market?

3. How does consumer purchase behavior in the context of digital payment apps vary based on demography and psychology?

1.4 Our Research Contribution

In this study, we have investigated the personal and psychological factors that influence the consumer behavior of the Bangladeshi market in the context of the digital payment industry and how digitalization of financial transactions has affected and will affect the consumer behavior of this region. The primary contributions of our study are presented below. Our study,

- Reveals the personal and psychological factors that have major influence on consumer behavior in the context of the digital payment industry of Bangladesh.
- Showcases the effects of digital payment app usage on the purchasing behavior of an average consumer in Bangladesh.
- Provides a roadmap to conduct surveys and analyze user generated real world data using modern state of the art methods to get deeper insights and hidden patterns.
- Provides a framework that can be applied to other contexts and industries for market research and consumer behavior analysis.

1.5 Research Organisation

The study is organized and segmented into further sections. Each section is dedicated to explaining a crucial part of this study. Firstly, we have the Literature Review section, which consists of related works of our study. Then, we have the Background section that explains the models, techniques and metrics used to analyze the dataset. After, in the Research Methodology section, we will discuss the methodology we used including how the survey was designed, how we collected data, the actual dataset along with how the data was analyzed both qualitatively and quantitatively. Next, in the Result section, we have looked into the results of our analysis and the performance of the models used. Furthermore, we have discussed how the results answer our research questions and the limitations of our study in the Discussion and Limitations section. Finally, the conclusion and future prospects of this study are elaborated in the Conclusion and Future Work section.

Chapter 2

Literature Review

In this section, we have discussed some of the previous studies we have reviewed to enrich our knowledge related to our topic. As our thesis work is highly inter – disciplinary, we have divided our attention to four key parts of our thesis topic which are consumer purchasing behavior, user review analysis, sentiment analysis and HCI and digital payment apps.

2.1 Consumer Purchasing Behavior

In this section, we have focused on consumer purchasing behavior analysis. Here, we looked into studies related to the 5 stages of consumer purchasing behavior and factors that influence consumer purchasing behavior such as personal, social and psychological factors. Moreover, we look into studies related to the application of consumer purchasing behavior. Lastly, we have put more emphasis on services and studies that incorporate survey questionnaires to obtain data.

Firstly, Omar and Atteya [17] investigated the impact of three digital marketing channels, namely Email Marketing, Mobile Marketing, and Retargeting, on different stages of the consumer decision process in Egypt. The authors organized an online survey and performed statistical analysis to examine the connection between digital marketing methods and consumer decision-making. Results indicate that Email Marketing positively influences consumer decisions in the post-purchase and information research stages, but negatively impacts the purchase decision stage. In contrast, Mobile Marketing has a negative impact on consumer decisions in all stages, while Retargeting has a significant positive impact on the evaluation stage, as well as information research, recognition, purchase decision, and post-purchase stages. Notably, Retargeting emerged as the most influential variable, with the strongest effect observed in the evaluation stage. On the other hand, Mobile Marketing was found to have the least impact on the consumer buying decision process and had a negative effect on the purchase decision.

Now, Let's look into papers discussing the application of consumer purchase behavior data. For example, Choudhury and Nur [6] aimed to identify potential customers in the grocery industry of Bangladesh, using sales data from a super shop in Sylhet called Taradin. The authors constructed engineered features and compared the results with native ones, utilizing five machine learning classifiers, including a deep learning-based model called the Multi-Layer Perceptron (MLP) classifier. The results indicated that the engineered features outperformed the native features, with the MLP classifier achieving 99.41% accuracy. Conversely, the best performance for the native features was demonstrated by Logistic Regression, with an accuracy of just 56.78%. The study's main limitation is the lack of generalizability for online grocery industries, which have significantly more personalized features available.

Similarly, Li et al. [10] focused on determining the segments for which email marketing would result in higher potential sales for a multinational bicycle company by applying consumer purchasing behavior analysis and Machine Learning techniques. The researchers utilized existing customer information and attributes, applied both supervised and unsupervised machine learning algorithms, like Decision Trees, Cluster Analysis, and Naive Bayes Classifier. According to the study, Decision Tree showed the best performance. However, the main limitation acknowledged by the researchers was that they only worked with the company's historical data and did not consider other complex market environment factors.

Likewise, Tabianan, Velu and Ravi [31] utilized the K-Means clustering algorithms to segment customers based on their purchase behavior data within the Malaysian e-commerce industry. The results showed that using more stable data, for example, nominal values smoothed by the Z score, improved the stability and variability in constructing clusters over time. The researchers also identified three maps generated using the hierarchical K-Means algorithm as the most representative of profitable segments. Furthermore, the authors suggest that a deep learning clustering approach could be a potential future direction for research in this area.

Here, let's look into studies that focus on understanding the factors that influence consumer purchase behavior. For instance, Sharma, Singh and Pratt [18] aimed to investigate whether there is a difference in the intention of Baby Boomers and Millennials to purchase travel and tourism services online in Australia. To get the data, a survey was conducted using an extended version of the UTAUT2 model, and various psychological variables were considered. Results indicated that there is indeed a difference in the intention of purchasing travel online between the two generations, even in a developed country like Australia. To assess the effectiveness of the study, R2 values were used, which indicates the percentage of variance in the construct accounted for by the model. The R2 values increased from 33 percent to 38 percent with the inclusion of attitude and to 64 percent with the addition of compatibility, innovativeness, perceived risk, and perceived trust. The R2 values were categorized into weak (19 percent), moderate (33 percent), and substantial (67 percent) levels. The increase in R2 highlights the importance of the constructs in predicting online purchase intention. The researchers recommended exploring the underlying reasons for the differences in consumer behavior between generations and conducting further research on other industries and countries due to contextual differences that may affect the results.

Similarly, the primary objective of Afroj, Hasan and Fuad [27] was to gain insights into the characteristics of ride-hailing users, their trip behaviors, the impact of ride-hailing services, and service quality. The study specifically focuses on Uber as the chosen ride-hailing company and employs survey interviews conducted in the hotspots of Dhaka, utilizing a simple random sampling method. The results indicate that Uber is predominantly favored by a wealthier and younger demographic. Moreover, the preference for Uber over other transportation options is influenced by factors such as improved accessibility, safety, comfort, and the availability of emergency assistance. Additionally, Uber users often choose to share their trips with individuals they are familiar with. The study also reveals the users' expressed need for enhancements in the facilities provided within Uber cars and improvements to the GPS system in the app. However, the research has certain limitations, including the lack of access to key user information due to privacy concerns as Uber declined to share such data. Furthermore, the study does not incorporate information regarding surge pricing. In my opinion, while the study effectively covers personal factors influencing user behavior, it primarily focuses on user motivation and neglects to delve into other psychological factors that play a role in the user decision-making process. Additionally, the study is limited by its geographical constraints and could benefit from the utilization of advanced data analysis techniques.

Likewise, the aim of Himel et al., [22] was to gain insights into the users' attitude and intention regarding the adoption of digital financial services in Bangladesh. To achieve this objective, the authors employed structured survey questionnaires and group discussions. They utilized two theoretical frameworks, namely the Technology Acceptance Model (TAM) and the Innovation Resistance Theory (IRT), to validate their findings. Based on their analysis, the researchers established that perceived usefulness, ease of use, and trust positively influence user attitudes, while barriers to adoption have a negative impact. However, it is important to note that this study has certain drawbacks. For instance, the sample size was relatively small, comprising only 196 participants, and the non-probability sampling method used may introduce biases and restrict the generalizability of the results. Additionally, this study focuses solely on behavioral intention and does not examine actual usage behavior. Moreover, it primarily investigates the effects of perceived usefulness, perceived ease of use, perceived trust, and barriers to adoption on attitudes and intentions.

2.2 User Review Analysis

In this part of the study, we look into the studies related to user review analysis mostly in context of user behavior analysis and services.

Firstly, Chen et al. [15] investigated user satisfaction in online education platforms in China during the Covid-19 pandemic. A survey was conducted among students, and web crawling was used to gather data. Emotions of the respondents were analyzed through the NLPIR parser, and a user satisfaction index was constructed. A Back Propagation Neural Network was utilized to predict user satisfaction, achieving an accuracy rate of 77.5%. Moreover, the authors found that personal factors of users did not directly impact satisfaction, but platform availability had the greatest influence. Additionally, they suggested that improving platform technology and developing various interactive formats could enhance learning efficiency and improve education quality. Lastly, they suggested that future research should consider the perspectives of parents and teachers, as well as improvements in questionnaire design and algorithm development.

Similarly, Cheng and Jin [5] aimed to understand the factors that influence user experience and satisfaction on Airbnb by analyzing online reviews in the context of Sydney, Australia. The authors used text mining and sentiment analysis to gather insights into the most important attributes that Airbnb users consider when assessing their experience. The study found that amenities, location, and host are the most important factors, and that Airbnb guests tend to follow instructions and guides given by the hosts more often than those given by hotels. In addition, good communication was found to play an important role in building trust and enhancing user experience. The study also highlights the importance of incorporating other variables, such as ratings and property descriptions, and taking user information into account for further insights. Looking into other regions, particularly in developing countries, can provide cross-cultural insights and more generalized ideas. The authors note that their dataset was positively biased, which could limit the generalizability of their findings.

Conversely, Zhong et al. [14] analyzed user reviews of the iPhone 5 in different cultural settings, specifically in the USA, UK, and India, to investigate how cultural factors influence consumer purchasing behavior. By using web crawling to gather data and NLP techniques for data preprocessing, keyword extraction, and sentiment analysis, the authors were able to provide insights that can help Apple improve their product development and marketing strategies for targeting diverse markets. However, one of the limitations of this study, according to the authors, is that they relied on past data and suggest that using real-time big data could increase the effectiveness of future studies.

Now, let's focus on studies that analyze user reviews of different apps. Firstly, Mahmud et al. [30] aimed to analyze user-generated reviews of ride-sharing mobile applications using sentiment analysis techniques. The authors collected data from Google Play Store reviews and preprocessed it using NLP techniques. Next, they used deep learning methods, including CNN, LSTM, and DistilBERT, to classify the reviews into positive, negative, and neutral categories. Their findings showed that DistilBERT outperformed the other models in sentiment analysis. However, the study had limitations, such as using only one data source, not considering reviewer information, and not conducting a more in-depth analysis of the data to gain a better understanding of user perception.

Similarly, Balakrishna, Lok and Rahim [20] analyzed user reviews of digital payment apps popular in Malaysia using various ML and NLP techniques. Their objective was to understand user sentiment and emotions and to identify emerging topics related to digital payment. To do so, the study utilized supervised machine learning models such as SVM, random forest and naive Bayes classifiers. Moreover, unsupervised learning was also employed to identify top emerging topics. Results indicate that SVM and random forest classifiers performed equally well and outperformed the naive Bayes classifier. However, the accuracy of all models was found to be unsatisfactory, possibly due to the noisy nature of the social media data obtained from Google Play Store reviews. Future work suggested includes local language review analysis and adopting a hierarchical approach of first conducting sentiment analysis and then applying emotion analysis on the results.

Likewise, Fransiska, Rianto and Gufroni [16] employed sentiment analysis to analyze Google Play Store reviews of an Indonesian app Provider by.U. From their literature review they found that SVM is widely regarded as the best text classification model and that TF-IDF is the most effective feature extraction method for text analysis. To determine the impact of TF-IDF and find the hyperparameters that yield the best accuracy, the authors utilized hyperparameter tuning. Their research findings indicate that a 5-fold cross-validation with TF-IDF produced the most favorable results. Additionally, while TF-IDF enhanced accuracy, it was not considered a significant improvement. As a potential avenue for future research, the authors suggested utilizing sentiment dictionaries or manual labeling of data and comparing different ML models and feature extraction methods.

Additionally, let's look into a study that concentrates on product reviews on an online marketplace. For example, the objective of this study, by Haque, Saber and Shah, [4] was to analyze product reviews of various items sold on Amazon. The authors used Natural Language Processing (NLP) techniques for feature engineering and applied different supervised learning classifiers to classify the reviews as positive or negative. Moreover, the study also employed active learning, k-fold cross-validation, and varying train-test ratios to achieve the most accurate results. Lastly, the authors reported that 10-fold cross-validation yielded the best results and the SVM model was the top-performing classifier.

Lastly, this paper takes a different approach to analyze user reviews. The objective of the research, by Mukherjee and Bhattacharyya, [2] was to analyze customer reviews on social media platforms to extract opinions about specific product features. The authors argued that analyzing each sentence of a review could provide a better understanding of customer sentiment than considering the review as a whole. Moreover, the study aimed to focus on domain independent parameters. Despite data limitations, the approach achieved high accuracy. However, it should be noted that the system was unable to evaluate implicit sentiment that is dependent on the product domain.

2.3 Sentiment Analysis

In this part of the study, we mainly looked into the studies related to sentiment analysis and tried to analyze the techniques of sentiment analysis and their efficiencies.

In the study, by Rahman and Hossen, [11], the primary objective was to implement sentiment analysis on movie review data using various machine learning classifiers. The researchers collected a dataset consisting of 2000 movie reviews, with an equal number of positive and negative reviews. Data preprocessing steps were performed, including the removal of URLs, brackets, numbers, punctuation marks, and stop words. Next, tokenization was employed to divide the text into smaller units, such as sentences or words. Then, feature vectors were created, which involved classifying each review as positive or negative and identifying opinion words using parts of speech tags. Emotions were determined based on these opinion words, and sentiment scores were used to determine positive and negative words. Machine learning techniques were applied for classification, with 1400 movie reviews used for training and 600 for testing. The results indicated that the Multinomial Naïve Bayes classifier outperformed the Support Vector Machine classifier, although this was dependent on the selected parameters and the size of the training dataset. The accuracy of the Multinomial Naïve Bayes classifier was reported to be 88.5%, which was higher than the other classifiers.

Similarly, Soumik et al. [7] focus on sentiment analysis using text classification techniques applied to a Bangla dataset. The researchers introduce the TF-IDF (Term Frequency-Inverse Document Frequency) algorithm, which converts words into feature vectors. They develop a web crawler to collect data from the Google Play Store, encountering challenges such as misspellings, slang words, and lexical variations. The collected data was stored, and information such as name, rating, and review features were extracted. Manual annotation of the reviews was conducted, assigning numerical values. Here, five-fold cross-validation was employed to obtain unbiased results. Data preprocessing involved removing white spaces, sorting according to the standard sorting order by Bangla Academy, collecting and removing stopwords, and assigning weight to common data using frequency distribution. The TF-IDF approach assigns less weight to words that are irrelevant for sentiment analysis purposes. However, the study faces limitations, such as the lack of proper tool sets for Bangla data preprocessing, leading to manual cleaning of the data. Additionally, some Bengali words are converted to ASCII during processing and are subsequently removed manually. Lastly, The use of a lexicon-based approach for determining the polarity of tweets or comments is not viable due to the unavailability of dictionaries containing positive and negative words in the Bangla language.

Similarly, Tuhin et al. [12] introduce sentiment analysis algorithms specifically designed for Bangla language datasets. It presents an automated process for sentiment analysis, focusing on six emotions. The study proposes two machine learning techniques: Naive Bayes (NB) and a topical approach to extract emotions from Bangla text at both the article and sentence levels. The researchers utilize a manually created dataset comprising 7500 Bangla sentences. During the data cleaning process, mixed sentences, special characters, and punctuation marks were ignored. By running the NB algorithm, emotional scores were obtained for each word corresponding to each emotional class. These scores were then used to determine the level of emotions for a complete sentence. The study also employs a binning technique to enhance performance by providing a value for missing words in the dataset. Additionally, a generic topical approach is utilized for extraction purposes. The primary process involves identifying the presence of six emotion-related words in the training dataset, calculating the value of each emotion feature accordingly. The researchers conclude that the topical approach achieves higher accuracy compared to the NB classification algorithm, both at the sentence level and article level.

Likewise, Alam, Rahoman and Azad [3] present a framework for sentiment analysis of Bangla texts using a Convolutional Neural Network (CNN) model. The proposed model achieves a classification accuracy of 99.87%, surpassing the existing state-of-the-art Bangla sentiment classifiers by 6.87%. The CNN model is trained to classify sentiment from Bangla comments, distinguishing between positive and negative sentiments. Since no publicly available Bangla dataset exists, the researchers generated a small dataset of Bangla sentences for training and testing the model. They collected Bangla comments from various sources, filtered them, and replicated them to increase the comment count. After preprocessing, the comments were utilized to train the CNN model. The CNN approach achieved a training accuracy of 99.87%. The paper highlights that their obtained accuracy is the highest observed for analyzing Bangla sentiment.

Lastly, Naim [24] mainly focuses on aspect-based sentiment analysis (ABSA) and it's implementations in Bangla language. ABSA mainly extracts aspect terms from texts and then identifies sentiment polarity which are connected with those terms. Due to lack of resources it is hard to implement ABSA in Bangla language. This research used PSPWA for aspect term extraction from cricket and restaurant datasets. PSPWA mainly gives priority to noun's and introduces them as aspect terms. Also Supervised learning algorithms and Convolutional Neural Network (CNN) have been used here. TF-IDF is applied here for data preprocessing and for feature engineering. Finally, the research shows improved results using aspect term extraction then other processes but also acknowledges the need for future work on sentiment identification.

2.4 HCI and Digital Payment Apps

In this section, we will look into the literature related to digital payment apps in HCI and try to understand the techniques used for their effectiveness and limitations. Moreover, we will dive deep into the realm of survey design and survey data analysis.

In their study, Lewis and Perry, [9] investigates the potential advantages and obstacles associated with the utilization of digital platforms for managing personal finances. This paper gives a comprehensive investigation into individuals' daily financial behaviors, encompassing expenditure, savings, sharing, and budgeting practices. It elucidates the influence of social and environmental elements on these behaviors, as they are manifested and imbued with significance within the realm of everyday existence. According to them, there is an observed increase in the level of attention given to financial practices and interactions, the adoption of digital payment and wallet systems, and the availability of digital information that is pertinent to the analysis of routine financial management. Moreover, it emphasizes the significance of comprehending the social and contextual elements that influence individuals' financial behaviors in their daily lives. The authors utilize a diverse array of sources in order to offer valuable insights into the optimal design of digital financial systems, with the aim of enhancing their capacity to effectively meet individuals' financial management requirements.

Then the study of [21] Singhal and Gupta (2021) investigates the effects of the COVID-19 pandemic on the utilization of digital payment services in rural regions of India. The researchers did a secondary research study to examine the effects of the pandemic on digital payment systems in rural areas. The literature review section of the study presents a comprehensive examination of the notion of digital payments and its significance within the contemporary digital landscape. The writers emphasize the advantages of digital payments, including their convenience, speed, and security, and discuss the transformative impact they have had on transactional behavior. The authors additionally examine the obstacles encountered by both organizations and individuals when embracing digital payment methods, including limited knowledge, inadequate infrastructure, and diminished confidence. Subsequently, the authors proceed to examine the ramifications of the COVID-19 epidemic on digital payment services. It is observed that the COVID-19 pandemic has expedited the uptake of digital payment services, as individuals are actively avoiding physical contact and cash transactions in light of the potential risk of infection. The authors reference other studies that have documented a rise in the utilization of digital payment systems within the epidemic, with a particular emphasis on rural regions. The literature research additionally underscores the difficulties encountered by businesses in the process of deploying digital payment systems throughout the pandemic. The authors observe that a significant number of small enterprises situated in rural regions have challenges in terms of inadequate infrastructure and limited resources, resulting in their inability to embrace digital payment systems. Consequently, this has impeded the progress and expansion of digital payments within these areas. The authors also examine the significance of cultivating trust and raising knowledge among individuals as a means to promote the adoption of digital payment systems. In general, the literature study offers a thorough examination of the notion of digital payments, encompassing its advantages and obstacles, as well as the influence of the COVID-19 epidemic on its implementation in rural regions. The writers have referenced other pertinent studies in order to substantiate their claims and establish a robust framework for their research investigation.

After that, Zhang, Luximon and Song [13] aim to examine the influence of perceived security, perceived control, interface design characteristics, and conscientiousness on the sustained usage of mobile payment services by consumers. The article commences by providing an overview of the swift advancement of mobile commerce (m-commerce) and the crucial significance of mobile payment services within this domain. The authors emphasize the limited understanding of the relationship between perceived security and the sustainable utilization of mobile payment services. They put forth a study model aimed at investigating the role of perceived security on consumers' ongoing use of mobile payment services. The literature review part of the article is a comprehensive summary of prior research conducted on the topic of felt security and the factors that influence it. The authors acknowledge that several studies have been conducted on the topic of perceived security. However, they highlight the lack of consistency in the findings, which might be attributed to the evolving definition of perceived security within the specific research contexts. The authors additionally conduct a comprehensive examination of prior research pertaining to interface design, perceived control, and conscientiousness, emphasizing their potential influence on perceived security and the sustained adoption of mobile payment services by consumers. The subsequent section of the article outlines the study strategy and methods employed to gather data from a sample of 252 participants via an online survey. The study used structural equation modeling to analyze the results, and the authors provide a discussion of the findings pertaining to the influence of perceived control, interface design characteristics, and conscientiousness on perceived security and users' continuing intention to utilize digital payment services. The article finishes by engaging in a conversation regarding the practical ramifications of the study for mobile service providers and designers. The authors propose that the retention of users in mobile payment services can be improved by boosting the perceived security. To achieve this, they offer design insights that focus on strengthening the perceptions of customisation and feedback design. The authors also highlight the significance of conscientiousness in users' assessment of interface design features and security measures.

Moreover, Zhou et al. [37] proposed a framework for the development of a central bank digital currency. This framework aims to integrate the advantages of digital payments, such as convenience, with the user-friendly nature of physical cash. This study examines the limitations associated with conventional digital payment systems and endeavours to propose a solution that is characterised by accessibility, security, and user-friendliness.The writers commence their discourse by examining the various rationales for individuals' preferences for either cash or digital payment modalities. The authors contend that the process of exchanging physical currency for digital forms of money is predominantly dependent on financial institutions, a circumstance that may prove difficult or unfeasible under specific circumstances. In order to tackle this matter, the authors put up a payment system that enhances the conversion process and enhances its accessibility to users in both the pre-transaction and post-transaction stages, irrespective of temporal or geographical limitations. The report also discusses the matter of effectively handling monetary change in retail environments. The authors examine two potential techniques by which merchants can generate additional revenue, contingent upon varying consumer preferences. The authors provide a proposed system that enables merchants to authenticate the legitimacy of cash transactions without relying on network connectivity. This solution is particularly advantageous in situations when offline payments are involved.In order to guarantee the integrity of the system, the authors put forth a trust-level model that allocates varying degrees of confidence to distinct entities inside the system. The design of the model aims to mitigate instances of double-spending and fraudulent activity, while concurrently facilitating transactions that maintain anonymity. In its whole, "Print Your Money: Cash-Like Experiences with Digital Money" presents a thorough and inventive resolution to the obstacles encountered in the realm of digital transactions. The article exhibits a commendable level of thorough research and proficient writing. Moreover, the proposed design possesses the capacity to fundamentally transform our conceptualization of currency and financial transactions.

According to Kaewkitipong et al. [29], the utilization of mobile payment services has experienced a notable surge in popularity in recent years, as an escalating number of individuals depend on their mobile devices for conducting financial transactions. Consequently, the significance of Human-Computer Interaction (HCI) and trust aspects in the ongoing utilization of mobile payment services has emerged as a subject of interest among scholars and practitioners alike. According to their research, numerous scholarly inquiries have been conducted to examine the significance of human-computer interaction (HCI) within the realm of mobile payment services. In general, existing literature indicates that the utilization of mobile payment services is influenced by both Human-Computer Interaction (HCI) and trust aspects. It is imperative for mobile application developers to prioritize the enhancement of usability and user experience in their applications, concurrently addressing the apprehensions of users regarding security and privacy. In the end, they conclude that by implementing this strategy, organizations can enhance users' confidence in their services and foster sustained utilization.

Furthermore, Tripathi [19] aimed to examine the socio-demographic factors that influence consumer attitude towards mobile payment apps and what are the barriers they face in the state of Gujarat, India. For this study, they collected data from 100 individuals using a well structured google form. Here, they used convenience sampling. Next, they analyzed the data and concluded that consumers are preferring mobile payments due to motivations such as cashback and discounts, convenience, recorded spending history, and reduced theft risk etc. Additionally, they identified barriers hindering the usage of digital payment apps such as transaction charges, trust issues and not having the knowledge to use the app. Lastly, the researchers pointed out that most of the respondents are students from the age of 21-25 with an annual income below 1 lack rupees.

Interestingly, Shree et al. [26] aimed to understand how perception and trust in digital payments, and experience with online frauds, affect the payment behavior of consumers in India. To answer this question, researchers performed an online survey using snowball sampling and collected data from 640 individuals. Their survey was well structured with 28 questions and 7 sections. Moreover, they used Multinomial Logistic Regression to analyze the data. According to their research, perception and trust indeed, have significant influence on the choice of paying digitally. However, when it comes to past experience with online frauds, it impacts consumer preference in case of grocery payments. Lastly, the main limitations of this study is the dataset is skewed with most responses coming from urban, educated and people from affluent backgrounds along with timing of the survey which was doing the pandemic which might have influenced the findings of this study.

Similarly, Kameswaran and Muralidhar [8] aimed to investigate the accessibility of cash and digital payment apps for people with visual impairments while using ride sharing services in India. For this, they conducted 30 interviews with people with visual impairments in metropolitan India. From their analysis, they found that neither cash nor digital payment apps are accessible to these people and they need to perform additional work to make these modes of payment more accessible. Lastly, they advise policymakers to make the digital payment apps more accessible and inclusive such that people from all walks of life use them properly.

Likewise, Pal et al. [25] aimed to understand the factors that influence the actual usage and future use intention of mobile payment technology in India. To do so, they conducted a survey across India and got 551 responses. Using the valence framework and the technology affordances and constraints theory (TACT), they found that convenience, reflection, security and awareness play a significant role in shaping usage of mobile payment apps. Lastly, they identified single country context and relying on self reporting data as the limitations of their study and suggested future research on other types of factors such as social and cultural factors that may influence mobile payment app usage.

Now, let's dive into the literature related to survey data analysis. Firstly, Natuhamya [35] talks about applying weights during work with survey data. It proposes a technique to calculate these weights to make sure the estimations are genuinely correct. This assists in forming better judgments from the facts we have. The article also provides a table and references to other research publications on survey data analysis.The article doesn't discuss about any drawbacks or negatives of the strategy they recommend. But it's probable that there are some flaws with the procedure that they didn't address in the text.

This study, by Alsayat [34] looks at how online writings by users, such as reviews and comments, could bring insight into the desires and satisfaction levels of tourists in Mecca, Saudi Arabia. The researchers examined several social media postings and reviews of Mecca hotels using computer systems to determine this. They also looked at the results of comparable studies conducted by other researchers. They gained knowledge from doing this that may be applied by hotels and others in the tourist sector to increase customer satisfaction and assist travelers in selecting the ideal accommodation in Mecca.

Moreover, Holt, Scott and Ewings [1] speak about performing a statistical test called chi-squared with survey data. They argue that when you employ this test with survey data, it might occasionally give you the erroneous findings. Instead, they recommend utilizing new methodologies that address the specific manner survey data is obtained. To prove this, they looked at findings from two huge nationwide polls. What they observed is that the traditional chi-squared test can make errors when employed with survey data. They advocate adopting an alternative test which performs better with this sort of data.

Lastly, The focus of Ko, Kim and Lee [23] is on people's intentions to utilize shared mobility services is included in this Paper. It examines the variables that affect this intention in both current and potential customers of these services. The characteristics of the respondents, who were polled in Gyeonggi Province, Korea, illustrate the many types of commuters in that country. They discovered which elements—such as education—really important when utilizing shared transportation options. The National Research Foundation of Korea and the KAIST-KU joint research center supported the authors' work, and there were no conflicts of interest. However, due to Gyeonggi-do's data security regulations, they were unable to release the data they utilized.

Chapter 3

Background

This section of the study provide detailed theories of the established model and techniques used to perform the comprehensive data analysis.

3.1 Chi Square Test of Independence:

The Chi-Square Test of Independence is a statistical test used to determine whether there is a significant association between two categorical variables. It is based on the chi-square statistic, which measures the difference between the expected and observed frequencies in a contingency table.

Below are the key concepts and steps involved in the Chi-Square Test of Independence:

Contingency Table: The data is organized into a contingency table (also known as a cross-tabulation or two-way table). The table displays the frequency distribution of two categorical variables. Each variable has its categories, and the intersections represent the joint frequencies.

Hypotheses:
1. Null Hypothesis (H_0): There is no significant association between the two categorical variables.
2. Alternative Hypothesis (H_1): There is a significant association between the two categorical variables.

Chi-Square Statistic
The chi-square statistic is calculated using the formula:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

where O_i is the observed frequency in cell i and E_i is the expected frequency in cell i.

Degrees of Freedom
The degrees of freedom (df) for a Chi-Square Test of Independence is determined by:

$$df = (r - 1) \times (c - 1)$$

where r is the number of rows and c is the number of columns in the contingency table.

Critical Value or P-Value: The chi-square statistic is compared to a critical value from the chi-square distribution with d_f degrees of freedom. Alternatively, the p-value can be calculated. A small p-value (< 0.05) suggests rejecting the null hypothesis.

Decision: If the chi-square statistic is greater than the critical value or the p-value is less than the significance level (e.g. 0.05), then the null hypothesis is rejected. Rejection implies evidence of a significant association between the two categorical variables.

3.2 Principal Component Analysis:

Principal Component Analysis (PCA) is a popular unsupervised machine learning model used for dimensionality reduction by creating principal components. It is used for finding patterns in data of high dimension. It expresses the data in such a way as to highlight their similarities and differences.

Covariance Matrix: Covariance measures the degree to which two variables change together.It is a measure of how much two dimensions vary from the mean with respect to each other.

Variance (Var(X))
For a random variable X, the variance is given by the formula:

$$\text{Var}(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Where:

- n is the number of observations.
- x_i is the i-th observation.
- $\bar{x}$ is the mean of the observations ($\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$).

Covariance (Cov(X, Y))
For two random variables X and Y, the covariance is given by the formula:

$$\text{Cov}(X, Y) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

Where:

- n is the number of observations.
- x_i and y_i are the i-th observations for variables X and Y.
- $\bar{x}$ and $\bar{y}$ are the means of X and Y ($\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ and $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$).

Eigenvalue Decomposition: Obtains the eigenvectors and eigenvalues of the covariance matrix.

Selection of Principal Components: Sort eigenvectors by eigenvalues and choose the top k components.

Projection: Transform the data using the selected principal components.

3.3 Decision Tree

A decision tree is a supervised machine learning algorithm used for both classification and regression tasks. It is a tree-like structure where each internal node represents a decision based on the value of a particular feature, each branch represents the outcome of that decision, and each leaf node represents the final decision or the target value.

Feature Importance in Decision Trees: Decision trees naturally provide a measure of feature importance based on how often a feature is used to split the data across all decision nodes. The importance of a feature is calculated by considering:

- **Gini Importance (for Classification):** For each feature, the improvement in Gini impurity is calculated over all the nodes where the feature is used for splitting. The more a feature reduces impurity, the more important it is.

  $$\text{Gini Importance} = \sum_{i=1}^{n} \frac{N_i}{N} \times \text{Gini}(i) \tag{3.1}$$

  where N_i is the number of samples at node i, N is the total number of samples, and Gini(i) is the Gini impurity at node i.

- **Feature Importance (for Regression):** For regression tasks, the feature importance is often computed based on the average improvement in the mean squared error that the feature brings across all nodes where it is used for splitting.

$$\text{Importance} = \frac{\text{Improvement in MSE}}{\text{Number of Nodes Splitting on the Feature}} \tag{3.2}$$

Features with higher importance values have a greater influence on the decisions made by the decision tree.

3.4 Random Forest

A Random Forest is an ensemble learning method that combines multiple decision trees to create a more robust and accurate model. Each tree in the forest is trained on a random subset of the training data and makes decisions independently. The final prediction is then determined by aggregating the predictions of all individual trees.

Feature Importance in Random Forest: Random Forests naturally provide a measure of feature importance based on how much each feature contributes to the accuracy of the trees in the forest. The importance of a feature is calculated by considering:

- **Mean Decrease in Impurity (for Classification):** For each tree, the impurity reduction (such as Gini impurity) provided by each feature is averaged over all trees. The higher the average impurity reduction, the more important the feature.

$$\text{Gini Importance} = \frac{1}{T}\sum_{t=1}^{T}\sum_{\text{features}}\frac{\text{Impurity Reduction}_{\text{feature}}}{\text{Number of Trees}} \tag{3.3}$$

- **Mean Decrease in Accuracy (for Regression):** For each tree, the decrease in accuracy (or increase in mean squared error) caused by each feature is averaged over all trees. The higher the average decrease in accuracy, the more important the feature.

$$\text{Importance} = \frac{1}{T}\sum_{t=1}^{T}\sum_{\text{features}}\frac{\text{Decrease in Accuracy}_{\text{feature}}}{\text{Number of Trees}} \tag{3.4}$$

  where T is the total number of trees for both cases in the Random Forest. Features with higher importance values have a greater influence on the overall predictive performance of the Random Forest.

3.5 Light Gradient Boosting Machine

LGBM (Light Gradient Boosting Machine) is a gradient boosting framework that uses tree-based learning algorithms. It is designed for speed and efficiency and is particularly effective in handling large datasets. LGBM builds trees in a level-wise manner, choosing the leaf with the maximum delta loss to grow at each level.

Feature Importance in LGBM: LGBM provides feature importance scores based on how often a feature is used in tree nodes and how much it contributes to the reduction in the objective function (such as mean squared error for regression or log loss for classification).

Calculating Feature Importance:

- **Split Gain (Reduction in Loss):** For each feature and each split in a tree, LGBM calculates the reduction in the objective function (e.g., mean squared error reduction for regression or log loss reduction for classification). The split gains across all splits for a specific feature are aggregated.
- **Feature Importance Score:** The feature importance score is calculated based on the sum of split gains over all the trees in the ensemble. It provides a measure of the overall contribution of each feature to the model's performance.

LGBM's feature importance is computed based on the actual improvement each feature brings to the model during training.

3.6 K-Means Clustering

K-Means clustering is an unsupervised machine learning algorithm used for partitioning a dataset into K distinct, non-overlapping subsets (clusters). The goal is to group similar data points together while keeping dissimilar points in different clusters.

Steps of K-Means Clustering

Initialization:

- Choosing the number of clusters K and randomly initialize K centroids. These centroids can be randomly selected data points or placed in a way that represents some knowledge about the data.

Assignment Step:

- Assigning each data point to the nearest centroid, forming K clusters.

Update Step:

- Recalculating the centroids of the clusters based on the mean of the data points assigned to each cluster.

Repeat:

- Repeating steps 2 and 3 until convergence, which occurs when the centroids no longer change significantly or a predefined number of iterations is reached.

Objective Function (Inertia): Let C_k represent the set of data points in cluster k, and μ_k represent the centroid of cluster k. The inertia (objective function) is given by:

$$\text{Inertia} = \sum_{k=1}^{K} \sum_{i \in C_k} \|x_i - \mu_k\|^2$$

where $\|x_i - \mu_k\|^2$ denotes the Euclidean distance.

3.7 TextBlob

TextBlob is a famous Python library used for text data processing. It provides an easy to use API for basic Natural Language Processing (NLP) tasks, such as sentiment analysis, part of speech tagging, translation etc. It is built on top of NLTK (Natural Language Toolkit) and Pattern.

3.8 Cronbach's Alpha

Cronbach's Alpha is a measure of internal consistency and reliability of a set of items in a survey questionnaire. It checks how closely related a set of items are as a group. The formula for Cronbach's Alpha is:
The formula is given by:

$$\alpha = \frac{k-1}{k} \left(1 - \frac{\sigma_T^2}{\sum \sigma_i^2} \right) \tag{3.5}$$

where:

k is the number of items,
σ_i^2 is the variance of the scores on the i-th item,
σ_T^2 is the variance of the total scores.

Chapter 4

Research Methodology

In any research project, the methodology is the heart and soul of the research as it sets the flow of the entire research. In our case, our comprehensive approach is displayed through the following diagram below along with detailed step by step explanation of the key elements of our detailed approach.

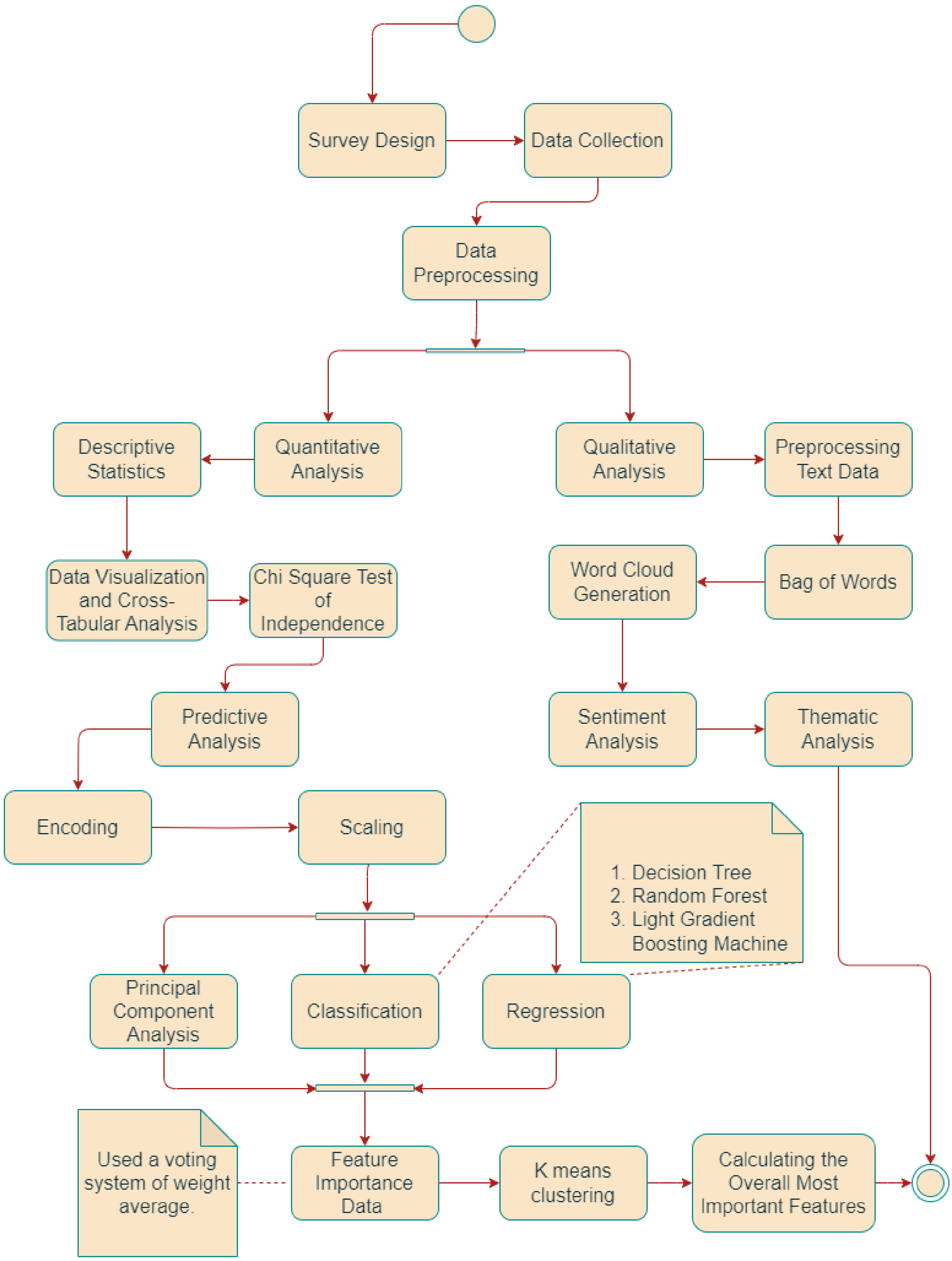

Figure 4.1: Workflow of Our Proposed Approach

4.1 Survey Design and Data Collection

The aim of this survey is to understand consumer purchase behavior and user perceptions related to digital payment apps. We conducted this survey using a google form by snowball sampling as it enables us to acquire information from hard-to-reach communities. We reached roughly 60 to 70 people then by using snowball sampling effect, we got 170 samples, here snowball sampling raises the sample size over time owing to referrals from current participants, assuring inclusion and representation [33].

Starting with a small initial sample of participants, snowball sampling operates by relying on those individuals to refer to other participants [33]. Iteratively, this process continues until a big enough sample size is reached, like a snowball rolling downhill. Snowball sampling starts with a small number of "seed" individuals to gather data. Requesting that they suggest additional individuals who meet the requirements for the study. Expands the sample through recommendations. Continues until the research's goals are satisfied or the data are saturated. It is especially helpful for researching populations that are challenging to reach using conventional sample techniques.

There are 5 sections in the survey, and each section has certain questions that must be answered. Let's discuss these sections and what they represent.

Section 1:
In this section, questions related to demography (age, gender, employment status and monthly income) are asked. Moreover, the usage status of the respondent is asked. Lastly, we directed the users to section 2 and the non user to section 3.

Section 2:
This section is structured with questions related to the psychology of the user. Here, we asked regarding the usage and satisfaction level across 10 apps using 2 multiple-choice grid questions with the rate level 0 signifying "never used" and the rate level 5 signifying "only used". Then we asked about :

- **Frequency of usage of digital payment apps** (MCQ)
- Why do they use them? (Checkbox)
- Why do they prefer the app(s) they use? (Checkbox)
- What type of transactions do they perform with these apps? (Checkbox)
- What are their preference criteria while choosing a particular app? (Checkbox)
- Thoughts and opinions. (Open ended)
- Problems faced and Suggestions. (Open ended)

After this section, users are directed to section 4.

Section 3:
This section is dedicated to non-users. Here, we asked for their reasons not to use digital payment apps and what may encourage them to adopt these apps through MCQs. After this section, the non users are directed to section 5.

Section 4:
In this section, we focused on user's consumer behavior and the effects of digital payment app usage. At first, we asked whether they consider digital payment apps the **most convenient option** of payment. Then, using linear scale questions, we asked:

- **Impact on financial management habits**
- **Influence of promotions and discounts while using these apps**
- **Confidence on security of these apps**
- **Likelihood of recommending to use digital payment while making a purchase**

After this section, the non users are directed to section 5.

Section 5:
In this section, we focused on the consumer behavior of the entire market as these questions are applicable for both users and non users. These are:

- **What mode of transaction do they prefer?** (MCQ)
- Concerns regarding digital payment apps. (MCQ)
- **Does usage of these apps increase impulse purchase?** (MCQ)
- **Does usage of these apps increase online spending?** (MCQ)
- Perception of the industry as a whole. (Open Ended)

The bolded questions are used as a metric to quantify consumer behavior and the rest of the question provides us with information regarding consumer demographics and psychology.

4.2 Dataset

How The Survey was Conducted and Designed

This survey explores the impact of digital payment app usage on consumer behavior focusing on purchase patterns and user perceptions. Conducted from July 28 to November 20, 2023, through snowball sampling we reached 170 participants. Snowball sampling, starting with a small seed group and expanding through referrals, was crucial for capturing insights from hard-to-reach communities [33]. This method allowed us to tap into diverse viewpoints and experiences related to digital payment apps.

The survey comprises five sections, each with specific questions. Before categorical inquiries, respondents provide basic information on age, gender, employment, monthly income, and app usage.

Participants from diverse professions contribute valuable insights, allowing us to understand digital payment app usage patterns. Their feedback on satisfaction, frequency of use, preferred apps, transaction types, and criteria aids in studying consumer behavior and financial decisions. We assure complete anonymity and data confidentiality, utilizing personal information solely for research purposes.

Observations: This section presents the comprehensive results of the survey, which aimed to study the impact of digital payment app usage on user behavior. The questionnaire covers demographic information and specifics about digital payment app usage.

1. Observations From Section One

In this section we mainly gathered the primary data of respondent before classifying into user part and non-user part.

Age : This section categorizes respondents by age to explore how age influences digital payment app usage and financial behavior. It assesses potential differences in preferences and behaviors between younger and older individuals.

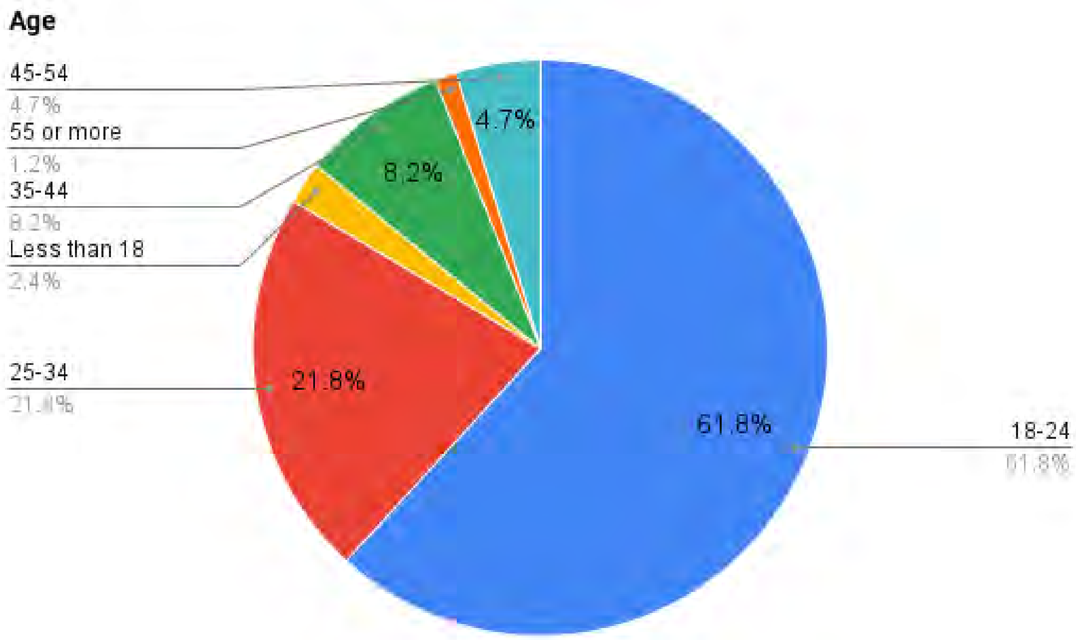

Figure 4.2: Age of Respondents

This graphic presents a pie chart showing age distribution among 170 respondents using a snowball sampling approach. The majority 61.8% are aged 18-24, followed by 21.8% aged 25-34. The survey had limited reach among those aged 55 or more, indicating a higher interest and responsiveness among younger participants (18-24 and 25-34).

Gender: Gender data is vital for studying how digital payment app usage influences financial behavior. It helps adapt services to diverse user needs and addresses gender-based disparities, promoting financial inclusion and equality.

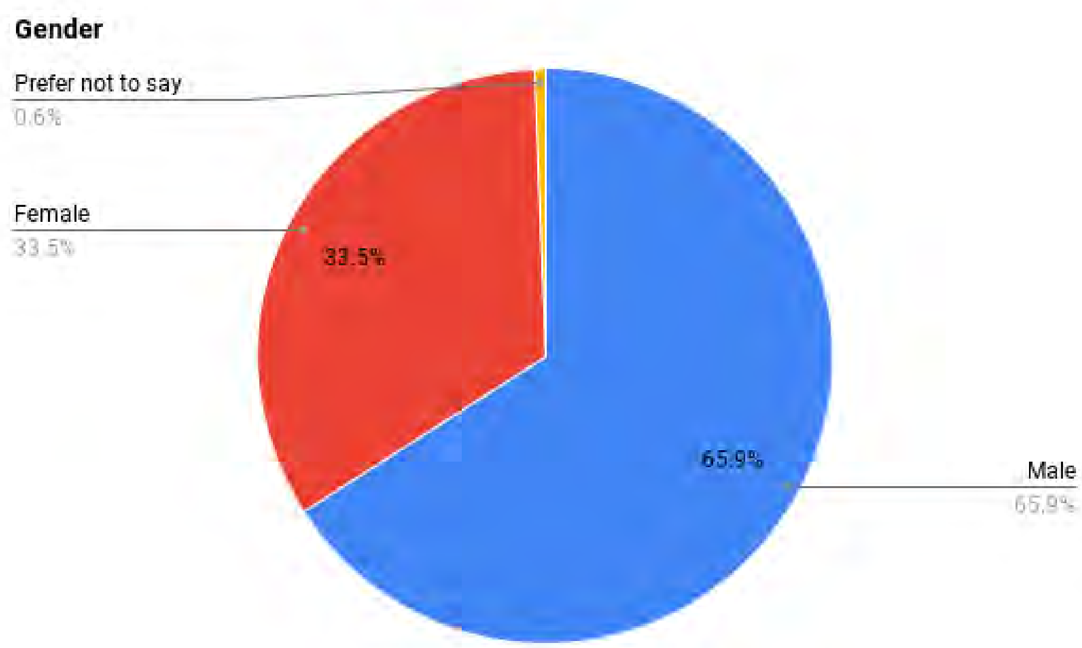

Figure 4.3: Gender of Respondents

The pie chart displays gender distribution among respondents, with 65.9% males (112) and 33.5% females (57). One respondent (0.6%) preferred not to disclose their gender. The higher number of male respondents suggests a potential correlation with increased digital payment app usage.

Employment Status: This survey explores how job types relate to digital payment use and impact financial behavior. Data helps tailor financial tools for diverse employment categories like freelancers managing irregular income and full-time workers handling routine bills and long-term planning. Insights lead to more efficient and specialized financial services for each work situation.

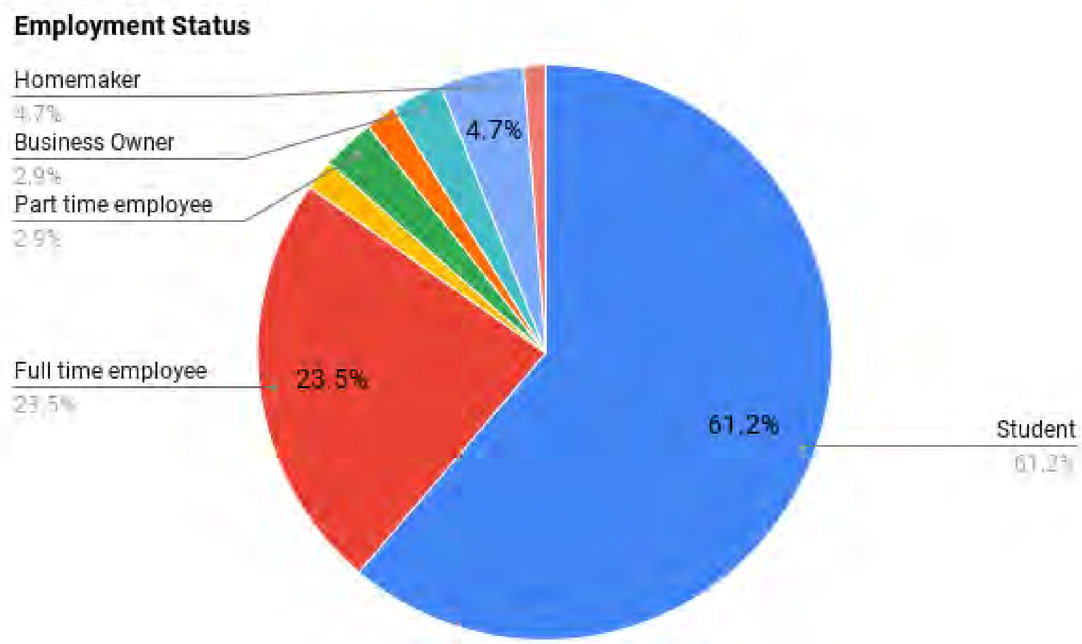

Figure 4.4: Employment Status of Respondents

The pie chart reveals that 61.2% (104 out of 170) of respondents are students, followed by 23.5% (40) as full-time employees. Part-time employees account for 2.9% (5), slightly more than homemakers. Other categories include unemployed (1.8%, 3), business owners (2.9%, 5), freelancers (1.8%, 3), and retired individuals (1.2%, 2). Notably, there are no self-employed participants. The survey, conducted through snowball sampling, involves individuals from eight different professions emphasizing diverse perspectives on digital payment applications.

Monthly Income: Analyzing respondents' monthly incomes is key to understanding the link between financial behavior and digital payment app usage, offering insights into patterns across socioeconomic groups and how these apps aid financial management at different income levels.

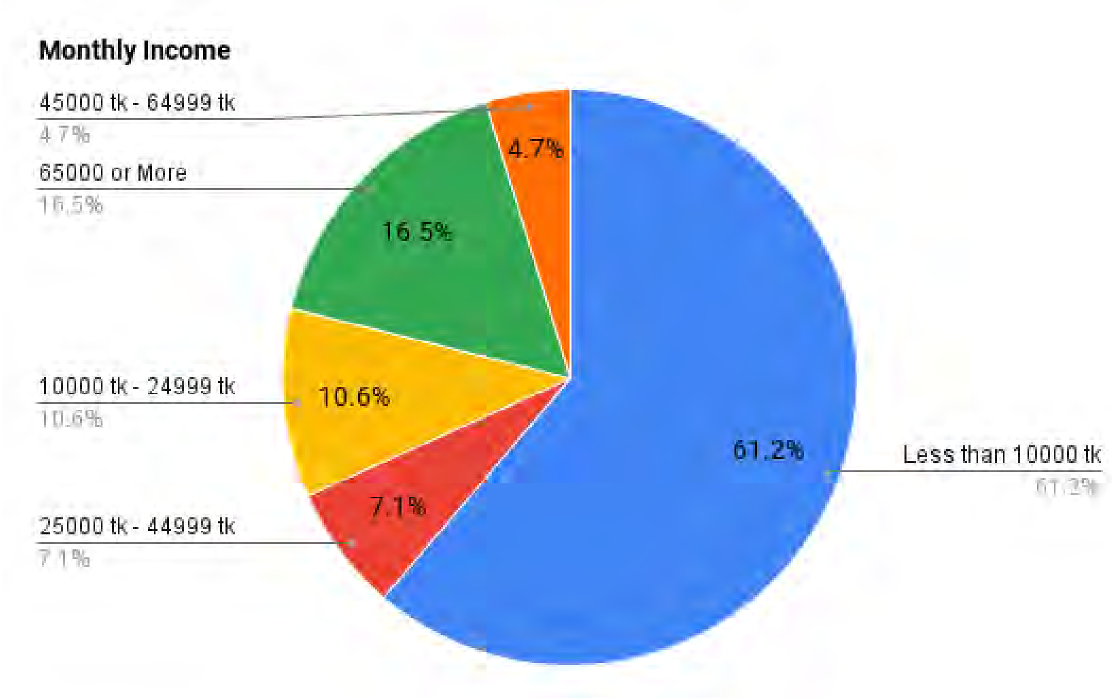

Figure 4.5: Monthly Income of Respondents

The pie chart illustrates that 61.2% of respondents earn less than 10,000 taka monthly, indicating a majority with lower incomes. 10.6% earn between 10,000 to 24,999 taka, 7.1% between 25,000 to 44,999 taka, and 4.7% between 45,000 to 64,999 taka. Surprisingly, 16.5% earn 65,000 taka or more monthly. This categorizes respondents into three income groups: low, medium, and high. Regardless of income, individuals show awareness of digital payment apps, reflecting societal influences. The data highlights widespread exposure and awareness of digital payment apps in society.

If Responders are Current Users or Not: This question determines if respondents use digital payment apps, enabling further exploration of how usage impacts financial behavior.

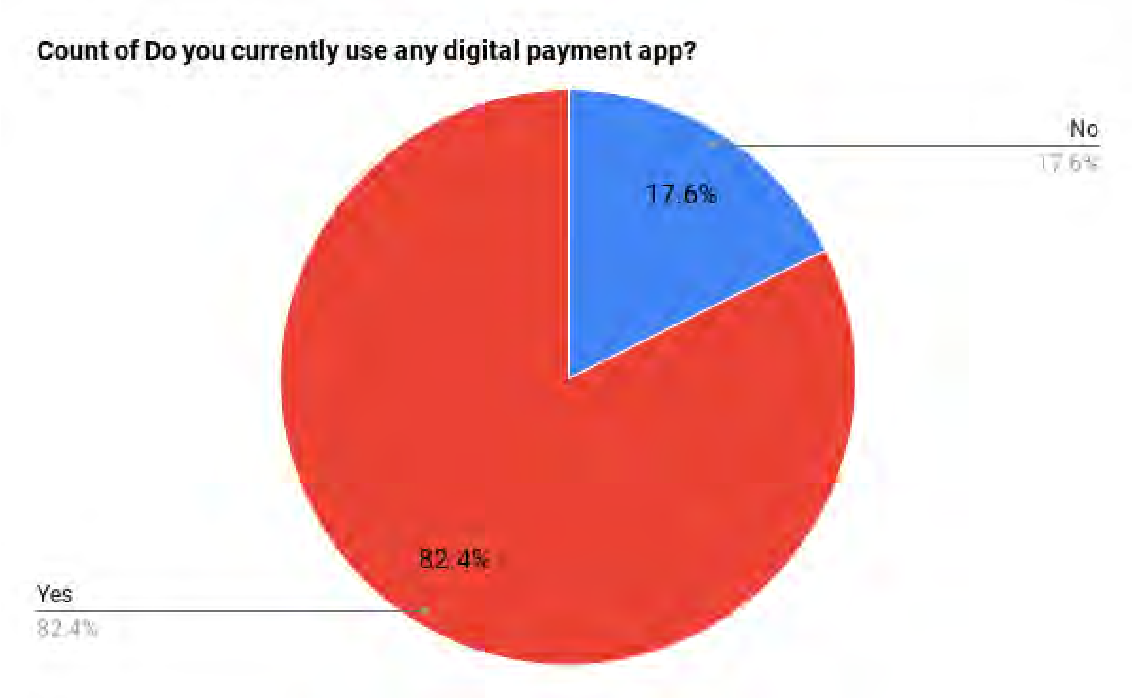

Figure 4.6: Survey Response to the Following Query “Do you now use any digital payment app?”

In the pie chart with 170 participants, 82.4% use digital payment apps, while 17.6% do not. The popularity of digital payments is attributed to their convenience, speed, and global shift towards cashless transactions, accelerated by the COVID-19 pandemic. The 17.6% not using these apps may have concerns about privacy, security, limited access to technology, or a preference for traditional payment methods.

2. Section Two Observations Analysis

Mainly, this section is designed fully based on the individuals who use digital payment applications and based on their demands and activities.

Digital Payment App Usage: Respondents rated their digital payment app usage from 0 to 5, with 5 indicating exclusive use and 0 indicating no use. Bkash emerged as the most widely used app, followed by Nagad, NexusPay, and OK Wallet. Upay, Rocket, MCash, SureCash, Trust Axiata Pay, MyCash, and OK Wallet had lower usage. Notably, Bkash had the highest usage, while MCash and SureCash were the least used.

User Satisfaction: A satisfaction survey measured user feedback on digital payment apps. About 83% of respondents used Bkash, with 60 completely satisfied users. Nagad was the second-highest rated app, while satisfaction for other apps, except Bkash, was generally low. Nagad and Rocket had similar, less satisfied user numbers. Overall, satisfaction levels for apps other than Bkash consistently declined.

Frequency of Usage: The study aimed to understand how often respondents use digital payment apps in their daily lives. Answers from 140 participants revealed that 45% use digital payment apps once or twice a week, with 32.1% using them daily. Only 12.1% use them occasionally, with 10% using them once or twice a week. The majority of respondents use digital payment apps frequently in their daily lives.

Reasons for Usage: Respondents cited reasons for using digital payment apps. Most respondents (118 out of 140) use digital payment apps for convenience, with additional reasons including speed (53.6%) and contactless payments (54.3%). Some also cited security, rewards, credit access, mobile recharge, and subscription purchases. Overall, consumers employ digital payment apps for their convenience and diverse functionalities.

Preferences for Specific Apps: Respondents were asked why they prefer the particular digital payment app(s) they use. Most significant factors included safety, ease of use, popularity, availability, brand recognition, and integration with other apps. "Easy to use" emerged as the most significant factor, highlighting the importance of user-friendly design for accessibility and convenience.

Types of Transactions: The question explored the range of financial activities facilitated by digital payment apps. Most respondents (113 out of 140) used digital payment apps for online purchases, and 111 used them for money transfers, highlighting the convenience of digital transactions. Users also engaged in various activities like donations, mobile recharge, in-store purchases, subscriptions, and peer-to-peer transactions.

Criteria for Choosing Apps: When choosing digital payment apps most respondents prioritized ease of use (105 out of 140). Other important factors included lower fees, better security, integration with other apps, customer service, rewards, features, and speed.

These observations provide a comprehensive understanding of the usage patterns, preferences, and satisfaction levels among respondents offering valuable insights for app developers and companies in the digital payment app landscape.

3. Observations from section three

This section focuses on non-user responses exploring the reasons behind not using digital payment apps and identifying features that could be improved to encourage their adoption.

The reason behind not using any digital payment app: Individuals may choose not to use digital payment apps for various reasons, including personal preferences, access limitations, trust issues, and cultural or social concerns. Among 30 non-users, 19 provided reasons for not using digital payment apps. Main reasons included no need and a preference for other payment methods. Some cited a lack of trust, feeling unsafe, and not knowing how to use digital payment apps. Notably, none mentioned high fees as a concern.

Improvements, which non-users think should happen: To drive digital payment app usage, addressing user needs, preferences, and concerns is crucial and requires continuous improvement based on feedback. Respondents shared preferences for changes in digital payment apps each improvement accompanied by the number of respondents voicing that opinion. Preferences included lower fees (16.7%), improved safety features (23.3%), faster transaction processing (13.3%), better customer support and response time (16.7%), an increase in acceptance among merchants and businesses (10%), a user-friendly interface and navigation (36.7%), better rewards and incentives (26.7%), and greater integration with more apps and services (16.7%). Additionally, 40% of 30 respondents stated that they are not interested in using these apps in the near future. Respondents suggested improving the user-friendly interface and navigation to encourage future adoption.

4. Observations From Section Four

Observations from Section Four provide valuable insights into users' perspectives on digital payment apps.

Convenience Comparison with Traditional Payment Methods: In terms of convenience compared to traditional payment methods, the majority (75.7%) of respondents view digital payment apps as more convenient, with 19.3% considering them equally convenient and 5% finding them less so. These diverse opinions reflect individual experiences and preferences.

Impact of Digital Payment Apps on Financial Management Habits: Regarding the impact of digital payment apps on financial management habits, responses vary. A substantial 89.3% of respondents report a fair impact, while 52.1% note a significant impact. This variability is influenced by individual financial behaviors and app-specific features.

Influence of Promotional Offers on Purchase Decisions: Promotional offers play a notable role in influencing users' purchase decisions when utilizing digital payment apps. Approximately 37.9% of respondents indicate a fair influence, while 40% report a significant impact. This highlights the importance of promotions in shaping user behavior.

Confidence in Security Measures of Digital Payment Apps: In terms of confidence in the security measures of digital payment apps, 38.6% express high confidence, 15.7% have the highest confidence level, 28.6% report a medium level of confidence, and 17.1% have low confidence. This distribution reflects users' varying degrees of trust in the security measures implemented.

Likelihood to Recommend Digital Payment Apps: The likelihood of recommending digital payment apps is notably high, with 88.6% of respondents expressing a strong likelihood to recommend. Additionally, 18.6% would fairly recommend, while only 9.3% do not recommend these apps. This positive recommendation trend underscores overall user satisfaction.
These observations collectively portray a positive sentiment toward digital payment

apps, emphasizing their convenience, impact on financial habits, influence of promotional offers, security measures, and high likelihood of user recommendations.

5. Observations From Section Five

Observations from Section Five delve into users' overall experiences and opinions on digital payment apps.

Regarding the preferred mode of financial transactions, the survey reveals that 41.2% of respondents favor digital payment apps, while 38.2% prefer cash. Preferences for credit cards, debit cards, cheques, and other modes vary based on situational needs and individual choices.

Addressing concerns about digital payment apps, respondents primarily express worries about security and privacy (118 out of 170), followed by concerns about fraud and scams (103), technical issues (82), accessibility and compatibility (69), and trustworthiness (46). Security and privacy emerge as the most significant concerns among users.

Exploring the impact of digital payment apps on impulsive purchases, 63.3% of respondents agree that these apps contribute to increased impulsive buying. Only 8.9% disagree with this statement, emphasizing the potential influence of digital payment apps on users' spending behavior.

Assessing the belief that digital payment apps can elevate online spending, a majority of respondents (71.6%) agree with this statement. About 15.4% roughly agree, while 13% do not concur. This finding indicates a prevailing perception that the use of digital payment apps has the potential to boost online expenditures.

In summary, users express a diverse range of preferences and concerns regarding digital payment apps. While some favor these apps for financial transactions, concerns about security and privacy remain paramount. Additionally, there is a notable acknowledgment of the potential influence of digital payment apps on impulsive purchases and online spending among users.

Overall App Usage and Satisfaction: The survey shows many people use digital payment apps especially Bkash with 82.4% users. People like the apps for their convenience and speed leading to high satisfaction (82.4%). Respondents form a big part (61.2%) of users showing diverse income backgrounds. Concerns about security (69.4%) highlight the need for strong measures but confidence is still high (70.7%). Positive promotions (56.4%) and high recommendations (88.6%) indicate the apps play a vital role. There are various impacts on financial habits. To improve, focusing on security and personalized incentives could enhance user experiences and app adoption.

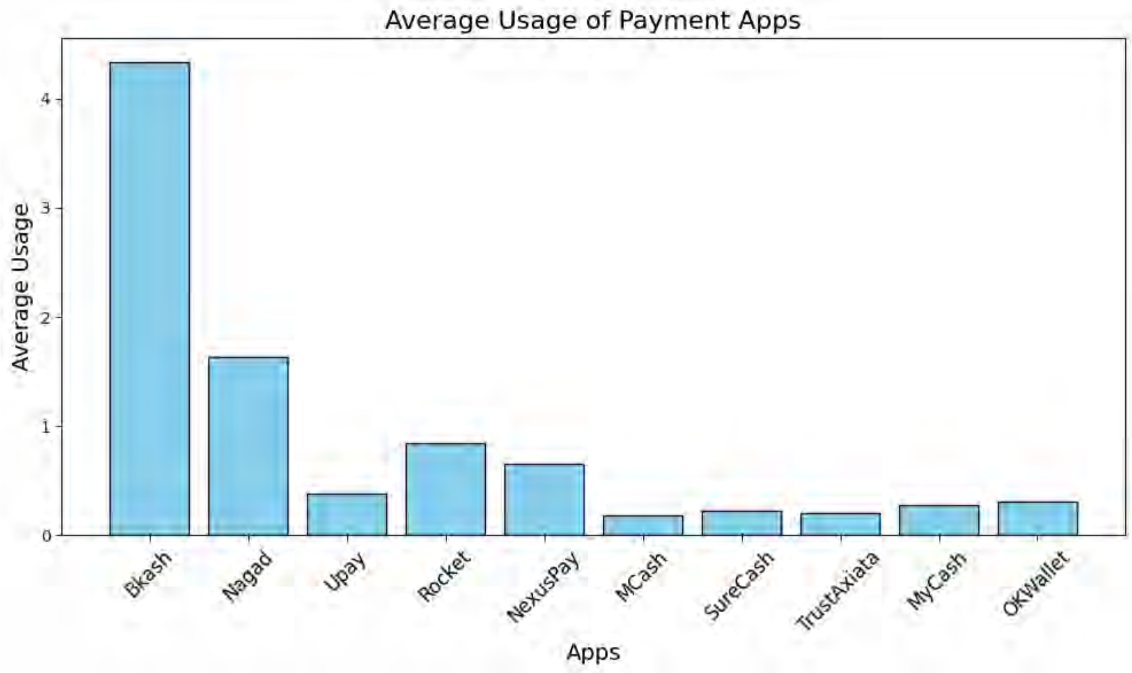

Figure 4.7: Average Usage of Payment Apps

In the app usage graph, shows the average usage of each app and the app satisfaction graph, the average satisfaction is shown. For the app satisfaction app, the 0 responses don't count since they represent that the person never used that particular app rather than their satisfaction level.

One of the observations is Bkash is the most popular app with an average usage of 3.0. Nagad is the second most popular app, with an average usage of 2.5. Basically, the graph shows that Bkash is the most popular payment app in Bangladesh. Other popular apps include Nagad and Upay.

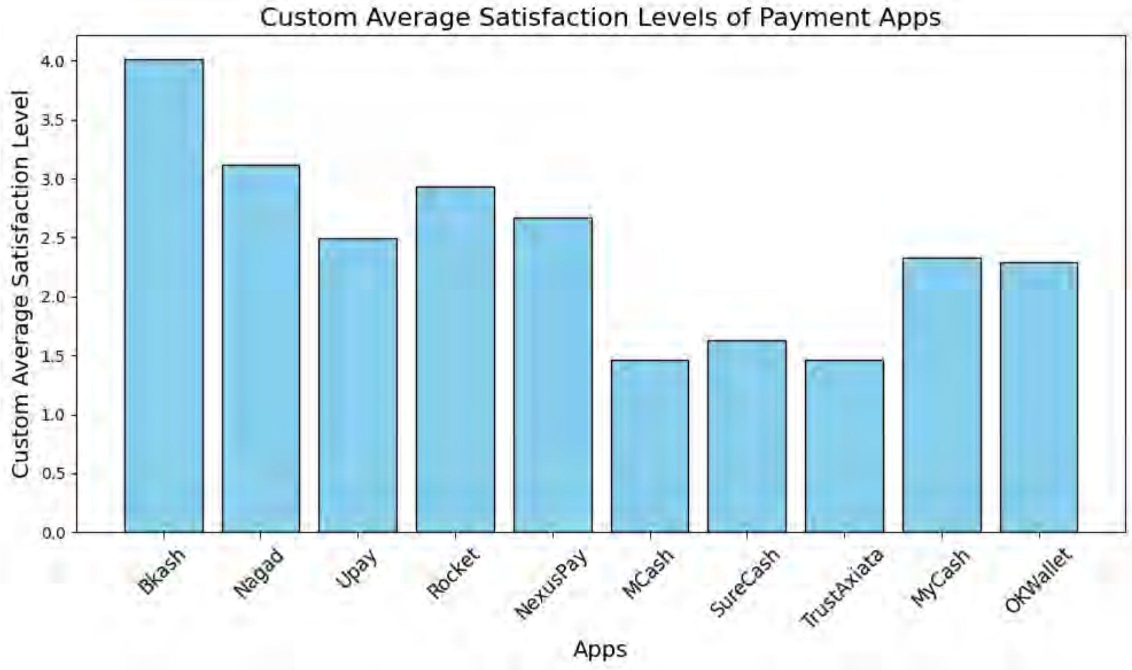

Figure 4.8: Customer Average Satisfaction Level of Payment Apps

Here, We can see the most satisfying payment app in Bangladesh is Bkash, with an average satisfaction level of 3.0. Nagad is the second most satisfying app, with an average satisfaction level of 2.5. Upay is the third most satisfying app, with an average satisfaction level of 2.0. Rocket, NexusPay, MCash, SureCash, TrustAxiata, MyCash, and OKWallet are all less satisfying, with average satisfaction levels of less than 2.0.

After doing this study the overall observation is Bkash and Nagad are the most used apps in Bangladesh so far and they are currently leading the market according to our data.

4.3 Data Preprocessing

As we are dealing with real world data, it is bound to have issues that are essential to handle properly to perform any further analysis. Let's discuss the steps taken to preprocess the data for a better and consistent analysis.

Shortening of Initial Column Names: At first, we shortened the initial column names since the initial data columns were the questions asked in the survey and were too long to work with. The simplified old names and their corresponding new names are listed below using the following table. As there are a lot of columns and most of them are renamed using intuitive names only the names that may seem a bit confusing are listed below:

Table 4.1: Shortening of Initial Column Names

New Names	Simplified Old Names
Usage_AppName	Usage status of a user for a particular app
Satisfaction_AppName	Satisfaction of a user for a particular app
FrequencyUsage	Frequency of usage of digital payment apps
ReasonUse	Why do they use them?
ReasonPreference	Why do they prefer the app(s) they use?
TransactionTypes	What type of transactions do they perform with these apps?
PreferenceCriteria	What are their preference criteria while choosing a particular app?
NoUseReason	Why don't they use them?
Encouragement	What may encourage them to use these apps?
TransactionPreference	What mode of transaction do they prefer?
ImpulsePurchases	Does usage of these apps increase impulse purchase?
OnlineSpending	Does usage of these apps increase online spending?

Handling Null Values: While observing the dataset, we found cases where some responses are labeled as null in the open ended questions. In this case, we imputed "No Comments". In case yes/no/maybe questions like ImpulsePurchases and OnlineSpending, we found 1 null value which we imputed to "Maybe" which was one of the options.

Dividing the dataset: We divided the dataset into 2 parts. One contained user data the other one contained non user data. Then we dropped columns that were not supposed to be in that dataset from each dataframe.

Handling Inconsistency: The raw dataset contained some inconsistencies. For example, in the usage question, some respondents claimed that they use an app whether as in the satisfaction question they claimed that they never used it and vice versa. To handle this, we assumed that the usage information was correct and then imputed 0 for non users of that app or the mean of user satisfaction for users of that app.

Handling Checkboxes: As our survey questionnaire contains a lot of checkbox type questions, it was essential to handle them properly. In our case, we used one hot encoding and added new columns where 0 means didn't choose and 1 means chose. Then we again shortened the column names with the following convention: Question_Option. For instance, if some chose the reason for using digital payment apps to be speed, the ReasonUse_Speed value is set to 1.

Changing Data Type: Apart from the open ended questions, usage status and satisfaction status, all the other columns are categorical in nature, we changed their data type to categorical.

Now as overall data preprocessing is done, it's time to start the analysis. In our case, the analysis is divided into 2 parts: quantitative and qualitative.

4.4 Qualitative Analysis

Here, we take the data from the open ended questions (thoughts and opinions, problems, suggestions and industry perception) and create a separate data frame. Now we will perform data analysis into 2 phrases: Preliminary analysis using NLP and Thematic analysis.

4.4.1 Preliminary Analysis

The intent of this section is to perform a basic sentiment analysis to understand the attitude and perception of consumers through the thoughts and opinions column and industry perception column. Also this section helps us to look into the data before actually doing thematic analysis and give a rough idea of the dataset.

Text Data Preprocessing: Before anything, we must preprocess our data. For this, we vectorized the data, removed all punctuation as well as stop words and also converted every character to lowercase.

Bag of Words and Word Clouds: Now we took the preprocessed data and performed a bag of words and created word clouds to generate a beautiful and intuitive way to look into the qualitative data.

Sentiment Analysis: Here, we used Textblob, a popular lexical based python library used for sentiment analysis. As sentiment analysis itself is a broad field of research, doing it more scratch can be very tricky and time consuming, distracting us from our main agenda. To avoid this, we used this method instead. Results of this analysis represented consumer attitude and perception towards digital payment apps in the later part of the analysis.

4.4.2 Thematic Analysis

In this section, we discuss the approach taken for thematic analysis of the qualitative data we receive from our survey. The intent of using thematic analysis was to dive deep into the attitude and perception of the consumer as well as to find the problems they face and the suggestions they have. This gives a more in depth understanding of consumer psychology and the impact the digital payment industry has made in the Bangladesh market.

Data Preparation and Initial Coding: To facilitate in-depth analysis, we adopted a collaborative approach. 3 of our team members were assigned a specific open-ended question for initial coding. Following careful skimming through the responses, individual thematic coding was undertaken. This initial coding phase employed an inductive approach, allowing themes to emerge organically from the data itself. This approach helped us to capture the essence of each participant's response properly.

Theme Generation: Following initial coding, a meeting was conducted among the individual coders. During this collaborative session, the individually generated codes were presented, compared, and discussed. Then through iterative dialogue and critical reflection, consensus was reached on the merging and refining of codes, leading to the identification of distinct themes. This process ensured the maximization of the richness of insights analysis through collaboration.

Thematic Reporting and Interpretation: The identified themes were subsequently named and defined, ensuring thematic clarity and coherence. Then, corresponding responses were then used to further elaborate on each theme, offering concrete evidence of thematic validity. Finally, the themes were interpreted within the context of our research objectives, providing valuable insights on consumer behavior.

4.5 Quantitative Analysis

In this phase of the analysis, we have considered all the quantitative data received from the close ended questions and from the preliminary qualitative analysis. At first, we also discretized the linear scale data from the consumer behavior section. This helped us to improve interpretability and helped us visualize our data properly. Additionally, we grouped similar classes from demographic features to reduce the classes with low number of instances. For example, we have grouped all types of employees to a single class named "Employee". Then, we used descriptive statistics such as mean, mode and median and with cross tabular analysis among demographic, psychological, user consumer behavior and market consumer behavior data to look deeper into the quantitative data. We used data visualization to aid our analysis and generated 100+ graphs and charts each giving us a detailed description of the corresponding part of data. The next part of our analysis is divided into 2 parts, inferential statistics using chi square test and predictive analysis using machine learning model.

4.5.1 Inferential Statistics

In this part of the analysis, we performed a chi square test of independence between the feature columns (demographic and psychological factors) and the output columns (metrics to quantify consumer behavior) and also between demographic factors and psychological factors. The intent using chi square test of independence was to check which factor individually has a significant association with which output column if any.

4.5.2 Predictive Analysis

In this section, we have leveraged the power of Machine Learning to find hidden patterns and to understand how the combination of personal and psychological factors as a combination influence distinct consumer behavior. Here, we created 2 dataframes. One of the user data and the other one for merged data which is the combination of user data and non user data. Next, we used Cronbach's Alpha, a metric to measure the reliability of survey data, to check the reliability of our data. Then we performed label encoding and prepared that data for predictive analysis.

We applied Principal Component Analysis (PCA) on the merged data after scaling it using Standard Scaling. For our study, we aimed for 95% information retention. Then we calculated the top 10 features of each component based on their absolute loading value. Additionally, the type (positive or negative) of their impact is also printed based on the sign of the loading value. We interpreted the results of PCA based on these following statements:

1. Each component represents segments or groups of people that have similar consumer behavior and mindset.
2. The features on top have more influence on the group/segment/cluster/component.
3. The loading shows how those features affect the consumer behavior of that group/component.
4. The bigger the Cumulative Explained Variance, the more important the group.

Classification and Regression Analysis

In this section, we used supervised learning algorithms, to understand which factors influence a particular consumer behavior when multiple factors are involved. This resembles a real world scenario, where all the personal, psychological, cultural and economic factors are involved. As the motivation involves understanding which features/factors play the major role in understanding consumer behavior, we used tree based ensemble learning models like Decision Tree, Random Forest and Light Gradient Boosting Machine (LGBM) which inherently calculate feature importance in order to make predictions. Moreover, we also performed automated hyperparameter tuning using Grid Search CV using Stratified K fold Cross Validation (since our data is imbalanced) in order to find the best accuracy. Lastly, our dataset consisted of both numerical (usage and satisfaction data) and categorical data (consumer behavior data), we had to use both the regression and classification variant of these models.

Feature Importance Calculation

In this section, we took the feature importance data from all the models for each output column and then tried to calculate a combined feature importance dictionary for each output column. To do so, at first, we scaled the feature importance data from each model using Min Max Scaling so that the scale of the feature importance data stays the same regardless of the model used and everything stays in the same range. Afterwards, we employed a voting system using a weighted average method where for classification models the weights are their accuracy score for that output column and for regression, it's the inverse of MSE. As a result, the weight of each model, in calculating the feature importance now depends on the performance of that model. This calculated weight average then becomes the combined feature importance data for that output column. Lastly, we visualize the combined feature importance for each distinct consumer behavior using a bar chart which includes the top 5 most important features.

Let's define the weights w_i based on the model performance:

- For classification models, w_i is the accuracy score of model i for the output column.
- For regression models, w_i is the inverse of the mean squared error (MSE) of model i for the output column.

So, the formula for the weight average can be expressed as:

$$WA_j = \frac{\sum_{i=1}^{n} w_i \times SFI_{ij}}{n}$$

where

$$w_i = \begin{cases} \text{Accuracy Score }_i & \text{for classification models} \\ \frac{1}{\text{MSE }_i} & \text{for regression models} \end{cases}$$

$$n \;=\; number\, of\, models$$

$$SFI \;=\; Scaled\, Feature\, Importance$$

K Means Clustering

Lastly, we took the feature importance of all the models for all the consumer behavior, scaled them using Min Max Scaling and created a new dataset where each column represents each factor's scaled importance across all the model and consumer behaviors. Then we transposed the dataframe and used the Elbow test based WCSS and found the reasonable value of K. Next, we applied K Means Clustering to cluster the feature importance data. Lastly, we calculated the top clusters based on the highest mean value and listed the feature names of those clusters to get the overall most important features across all the models which have the most influence on overall consumer behavior.

Chapter 5

Result

5.1 Experimental Setup

In order to conduct the computational experiments for this research, we leveraged the resources provided by the free tier of Google Colab. Google Colab, short for Colaboratory, offers a cloud-based platform with free access to a GPU (Graphics Processing Unit) and TPU (Tensor Processing Unit). The utilization of Google Colab facilitated seamless execution of machine learning algorithms and data analysis tasks, eliminating the need for substantial local computational resources. This cost-effective and accessible platform allowed us to efficiently run experiments, train models, and analyze results collaboratively in a cloud environment, contributing to the producibility and scalability of our research efforts.
Additionally, during our experiments, we utilized the following resources on Google Colab:

- Python 3 Google Compute Engine backend
- System RAM: 1.3 / 12.7 GB
- Disk Usage: 26.3 / 107.7 GB

5.2 Evaluation Metrics

5.2.1 Regression

MSE (Mean Squared Error)
MSE is a metric used to measure the average squared difference between the actual and predicted values. It is commonly employed to evaluate the performance of regression models. The formula for MSE is:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \tag{5.1}$$

where:

$$\begin{aligned} n &\text{ is the number of data points,} \\ Y_i &\text{ is the actual value,} \\ \hat{Y}_i &\text{ is the predicted value.} \end{aligned}$$

RMSE (Root Mean Squared Error)
RMSE is the square root of the Mean Squared Error. It is used to express the error in the same units as the predicted and actual values. The formula for RMSE is:

$$\text{RMSE} = \sqrt{\text{MSE}} \tag{5.2}$$

where MSE is the Mean Squared Error.
MAE (Mean Absolute Error)
MAE is another metric for measuring the average absolute difference between the actual and predicted values. It is less sensitive to outliers than MSE. The formula for MAE is:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|Y_i - \hat{Y}_i| \tag{5.3}$$

where:

$$\begin{aligned} n &\text{ is the number of data points,} \\ Y_i &\text{ is the actual value,} \\ \hat{Y}_i &\text{ is the predicted value.} \end{aligned}$$

R2 Score R2 Score, or the coefficient of determination, measures the proportion of the variance in the dependent variable that is predictable from the independent variables. The formula for R2 Score is:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2} \tag{5.4}$$

where:

$$\begin{aligned} Y_i &\text{ is the actual value,} \\ \hat{Y}_i &\text{ is the predicted value,} \\ \bar{Y} &\text{ is the mean of the actual values,} \\ n &\text{ is the number of data points.} \end{aligned}$$

5.2.2 Classification

Accuracy Score
Accuracy Score is a metric used in classification models to measure the ratio of correctly predicted instances to the total instances. The formula for accuracy is:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \tag{5.5}$$

Confusion Matrix
A Confusion Matrix is a table used in classification to evaluate the performance of a model. It shows the number of True Positives, True Negatives, False Positives, and False Negatives. The matrix is essential for calculating metrics like precision, recall, and F1 score.

5.2.3 Clustering

Elbow Test and WCSS (Within-Cluster Sum of Squares)
The Elbow Test is used in clustering algorithms (e.g., k-means) to determine the optimal number of clusters. WCSS is the sum of squared distances between each point and the centroid in a cluster. The Elbow Test involves plotting the WCSS for different numbers of clusters and selecting the "elbow" point where the rate of decrease slows, indicating an optimal number of clusters. The formula for WCSS is:

$$\text{WCSS} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \|x_{ij} - c_i\|^2 \tag{5.6}$$

where:

$$\begin{aligned} k &\text{ is the number of clusters,} \\ n_i &\text{ is the number of data points in cluster } i, \\ x_{ij} &\text{ is the } j\text{-th data point in cluster } i, \\ c_i &\text{ is the centroid of cluster } i. \end{aligned}$$

5.3 Experimental Findings

5.3.1 Qualitative Analysis

1. Preliminary analysis

For the preliminary analysis we have done basic sentiment analysis of the columns - thoughts and opinions, problems, suggestions and industry perception for understanding the attitude and perceptions of consumers thought, their opinions and also the industry perceptions. For that we have firstly vectorized the data and then generated a bag of words and word cloud for the easy visible look of our qualitative data. Below are the word clouds for users and non-users :

2. Sentiment Analysis

First of all we have categorized sentiment as if the score is higher than zero then it will be positive else if less than zero then it will be negative otherwise it will be neutral. Then we have categorized the subjectivity - if the score is higher than 0.5 then it will be subjective and if it is less than 0.5 then objective otherwise it will be neutral. After that we have labeled the rows with neutral where we imputed no comments and performed sentiment analysis using the textblob function and have generated sentiment score. Then we plotted all the sentiment scores corresponding to the vectorized columns to see the sentiment reports.

In summary, we found that most of the user's **Thoughts opinions** have **positive sentiments**, most of the **problems reports** of users have **neutral sentiments**, most of the **suggestion** of users have **neutral sentiments** and most of the **industry perceptions** of users have **positive sentiments**. Also we have got **positive sentiments** in the case of **Industry perceptions of the non-users**.

Figure 5.1: Word Cloud Users

Figure 5.2: Word Cloud Non-Users

3. Thematic Analysis:

Generating Initial Codes :
In this part we have gone through our collected responses of the questions :

1. Thoughts and Opinions regarding the App that user prefer most
2. Perceptions towards the Digital Payment Industry as a whole
3. Problems or Challenges that Users have encountered when using Digital Payment Apps
4. Suggestions for improving Digital Payment Apps

Firstly we have manually gone through all the responses and focused on organizing the collected data through a meaningful code which will turn the extensive information into a meaningful segment. We also used numbering and maintained a serial for our generated codes so that they can be easily accessible.

Generated Themes :
Then we have generated the final themes through comparing and matching the initial codes and this iterative process allowed us to refine the existing codes and generating overall themes for qualitative thematic analysis. Following are the themes that we got from our code analysis :

1. Reason Preference - Constraint Options
2. Availability and Portability
3. Reason Preference - Brand Recognition - Popularity
4. Accessibility and Integration
5. Technical Issue - Stability, Service, Delay
6. Industry Perception - Positive Future Prediction
7. Security Issues
8. Contactless Payments
9. Transaction - Financial issues
10. Network Issues - Connectivity

Review Themes :
Below are the brief explanation of each theme that we got in qualitative thematic analysis and the senses they hold :

1. **Reason Preference - Constraint Options :** It represents the reason behind for choosing that particular digital payment app by the users. It also represents the constrained options that the users faced . For example - if a user is always using Bkash and have never used any other app then that user is constrained with variation scope. Similarly if any user is using Nagad for his transaction purpose only because the charge is less then it represents his reason preference.

2. **Availability and Portability** It mainly represents the availability of payment apps like - cash out / cash in outlets are widely available, most of the shopkeepers accept Bkash payments. Also it represents the portability of digital payment apps like - they are the most convenient way for transaction, whenever and wherever a person can access his money.

3. **Reason Preference - Brand Recognition - Popularity :** It represents that users are choosing that particular app considering it's popularity. For example - Bkash is really easy to use and almost everyone use Bkash, thus it is convenient to do cash transfers with others. Also the theme represents that users make their choice of using which payment app considering their branding and popularity.

4. **Accessibility and Integration :** It represents that the payment app that the users are using is connected with various commercial sites and banks. For example, users can get money from their bank accounts and cards through Bkash. So they like to prefer Bkash most. Also Bkash have the highest integration with other apps which made it more suitable for daily usage. It represents some complaint responses such as - account creation process in Bkash is a bit too complex and also causes accidental transactions to wrong users, resetting pins are inconvenient.

5. **Technical Issue - Stability, Service , Delay:** It mainly represents some opinions and complaints that the users commented according to their using experiences. For example - frequent session outs, looks for app update very frequently, transaction failures, app hanging issues, pin recognition issues etc. Users also complained that some payment apps seemed to take a little longer time to process transactions, which makes them a bit anxious. Also sometimes they failed to use it in emergency cases for not having maintenance. Most of the users complained that offline access of Bkash app (using*247#) is quite slow and lacks many features. Also acceptance in the merchants, linking with cards are other technical issues that users faced. The agents don't want to cash in or out during late nights and not having any printable statement are the common service issues that most users have raised. Also auto logout without notifying is another service problem that users frequently faced. So we have tried to represent all complaints and issues like these through this theme.

6. **Industry Perception - Positive Future Prediction:** It represents all the comments that users made wishing a prospective future of that payment app industry. For example - some users commented that Bkash is booming and after 50 years hard cash will reduce significantly because of payment apps like Bkash. Also users commented that overall digital payment apps have very

progressive market and a growing industry and soon it will become the default mode of payment and will have a positive impact in modernization.

7. **Security Issues:** It mainly represents the security related issues of digital payment apps. For example - some users have commented that their accounts got hacked due to not becoming more cautious and for not having three step verification process like google accounts. Also it causes great hassles to them when they by chance forget the pin. It takes huge time by the admin authority to reset new pins. Some users have suggested for implementing three-step verification and bio-metric verification policies to every payment apps which has been recently implemented by Bkash.

8. **Contactless Payments:** It mainly represents all those comments that users made on the cash less payment facilities of payment apps. For example - in COVID-19 period coming in contact with hard cashes was a sever cause of virus attack. Also almost all the hard cashes that we found is mostly contaminated and unhygienic. So most of the users found the cashless payment options more safe and hygienic. Also some users commented that, through this they can pay any broken amount, no need to always pay round figures.

9. **Transaction - Financial issues:** It represents user's complaints and suggestions regarding the cash out fees and charges of digital payment apps. For instance - some users commented that Bkash has higher rate of charges than other payment apps and that is a reason behind their not using Bkash. Also Bkash has too much transaction cost, high charge for withdrawing cash. So they claimed that it would be more efficient for them if the charges could me more lessen.

10. **Network Issues - Connectivity:** It represents all the connection issues that users faced. For example - without internet connection they are unable to use these apps, sometimes it caused failed transactions and unnecessary charges. Sometimes they faces scammers and server related problems. Most of the users have focused on the problem of internet dependency of the payment apps.

5.3.2 Quantitative Analysis

1. Cross Tabular Analysis :

We have did cross tabular analysis among the demographic columns - age, gender, employment, the psychological feature columns and the output columns. We are adding here some of them as example.

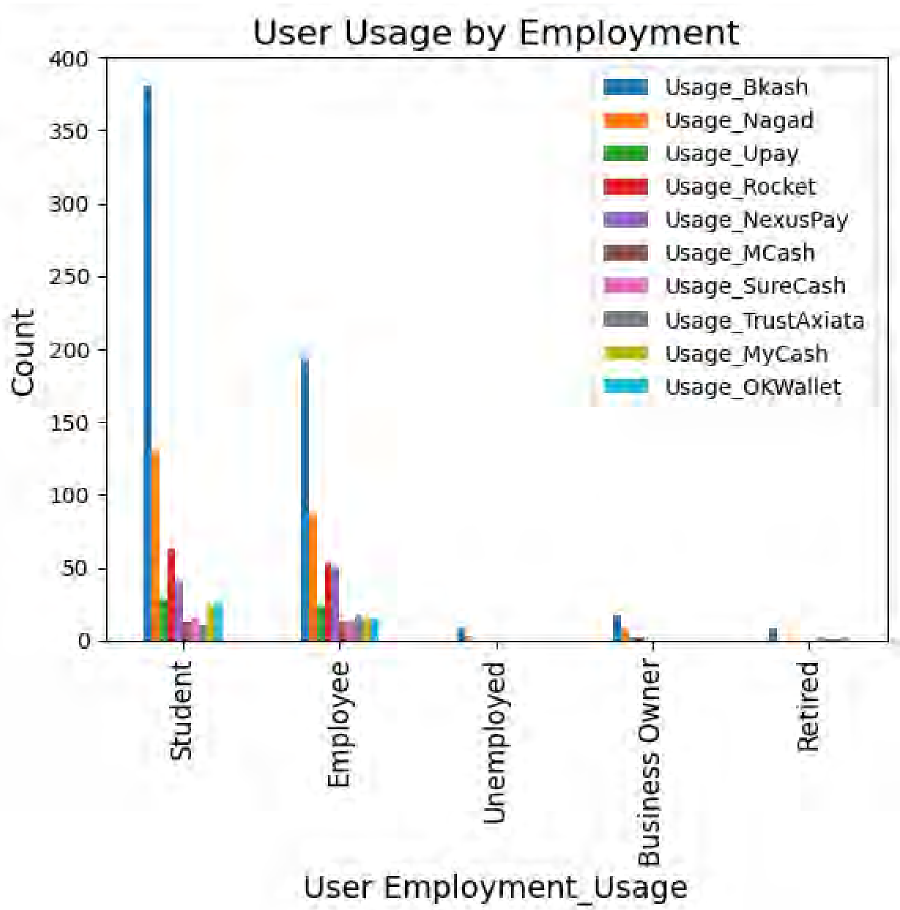

Figure 5.3: Usage vs Demography- User Employment

From the user usage and user employment cross tabular analysis graph given above we can see that according to our data, most of the users of digital payment apps are students and they mostly prefer to use Bkash and Nagad for their online transactions. According to the graph, the second most users of digital payment apps are employed persons and they mostly prefer to use Bkash and Nagad which is the similar preference as student users. Lastly the third most usage is done by the business owners and they are also likely to use Bkash for their transaction purposes.

Then we have made a comparison between the payment apps with respect to user satisfactions. According to the different income groups most of less than 10,000 tk income grouped people are more satisfied using Bkash rather than any other payment apps and least satisfied with MCash. The case is quite similar for the rest

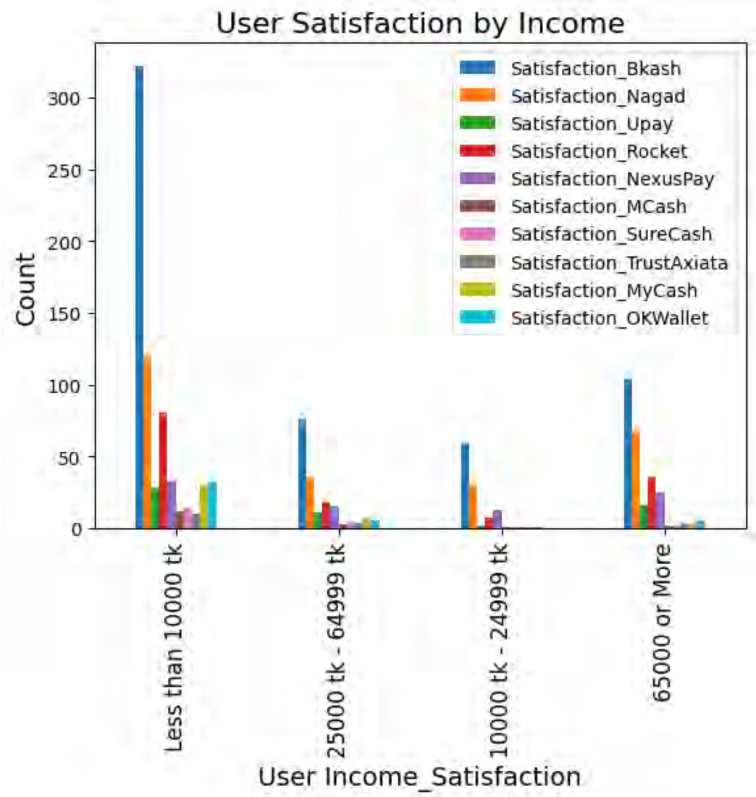

Figure 5.4: Satisfaction vs Demography- User Income

of the income group people. All of them are mostly satisfied with the experience of using Bkash rather than other payment apps.

According to the Transaction type vs demography graph, most of the less than 25 aged users prefer to use online purchase transactions, bill payment transactions and money transfer transactions and the 25-34 aged people are more interested in money transfer transactions more than online purchase and bill payment transactions. The aged people group are interested in digital payment apps for mostly their bill payment and money transfer transactions.

According to the preference criteria vs demography graphs, most of the male users prefer to use online transaction platforms that has lower fees, easier to use and have better security. In the case of female members, they mostly prefer easier to use apps, better security apps, lower fees apps and apps with better features.

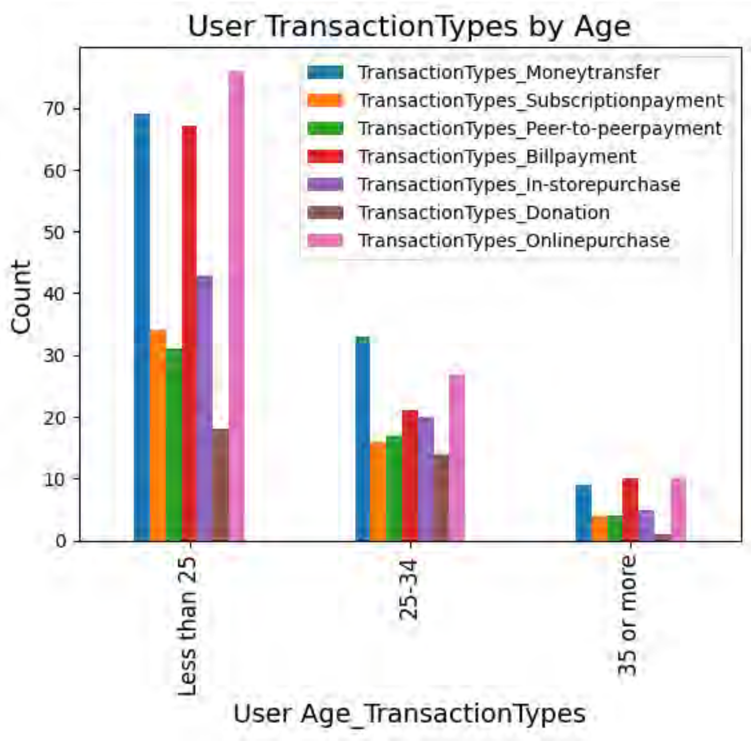

Figure 5.5: Transaction Type vs Demography- User Age

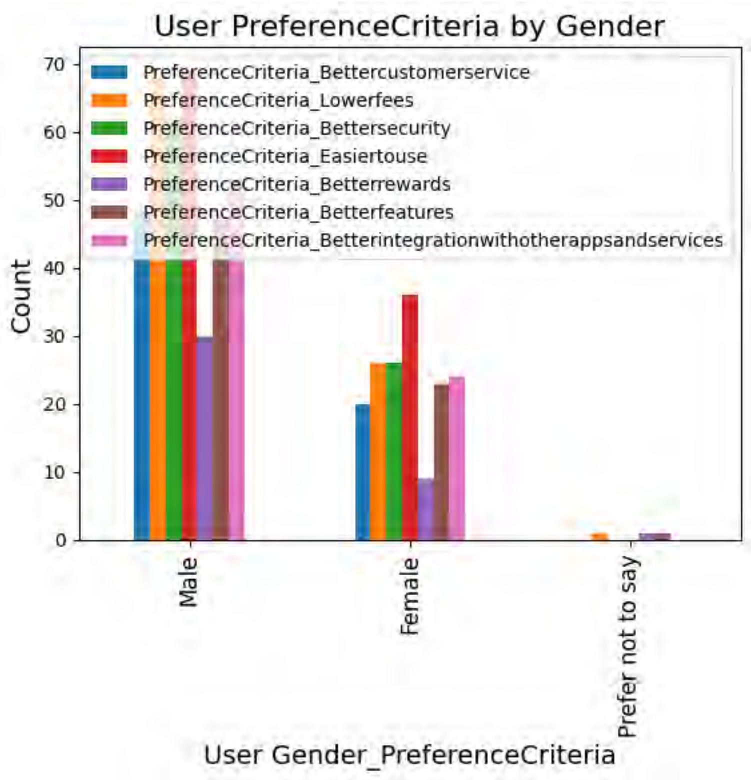

Figure 5.6: Preference Criteria vs Demography- User Gender

2. Chi Square Test of Independence:

The results of the Chi Square Test of Independence will help us to determine whether there is any evidence to reject the assumption of independence between two feature columns (demographic and psychological factors) and the output columns (metrics to quantify consumer behavior) and also between demographic factors and psychological factors. If the test is statistically significant, it suggests that there is a relationship between the variables.

2.1 Demography and Consumer Purchase Behavior:

2.1.1 For Users :

After conducting Chi Square Test of independence on demographic columns, including Age, Gender, Employment, and income in relation to various user-related target columns such as Frequency of Usage, Financial Impact, Recommendation Likelihood, Transaction Preference, and Impulse Purchases, we get that:

- **Age and Financial Impact:** A significant association exists between users' age and their reported financial impact. This suggests that age may influence how individuals perceive and are affected by financial considerations.
- **Gender, Recommendation Likelihood, and Impulse Purchases:** There is a meaningful association between users' gender and their likelihood to recommend digital payment app as a payment mode. Additionally, gender shows a significant association with users' tendency for impulse purchases. This implies that gender plays a role in shaping both the inclination to recommend and impulsive buying behaviors.
- **Employment, Recommendation Likelihood, and Transaction Preference:**

  Users' employment status is significantly associated with both their likelihood to recommend and their transaction preferences. This indicates that employed individuals may exhibit distinct patterns in terms of recommending products/services and their preferred transaction methods.

2.1.2 For Non-Users:

After conducting Chi-Square tests on demographic columns, including Age and Employment, concerning specific non-user-related target columns such as Transaction Preference. The statistical significance of these associations highlights the potential impact of age and employment status on the transaction preferences of individuals who are not actively using digital payment app.

- **Age and Transaction Preference:**

  A statistically significant association exists between the age of non-users and their transaction preferences. This suggests that age may play a role in shaping the transaction methods preferred by individuals who are not regular users.

- **Employment and Transaction Preference:**

  Similarly, there is a significant association between the employment status of non-users and their transaction preferences. This implies that the employment status may influence the preferred methods of transaction for individuals who are not actively engaged as users.

2.2 Psychological Factors and Consumer Purchase Behavior :

For Users: The following insights were obtained through Chi-Square tests conducted on user-related reasons for use and preferences, highlighting significant associations that can inform the understanding of user behavior and preferences -

- **Reason for Use - Security and Recommendation Likelihood**

  This suggests that users who prioritize security considerations may also be more likely to recommend to use digital payment apps.

- **Reason for Use - Security and Transaction Preference**

  Similarly, users who emphasize security in their reasons for using the digital payment apps, may have distinct preferences in how they conduct transactions.

- **Reason for Use - Easy to Use and Recommendation Likelihood**

  Again a statistically significant association exists between the preference for easy-to-use features and the likelihood of making recommendations. Users valuing ease of use may be more inclined to recommend digital payment apps.

- **Reason for Use - Preference Only Used this App and Recommendation Likelihood**

  Also a significant association is found between the preference for using only this app and the likelihood of making recommendations. Users exclusively using this app may be more inclined to recommend it.

- **Reason for Use: Access to Credit and Transaction Preference:**

  This indicates that users citing access to credit as a reason may exhibit specific preferences in their transaction methods.

- **Reason for Use: Rewards and Frequency Usage:** Users motivated by rewards may use digital payment apps more frequently.

- **Reason Preference: Transaction Fees and Impulse Purchases:** Users who prioritize cost considerations in their preferences may engage in more impulsive buying behavior.

- **Reason Preference: App Features and Promotional Influence:** Users valuing app features may be more influenced by promotional efforts.

- **Reason Preference: Safety and Security and Security Confidence:** There is a significant association between the preference for safety and security features and user's confidence in the security of this service.

- **Reason Preference: Rewards and Offers and Promotional Influence** Users motivated by rewards and offers may be more responsive to promotions.

The following table summarizes the significant associations between reasons for user preferences and their corresponding factors (each association is listed with only one user preference) -

Reason for User Preference	**Associated Factor**
Access to credit	Transaction Preference
Rewards	Frequency Usage
App features	Promotional Influence
Brand Recognition	Recommendation Likelihood
Easy to use	Recommendation Likelihood
Integration with other apps and services	Security Confidence
Only used this app	Online Spending, Recommendation Likelihood
Rewards and offers	Promotional Influence
Safety and Security	Security Confidence
Transaction fees	Impulse Purchases
Security	Recommendation Likelihood, Transaction Preference

Table 5.1: Associations Between Reasons for User Preference and Factors

2.3 Attitude and Consumer Purchase Behavior :

The result suggests that there is a statistically significant relationship between the subjective nature of thoughts and opinions expressed by users and their online spending behavior.
In practical terms, it could mean that users with more subjective thoughts and opinions are more likely to exhibit specific patterns in their online spending compared to users with less subjective expressions. This insight could be valuable for understanding how user attitudes and subjective perceptions might influence their online spending decisions.

2.4 User Perception and Consumer purchase Behavior :

The result - **there is a significant association between the Industry Perception Sentiment and Financial Impact** indicates that the way people perceive the industry (positively, negatively, or neutrally) is linked to how they assess the financial consequences or benefits associated with that industry. This insight could be valuable for understanding the interplay between public perception and financial considerations in a particular industry.

2.5 User Concerns and Consumer purchase behaviour

These associations provide insights into potential relationships and dependencies between users' concerns and various aspects of their experience, such as convenience, promotional influence, and security confidence.

- **Concerns Trust Worthiness and Is Most Convenient:** It means that users who express concerns about trustworthiness may have similar opinions about the convenience of digital payment apps.
- **Concerns Technical Issues and Promotional Influence:** Users who express concerns about technical issues may be influenced by promotions.
- **Concerns Technical Issues and Security Confidence**

  Users with technical concerns may also have specific views on the security aspect.

3. Predictive Analysis :

3.1 Cronbach's Alpha results :

Here's a brief explanation of the values that we got regarding the internal consistency and reliability of our survey after applying Cronbach's Alpha:

- **Overall Merged Data:**
  - Cronbach's Alpha: 0.649
  - 95% Confidence Interval: [0.57, 0.72]

  This means the set of items in the merged data has a moderate level of internal consistency.

- **User Data:**
  - Cronbach's Alpha: 0.806
  - 95% Confidence Interval: [0.758, 0.849]

  A high Cronbach's Alpha for user data indicates a strong internal consistency among the items in the scale. This suggests that the items in the user predictive data are reliably measuring the underlying construct.

3.2 PCA

Then we applied PCA on the merged data aiming for 95% information retention. Then we calculated the top 10 features of each component based on their absolute loading value. In the result, each component represents segments or groups of people that have similar consumer behavior and mindset. The features on top have more influence on the groups or other components . The loading shows how those features affect the consumer behavior of that group/component. In total, there are 27 components. The top 5 components are interpreted below:

Principal Component 1:

Top Negative Features:

- No Use Reason I don't trust digital payment apps: Negative (-0.1670)

- Encouragement Increase in acceptance in merchants and businesses: Negative (-0.1659)
- Encouragement Faster transaction processing: Negative (-0.1650)
- Encouragement Lower fees: Negative (-0.1645)
- Encouragement Better customer support and response time: Negative (-0.1643)

Interpretation: Component 1 is characterized by a strong negative association with reasons for not using digital payment apps, such as lack of trust and safety concerns. Additionally, negative associations with features like faster transaction processing and better customer support suggest a reluctance or dissatisfaction with these aspects.

Principal Component 2:

Top Negative Features:

- Employment: Negative (-0.5638)
- Age: Negative (-0.5236)
- Income: Negative (-0.4257)
- Concerns Fraud and Scams: Negative (-0.1665)
- Preference Criteria Better customer service: Negative (-0.1587)

Top Positive Features:

- Gender: Positive (0.0888)

Interpretation: Component 2 is predominantly characterized by negative associations with demographic factors like employment, age, and income. Additionally, concerns about fraud and scams, as well as preferences for better customer service, contribute to the overall negative nature of this component. Interestingly, gender has a positive association.

Principal Component 3:

Top Positive Features:

- Concerns Trust worthiness: Positive (0.4730)
- Concerns Technical Issues: Positive (0.4257)
- Concerns Fraud and Scams: Positive (0.4047)

- Concerns Accessibility and Compatibility: Positive (0.2457)
- Concerns Security and Privacy: Positive (0.2141)

Top Negative Features:

- Age: Negative (-0.2654)
- Employment: Negative (-0.2080)
- Industry Perception Sentiment: Negative (-0.1004)
- No Use Reason I don't know how to use digital payment apps: Positive (0.1097)

Interpretation: Component 3 is characterized by positive associations with concerns related to trustworthiness, technical issues, fraud, accessibility, and security. Negative associations with age, employment, and industry perception sentiment suggest that these factors contribute to a different aspect of user attitudes or concerns. The positive association with the reason for not using apps due to lack of knowledge implies a contrasting perspective.

Principal Component 4:

Top Positive Features:

- Concerns Technical Issues: Positive (0.3994)
- Concerns Fraud and Scams: Positive (0.3777)
- Industry Perception Sentiment: Positive (0.1975)

Top Negative Features:

- Concerns Accessibility and Compatibility: Negative (-0.3923)
- Concerns Security and Privacy: Negative (-0.3837)
- Concerns Trust worthiness: Negative (-0.3654)
- Gender: Negative (-0.3039)
- Preference Criteria Lower fees: Negative (-0.1439)
- Employment: Negative (-0.1198)

Interpretation: Component 4 is characterized by positive associations with concerns about technical issues and fraud. Gender has a negative association, indicating differences in perception based on gender. Negative associations with concerns related to accessibility, security, and trustworthiness, as well as preferences for lower fees and employment, contribute to the overall nature of this component.

Principal Component 5:

Top Positive Features:

- Gender: Positive (0.4835)
- Concerns Fraud and Scams: Positive (0.3131)
- Concerns Security and Privacy: Positive (0.2886)
- Reason Preference Safety and Security: Positive (0.1189)
- Transaction Types Bill payment: Positive (0.1105)

Top Negative Features:

- Concerns Accessibility and Compatibility: Negative (-0.5493)
- Income: Negative (-0.3089)
- Industry Perception Sentiment: Negative (-0.1989)
- Preference Criteria Easier to use: Negative (-0.1182)
- Age: Positive (0.1168)

Interpretation: Component 5 is characterized by a negative association with concerns about accessibility and compatibility. Gender, fraud concerns, security and privacy concerns, and a preference for safety and security contribute positively to this component. Negative associations with income, industry perception sentiment, and preferences for easier use suggest contrasting perspectives.

These interpretations provide insights into the underlying factors contributing to each principal component.

3.3 Regression Analysis :

Below is the performance report of each model in Regression Analysis . We have used MSE (Mean Squared Error), RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), R2 (R-squared) for measuring the error of the Regression models.

Table 5.2: Analyzing Regression Model Data for Bkash and Nagad

Models	Usage Bkash	Usage Nagad	Satisfaction Bkash	Satisfaction Nagad
Decision Tree	MSE : 1.897 RMSE : 1.377 MAE : 0.915 $R\hat{2}$: 0.404	MSE : 4.213 RMSE : 2.052 MAE : 1.490 $R\hat{2}$: -1.805	MSE : 1.332 RMSE : 1.154 MAE : 0.809 $R\hat{2}$: 0.461	MSE : 4.421 RMSE : 2.103 MAE : 1.519 $R\hat{2}$: -1.259
Random Forest	MSE : 1.294 RMSE : 1.137 MAE : 0.848 $R\hat{2}$: 0.593	MSE : 2.424 RMSE : 1.557 MAE : 1.308 $R\hat{2}$: -0.614	MSE : 0.778 RMSE : 0.882 MAE : 0.641 $R\hat{2}$: 0.686	MSE : 2.401 RMSE : 1.549 MAE : 1.305 $R\hat{2}$: -0.227
LGBM	MSE : 1.426 RMSE : 1.194 MAE : 0.849 $R\hat{2}$: 0.552	MSE : 2.257 RMSE : 1.502 MAE : 1.241 $R\hat{2}$: -0.503	MSE : 0.909 RMSE : 0.953 MAE : 0.707 $R\hat{2}$: 0.633	MSE : 2.696 RMSE : 1.642 MAE : 1.391 $R\hat{2}$: -0.378

As we can see, Random Forest performed the best among the models used. The performance of LGBM is also pretty good. However, Decision Tree performed poorly compare to other models.

Now, let's look a scatter plot as an example of model performance. Here Actual value means the true Labels and Predictive values means the values generated by the machine learning models. For example. we have added here the graph of Satisfaction Bkash for Random Forest as an example.

The actual satisfaction values are shown on the x-axis, and the predicted satisfaction values are shown on the y-axis. Each point on the graph represents a different Bkash user.

The diagonal line represents the perfect prediction line. If the model is perfect, all of the points would fall on this line or will be scattered around the line . However,

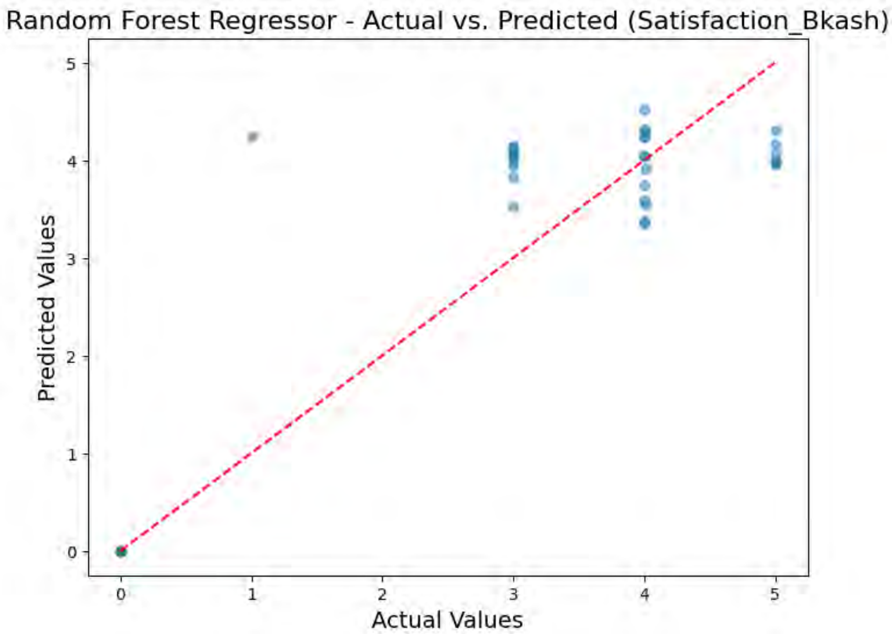

Figure 5.7: Random Forest - Actual vs Predicted (Bkash Satisfaction)

as you can see, the points are scattered above and below the line. This means that the model is doing a good job in the case of predicting satisfaction.

Overall, the graph shows that the Random Forest regression model is doing a good job of predicting the satisfaction of Bkash users.

The values that we got from analyzing Regression model for Satisfaction Bkash are below :

- **MSE: 0.77802** indicates an average squared error of 0.77802 units between the actual and predicted values.
- **RMSE: 0.88205** is the square root of MSE, representing the average error in the same units as the target variable.
- **MAE: 0.64075** measures the average absolute difference between actual and predicted values.
- **R2: 0.68578** indicates that 68.578% of the variance in the target variable can be explained by the model. As we have got a moderate value, it suggests a moderate predictive power.

3.4 Classification Analysis :

Overall Accuracy:

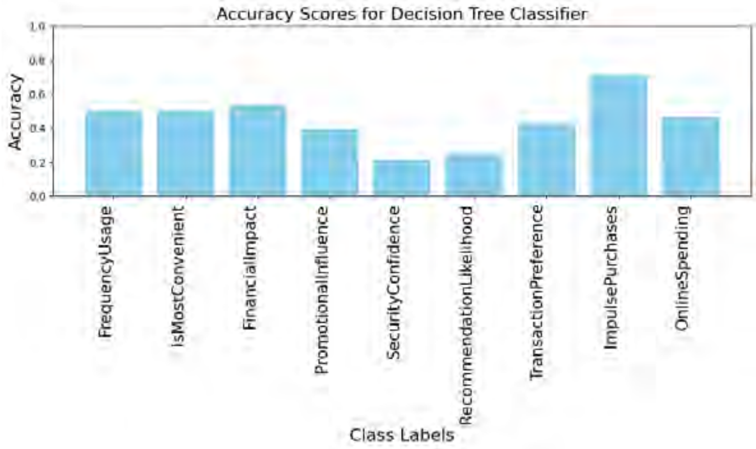

Figure 5.8: Accuracy Score for Decision Tree Classifier

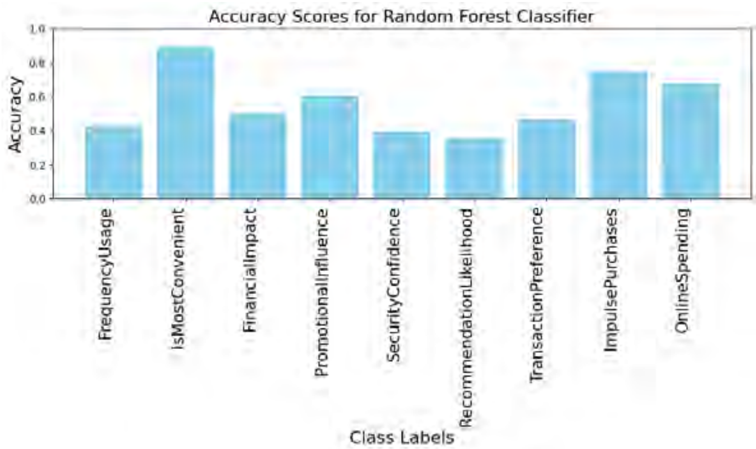

Figure 5.9: Accuracy Score for Random Forest

The graph Decision Tree classification indicates moderate accuracy across most class labels. The majority of bars reached almost 0.5 or higher on the accuracy scale, which is considered good performance for a Decision Tree classifier on real world data. On the other hand, most of the accuracy scores of Random Forest classifiers are above 0.5 accuracy scale and some of them reaches up to 0.8 and 0.9. Similar case happened for the case of LGBM classifiers. Most of the labels are around 0.4 to 0.5 and some of them reaches up to 0.8 and 0.9. According to the accuracy score results the Random Forest models provides best overall output. In case of merged data, the average accuracy to be around 0.5 and all model performed somewhat evenly.

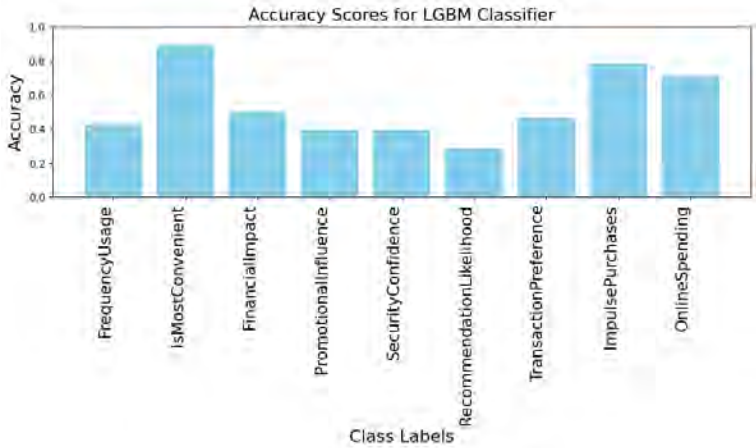

Figure 5.10: Accuracy Score for LGBM

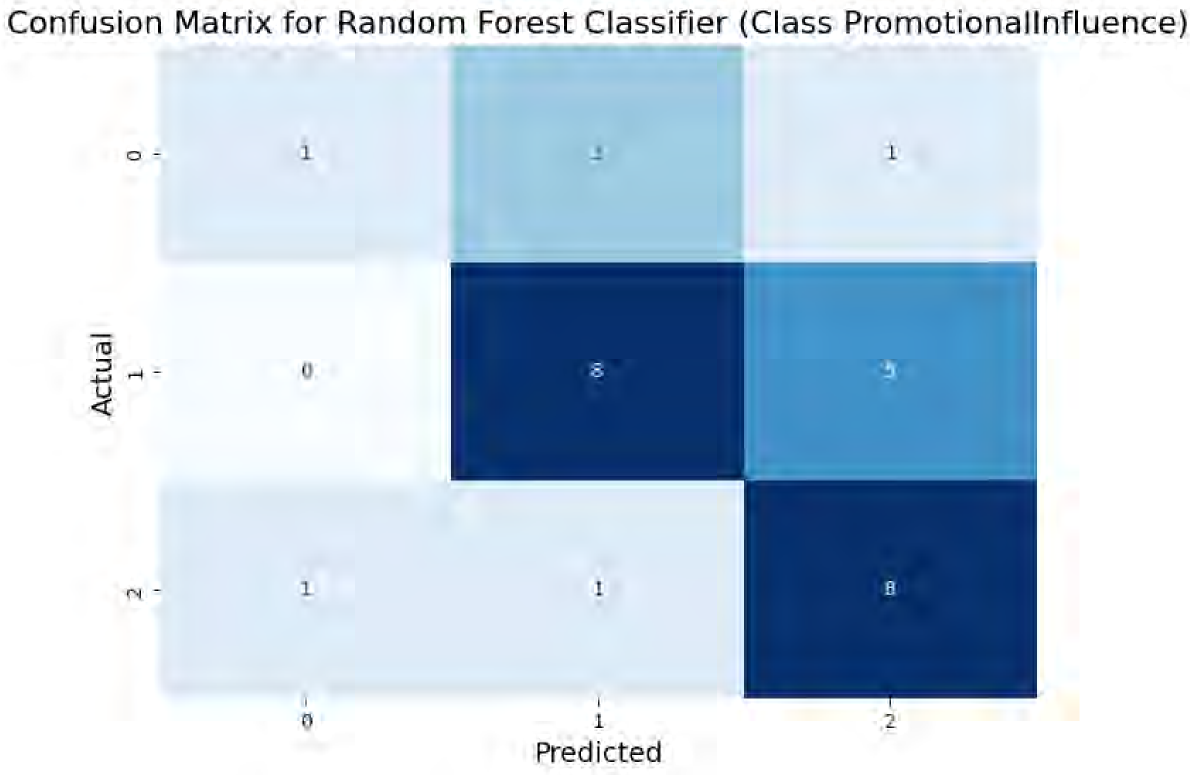

Figure 5.11: Confusion Matrix for Random Forest Classifier(Promotional Influence)

The above confusion matrix represents the performance of Random Forest that has been used to predict promotional influence. The rows represent the actual classes, and the columns represent the predicted classes.
The numbers written in each cell of the matrix show the number of data points that were classified into each combination of actual and predicted classes. For example, the cell in the third row and third column shows that 8 data points were actually in the "1" class and were also predicted to be in the "1" class.
The diagonal element of the matrix shows the number of correctly classified data points. The rest of the elements of the matrix display the number of incorrectly classified data points. For example, the cell in the first row and second column shows that 3 data points were actually in the "0" class but were predicted to be in the "1" class.

3.5 Feature Importance :

From the regression analysis report, we have generated feature importance dictionary which represents all the features and their importance . Through this we can know which factor or feature have the most influence on the particular output columns. However, it dosen't tell us whether the influence is negative or positive. In the table, The top 5 features influencing the output columns are listed in descending order (the higher the feature the more influence it has).

Table 5.3: Bkash and Nagad Combined Features (Usage)

Usage Bkash	**Usage Nagad**
Thoughts Opinions: Sentiments	Reason Preference: Only Used This App
Reason Use: Speed	Reason Preference: Availability
Reason Use: Rewards	Reason Preference: Integration with other apps and services
Preference Criteria: Better Features	Preference Criteria: Easier To Use
Preference Criteria: Better Security	Concerns: Fraud And Scams

Table 5.4: Bkash and Nagad Combined Features (Satisfaction)

Satisfaction Bkash	**Satisfaction Nagad**
Thoughts Opinions: Sentiments	Reason Preference: Only Used This App
Reason Use: Speed	Reason Preference: Availability
Preference Criteria: Better Security	Reason Preference: Integration with other apps and services
Reason Use: Rewards	Concerns: Fraud And Scams
Transaction Types: In Store Purchase	Preference Criteria: Easier To Use

We also generated feature importance dictionary to figure out which are the top features that influence the consumer behavior of users as well as the market itself using classification models. The results are summarized below.

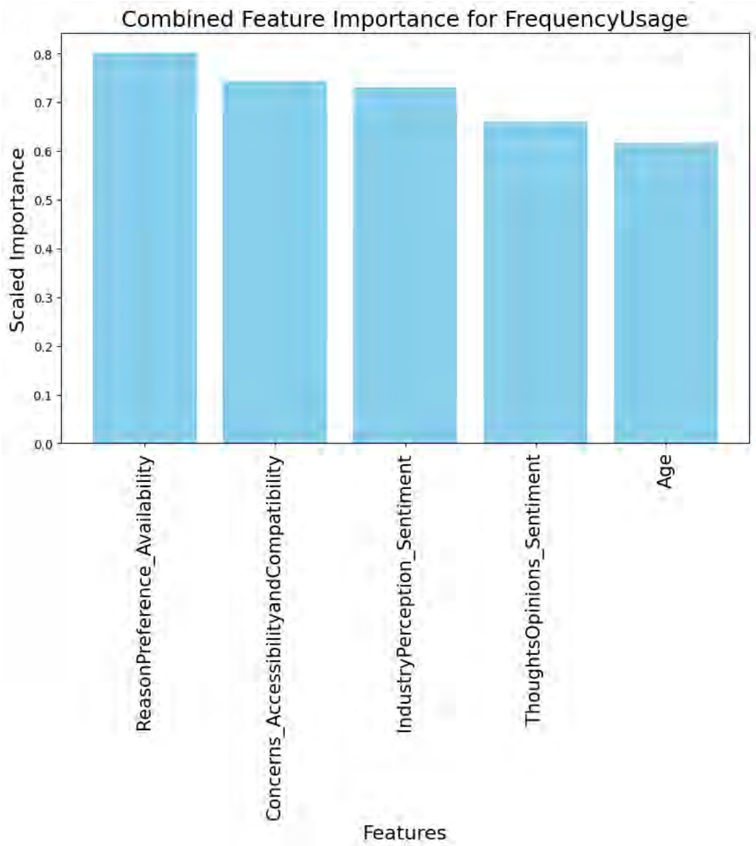

Figure 5.12: Combined Feature Importance for Frequency Usage

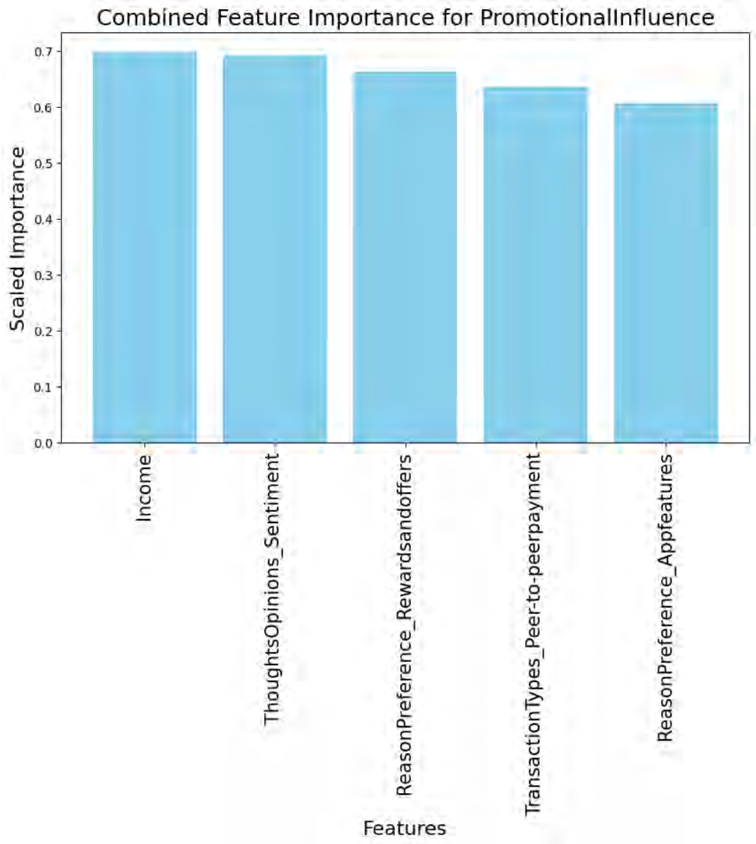

Figure 5.13: Combined Feature Importance for Promotional Influence

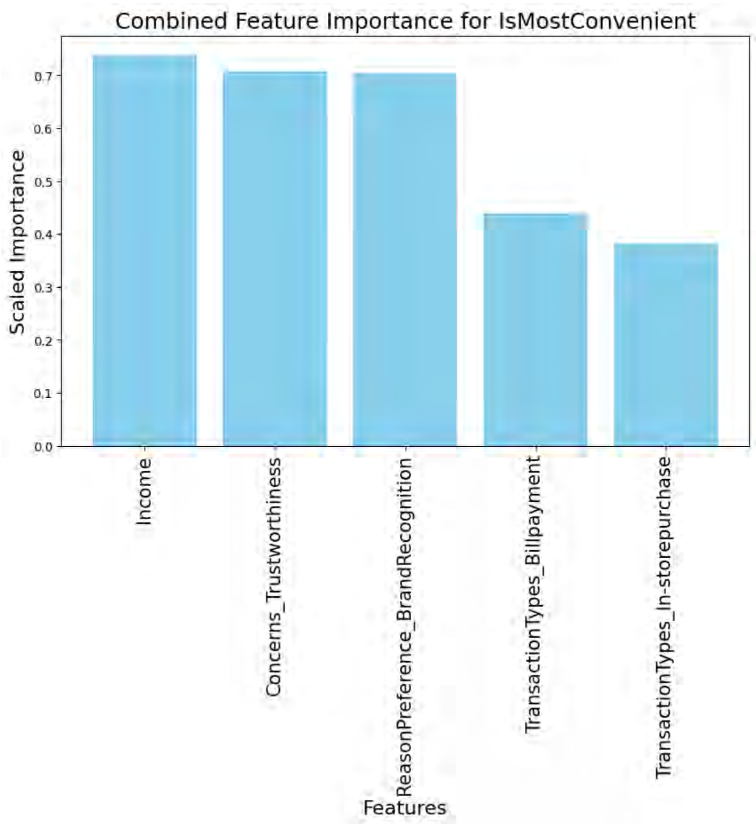

Figure 5.14: Combined Feature Importance for Is Most Convenient

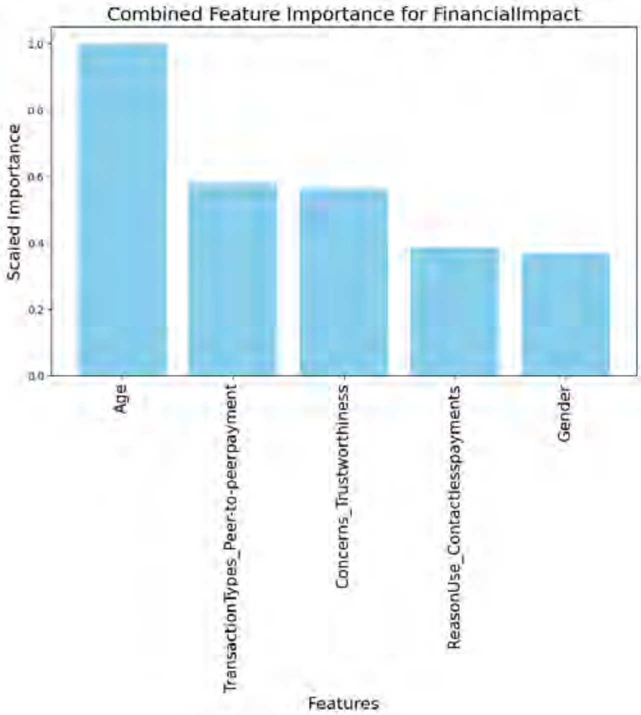

Figure 5.15: Combined Feature Importance for Financial Impact

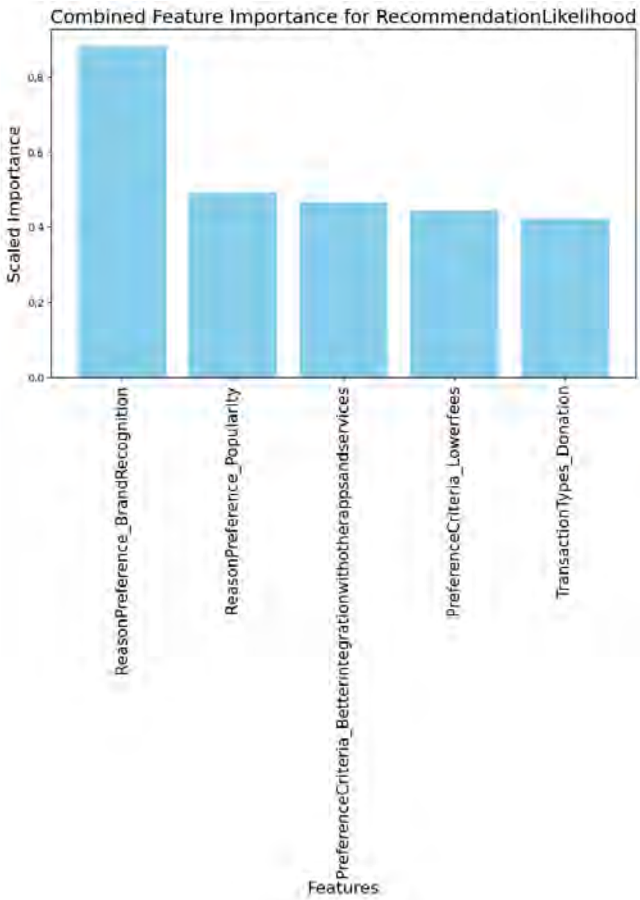

Figure 5.16: Combined Feature Importance for Recommendation Likelihood

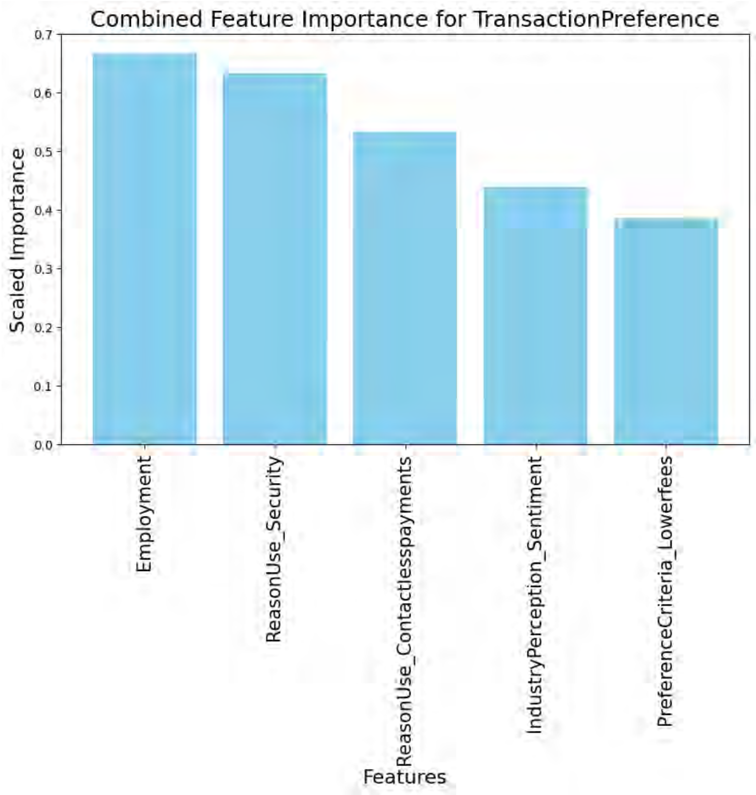

Figure 5.17: Combined Feature Importance for Transaction Preference

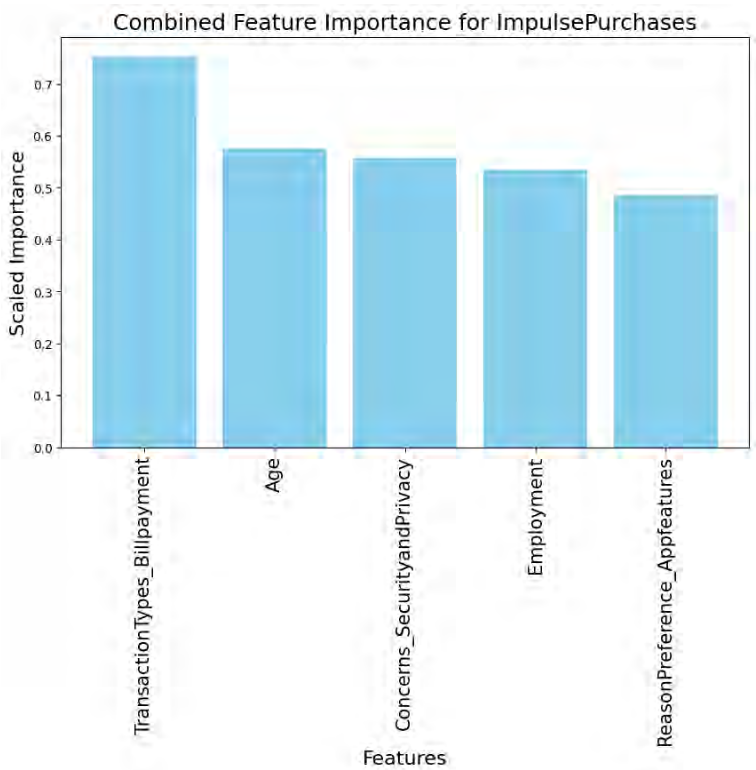

Figure 5.18: Combined Feature Importance for Impulse Purchase

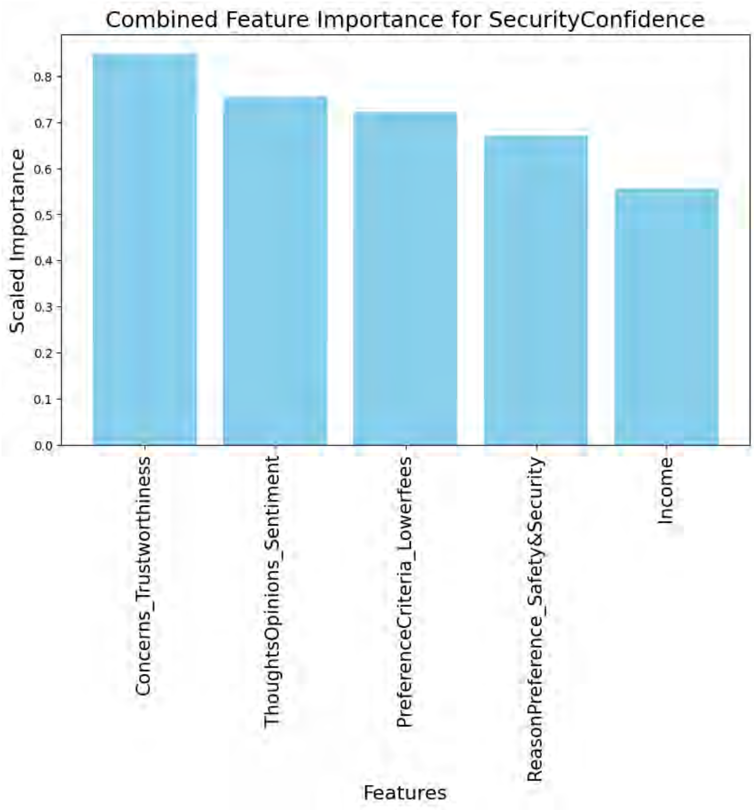

Figure 5.19: Combined Feature Importance for Security Confidence

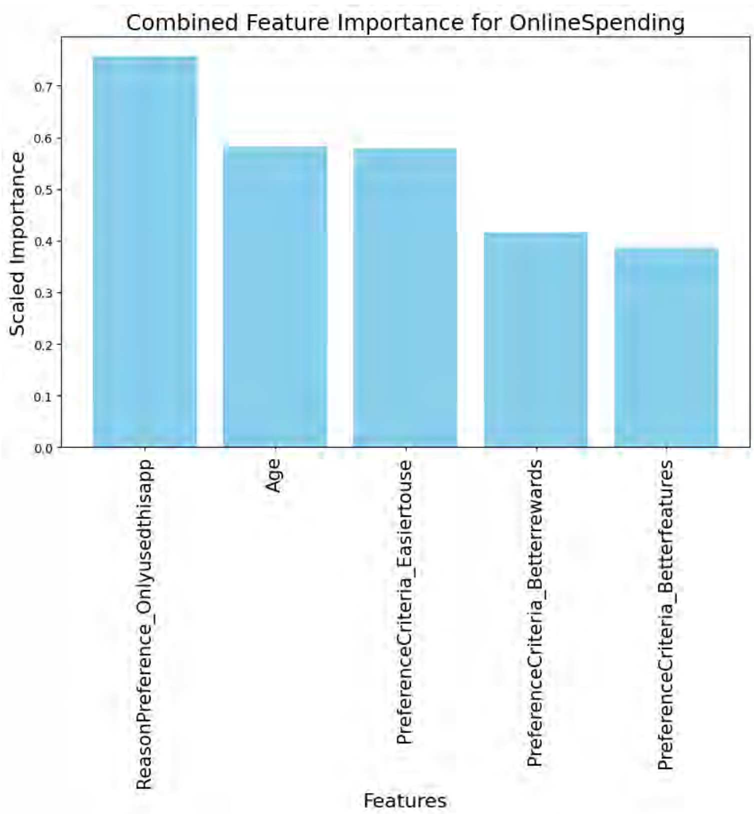

Figure 5.20: Combined Feature Importance for Online Spending

Table 5.5: Combined Feature Importance for Consumer Behavior of the Market

Output Columns	**Important Features (Descending order)**
Transaction preference	• Reason use: security • Employment • Thoughts and opinion sentiment • Transaction type: bill payment • Age
Impulse purchase	• Age • Reason preference: transaction fees • Concerns: trustworthiness • Industry perception sentiment • Transaction types: in store purchases
Online spending	• Reason preference: availability • Income • Reason preference: only used app • Industry perception sentiment • Encouragement: integration

Lastly, we took the feature importance of all the models for all the consumer behavior, scaled them using Min Max Scaling and created a new dataset where each column represents each factor's scaled importance across all the model and consumer behaviors. Then we transposed the dataframe and used the Elbow test based on WCSS and found the reasonable value of K. In our case, the value was 6. For any K below 6, results a high WCSS. On the other hand, for any K above 6, it causes the low number features for the top clusters.

In the end, after applying K Means Clustering and doing necessary calculation to find the top 3 clusters (in our case, cluster number 3, 5, 0), we found that the following 13 features are the overall most importance features that influence consumer behavior:

- Reason Preference Only used this app
- Age

- Income
- Reason Preference Availability
- Reason Preference Brand Recognition
- Concerns Accessibility and Compatibility
- Concerns Technical Issues
- Thoughts Opinions Sentiment
- Industry Perception Sentiment
- Reason Use Security
- Reason Use Contactless payments
- Transaction Types In-store purchase
- Transaction Types Bill payment

So according to the predictive analysis, among 53 features, these 13 features are the most important features which means these features have the most influence in consumer purchase behavior in the context of the digital payment industry of Bangladesh.

Chapter 6

Discussion

The exploration of psychological factors influencing consumer purchasing behavior in the digital payment industry in Bangladesh has produced insightful findings. Security reigns supreme, including availability, brand value and integration weave their own threads. Principal Component Analysis unravels user segments and reveals a surprising negative association with accessibility concerns, enriching the adoption narrative. Predictive analysis provide insights of current state of the digital payment landscape along with potential future trends, while nuanced findings, like the aforementioned negativity, refine our understanding of consumer decision-making. Though limitations exist, this research, through its comprehensive integration of findings, methodology, objectives, and data, lays a cornerstone for shaping strategies, refining user experiences, and fostering the sustainable growth of Bangladesh's evolving digital payment landscape.

6.1 Psychological Factors Influencing Consumer Behavior

The examination of consumer behavior in the context of digital payment apps in Bangladesh incorporates a comprehensive exploration of psychological factors through both qualitative (thematic analysis) and quantitative analyses. This dual approach provides a nuanced and multifaceted understanding of the intricate dynamics shaping users' decisions and preferences. Participants frequently expressed concerns regarding the security and privacy aspects of digital payment apps. This thematic analysis revealed a robust psychological factor where users prioritize the safety and protection of their personal information. This underlines the critical role of trust in influencing consumer behavior. Thematic analysis also brought to light the significance of transaction types, such as bill payments, in shaping consumer preferences. Additionally, preferences for availability, brand value and popularity, emerged as influential psychological aspects guiding user choices. The positive attitude expressed by users towards convenience and time-saving aspects further underscores the impact of these factors on adoption. Negative sentiments related to accessibility and compatibility issues were highlighted through the thematic analysis. Users' concerns regarding ease of access and compatibility with various devices emerged as influential psychological factors affecting their behavior. This reveals the importance of user-friendly interfaces and compatibility in driving consumer adoption. Quantitative analysis shad light on the features that greatly influence the consumer purchase behavior. Factors such as security, age, and income played significant roles in predicting user satisfaction and usage patterns. The evidence provided by regression analysis adds precision to the understanding of the psychological factors influencing consumer choice. For instance, usage of Bkash is greatly influenced by factors like speed, security and user attitude. On the other hand, in case of Nagad, availability, integration, user experience and user's leak of exposure to multiple apps play an vital role. Accuracy scores from the classification analysis, particularly using Random Forest, highlighted the effectiveness of models in predicting user behaviors. The exploration of accuracy scores and confusion matrices offered insights into the reliability of the models in different contexts, providing quantitative validation to the identified psychological factors. Feature importance analysis revealed key factors influencing different aspects of consumer behavior. Demographic factors such as age and income along with psychological factors such as preferring availability and brand recognition, security and technical concerns, users' attitude and perception all emerged as crucial in determining user behavior. This quantitative validation reinforces the significance of these psychological factors in shaping user decisions. The integration of qualitative and quantitative insights offers a holistic understanding of the psychological factors influencing consumer behavior in the adoption of digital payment apps in Bangladesh. Themes identified through qualitative analysis, such as security concerns, technical concerns, considering availability and brand value align seamlessly with the empirical evidence derived from quantitative analyses. This comprehensive approach enriches the study, providing depth and breadth to the exploration of psychological factors driving digital payment adoption. Overall, the combined findings emphasize the complex interplay of psychological factors as influential drivers of consumer behavior in the digital payment landscape of Bangladesh. This integrated understanding contributes to the development of targeted strategies and interventions that address the multifaceted nature of consumer decision-making in the adoption of digital payment technologies.

6.2 Effects of Digital Payment App Usage

The examination of the survey responses presents a comprehensive understanding of the multifaceted effects of digital payment app usage on consumer behavior. These effects encompass financial impact, impulse purchases, online spending, and promotional influence, revealing a nuanced interplay of demographic factors, sentiments, and transaction preferences. The survey data highlighted a significant financial impact on consumer behavior due to digital payment app usage. Users reported a notable shift towards online transactions and a preference for digital payment methods. Predictive analysis identified crucial factors such as age, gender, trustworthiness and peer to peer transaction types, underscoring the influence of these elements on users' financial behaviors. The feature importance analysis further emphasized their significance. The dataset and results revealed that digital payment apps contribute to increased impulse buying behavior among users. A substantial percentage of respondents acknowledged the role of these apps in prompting spontaneous and unplanned purchases. Predictive analysis demonstrated the importance of factors like age, employment, app features, security and privacy concerns in predicting impulse purchases. The feature importance analysis underscored the relevance of these elements in understanding and predicting impulsive buying behavior. Predictive models provided insights into online spending behavior influenced by digital payment app usage. Factors such as availability, easy to use, app features, age, income, and user perception were identified as influential in predicting the extent of online spending. Examination of promotional influence highlighted the significance of factors including income, user attitude, rewards and app features. These elements played critical roles in predicting users' responses to promotions. The document emphasized the importance of tailoring promotional strategies based on user demographics and sentiments. The interconnected nature of these effects has become apparent, emphasizing the need for targeted strategies and interventions based on a nuanced understanding of users' financial, impulsive, and promotional responses. Overall, our study underscore the transformative effects of digital payment app usage on consumer behavior. These effects extend beyond mere financial transactions, encompassing impulse buying tendencies, online spending patterns, and the significant influence of promotions and recommendations. Understanding and leveraging these insights are crucial for businesses and policymakers seeking to navigate and respond effectively to the evolving landscape of consumer behavior driven by digital payment app adoption.

6.3 Limitations

The research on the psychological factors influencing consumer behavior in the adoption of digital payment apps in Bangladesh, while insightful, is subject to several limitations that warrant consideration. These limitations encompass the sample size and representativeness, potential sampling bias, reliance on survey methodology, cultural specificity, limited coverage of psychological factors, model constraints, scope of digital payment apps, interpretation of PCA components, and external factors. The study's reliance on a sample size of 170 participants may limit its ability to fully capture the diverse demographics of Bangladesh. A larger and more representative sample would enhance the generalizability of the findings. If not carefully executed, the snowball sampling process may introduce bias. An online survey, for example, might exclude individuals without internet access, potentially skewing the results. The use of self-reported survey data introduces the possibility of response bias, where participants may provide socially desirable answers or inaccurately recall their attitudes and behaviors. This may impact the reliability of the findings. Moreover, rhe study's focus on Bangladesh may limit the generalizability of findings to other cultural settings. As a result, cultural nuances in consumer behavior might not be fully captured, affecting the external validity of the study. While the study identifies key psychological factors, it may not cover the entire spectrum of influences on consumer behavior, potentially overlooking nuanced aspects that could further enrich the analysis. The regression and classification models, while informative, have inherent assumptions and limitations. Deviations from assumptions, such as non-linearity, may impact the accuracy of predictions and interpretations. Consumer behaviors are dynamic, and the study's cross-sectional design might not capture changes over time. As a result, the findings may lack temporal relevance, especially in a rapidly evolving digital payment landscape. Additionally, Principal Component Analysis (PCA) results are subject to interpretation, and different researchers may interpret components differently. This subjective nature introduces a degree of variability in the analysis. The study may not fully account for external factors such as economic changes or government policies that could influence consumer behavior. Considering these factors could provide a more comprehensive understanding. The research might not extensively explore industry perspectives and strategies employed by digital payment service providers. Incorporating industry insights could contribute additional dimensions to the analysis. The inherent noise within real-world survey data, even after preprocessing efforts, has impacted the performance of models. For example, Random Forest, the best performing model in our case, showcase an average accuracy score of approximately 60% even after rigorous hyper-parameter tuning. Furthermore, the confusion matrices reveal issues such as overfitting. Acknowledging these limitations is crucial for a nuanced interpretation of the study's findings and serves as a foundation for future research endeavors to address these constraints and advance the understanding of consumer behavior in the context of digital payment app adoption.

Chapter 7

Conclusion and Future Work

In this comprehensive study exploring the psychological factors influencing consumer purchase behavior in the context of digital payment app adoption in Bangladesh, a nuanced understanding has been achieved. The research utilized a diverse set of methodologies and analytical tools to unravel the intricate dynamics shaping consumer purchase behavior. Principal component analysis (PCA) revealed distinct clusters of factors influencing consumer behavior, spanning security and privacy concerns, transaction preferences, and demographic variables. Notably, concerns regarding accessibility and compatibility emerged as significant negative factors, highlighting the multifaceted nature of consumer decision-making in the digital payment landscape. Insights from regression analysis showcased the superior predictive capabilities of the Random Forest model. Scatter plots illustrated the model's effectiveness in predicting user satisfaction, laying a foundation for understanding the nuanced interplay of features contributing to consumer contentment. The classification analysis extended the investigation, showcasing commendable accuracy of decision tree, random forest, and Light Gradient Boosting Machine (LGBM) models. Random Forest models, in particular, demonstrated robustness in predicting consumer behaviors across various class labels. Feature importance analysis shed light on crucial determinants of consumer behavior, offering guidance for businesses and policymakers. Factors like availability, brand value, user sentiments, and security considerations emerged as pivotal influencers, providing a roadmap for future strategies aimed at enhancing user experience and satisfaction. Acknowledging the study's limitations, including inherent noise in real-world survey data and challenges in achieving higher predictive accuracy, the research contributes valuable insights to the evolving landscape of digital payment app adoption in Bangladesh. Building upon the present study, several avenues for future research can deepen our understanding and contribute to the evolving landscape of digital payments in Bangladesh like mitigating limitations associated with snowball sampling through controlled surveys and interviews to capture a representative sample with nuanced demographic distribution. Additionally, exploring alternative machine learning and deep learning models, alongside custom feature importance calculation algorithms, can unveil new perspectives on psychological factors influencing consumer behavior. Moreover, conducting a longitudinal study to track the evolution of psychological factors over time can provide insights into changing consumer preferences, concerns, and behaviors. Furthermore, extending the study to encompass cross-cultural comparisons, will contribute to a more comprehensive understanding of psychological factors and inform tailored strategies for different user segments. In addition to this, investigating the influence of external factors, such as macroeconomic conditions, regulatory changes, and technological advancements, on consumer behavior can add complexity and anticipate industry shifts. Lastly, drawing on principles from behavioral economics can help us delve deeper into the cognitive biases and heuristics influencing consumer decisions in the digital payment domain. Embracing these future research directions can foster a more comprehensive understanding of the psychological factors at play in the evolving digital payment landscape. Addressing identified limitations and exploring these uncharted areas will contribute valuable insights to academia and industry stakeholders. In the end, as Bangladesh embraces digital payment solutions, the gleaned insights from our study and future research endeavors will offer stakeholders valuable perspectives for refining strategies, improving user experiences, and fostering sustainable growth in today's digital era.

Bibliography

[1] D. Holt, A. Scott, and P. Ewings, "Chi-squared tests with survey data," *Journal of the Royal Statistical Society: Series A (General)*, vol. 143, no. 3, pp. 303–320, 1980.

[2] S. Mukherjee and P. Bhattacharyya, "Feature specific sentiment analysis for product reviews," in *Computational Linguistics and Intelligent Text Processing*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 475–487.

[3] M. H. Alam, M.-M. Rahoman, and M. A. K. Azad, "Sentiment analysis for bangla sentences using convolutional neural network," in *2017 20th International Conference of Computer and Information Technology (ICCIT)*, IEEE, 2017.

[4] T. U. Haque, N. N. Saber, and F. M. Shah, "Sentiment analysis on large scale amazon product reviews," in *2018 IEEE International Conference on Innovative Research and Development (ICIRD)*, IEEE, 2018.

[5] M. Cheng and X. Jin, "What do airbnb users care about? an analysis of online review comments," en, *Int. J. Hosp. Manag.*, vol. 76, pp. 58–70, 2019.

[6] A. M. Choudhury and K. Nur, "A machine learning approach to identify potential customer based on purchase behavior," in *2019 International Conference on Robotics,Electrical and Signal Processing Techniques (ICREST)*, IEEE, 2019.

[7] M. M. Jawad Soumik, S. Salvi Md Farhavi, F. Eva, T. Sinha, and M. S. Alam, "Employing machine learning techniques on sentiment analysis of google play store bangla reviews," in *2019 22nd International Conference on Computer and Information Technology (ICCIT)*, IEEE, 2019.

[8] V. Kameswaran and S. H. Muralidhar, "Cash, digital payments and accessibility," *Proceedings of the ACM on Human-computer Interaction*, vol. 3, no. CSCW, pp. 1–23, Nov. 2019. DOI: 10.1145/3359199.

[9] M. Lewis and M. Perry, "Follow the money: Managing personal finance digitally," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–14.

[10] J. Li, S. Pan, L. Huang, *et al.*, "A machine learning based method for customer behavior prediction," *Tehnički vjesnik*, vol. 26, no. 6, pp. 1670–1676, 2019.

[11] A. Rahman and M. S. Hossen, "Sentiment analysis on movie review data using machine learning approach," in *2019 International Conference on Bangla Speech and Language Processing (ICBSLP)*, IEEE, 2019.

[12] R. A. Tuhin, B. K. Paul, F. Nawrine, M. Akter, and A. K. Das, "An automated system of sentiment analysis from bangla text using supervised learning techniques," in *2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS)*, IEEE, 2019.

[13] J. Zhang, Y. Luximon, and Y. Song, "The role of consumers' perceived security, perceived control, interface design features, and conscientiousness in continuous use of mobile payment services," *Sustainability*, vol. 11, no. 23, p. 6843, 2019.

[14] Q. Zhong, S. Liang, L. Cui, H. K. Chan, and Y. Qiu, "Using online reviews to explore consumer purchasing behaviour in different cultural settings," en, *Kybernetes*, vol. 48, no. 6, pp. 1242–1263, 2019.

[15] T. Chen, L. Peng, X. Yin, J. Rong, J. Yang, and G. Cong, "Analysis of user satisfaction with online education platforms in china during the COVID-19 pandemic," en, *Healthcare (Basel)*, vol. 8, no. 3, p. 200, 2020.

[16] S. Fransiska, R. Rianto, and A. I. Gufroni, "Sentiment analysis provider by. u on google play store reviews with tf-idf and support vector machine (svm) method," *Scientific Journal of Informatics*, vol. 7, no. 2, pp. 203–212, 2020.

[17] A. M. Omar and N. Atteya, "The impact of digital marketing on consumer buying decision process in the egyptian market," *Int. J. Bus. Manag.*, vol. 15, no. 7, p. 120, 2020.

[18] S. Sharma, G. Singh, and S. Pratt, "Does consumers' intention to purchase travel online differ across generations?: Empirical evidence from australia," *Aust. J. Inf. Syst.*, vol. 24, 2020.

[19] S. Tripathi, "A study on adoption of digital payment through mobile payment application with reference to gujarat state," *International Journal of Trend in Scientific Research and Development*, 2020.

[20] V. Balakrishnan, P. Y. Lok, and H. Abdul Rahim, "A semi-supervised approach in detecting sentiment and emotion based on digital payment reviews," en, *J. Supercomput.*, vol. 77, no. 4, pp. 3795–3810, 2021.

[21] A. Gupta and R. Singhal, "Impact of covid-19 on digital payment services at towns and villages," *IJCRT2106045 International Journal of Creative Research Thoughts (IJCRT)*, 2021.

[22] M. T. A. Himel, S. Ashraf, T. A. Bappy, M. T. Abir, M. K. Morshed, and M. N. Hossain, "Users' attitude and intention to use mobile financial services in bangladesh: An empirical study," en, *South Asian Journal of Marketing*, vol. 2, no. 1, pp. 72–96, 2021.

[23] E. Ko, H. Kim, and J. Lee, "Survey data analysis on intention to use shared mobility services," *Journal of Advanced Transportation*, vol. 2021, pp. 1–10, 2021.

[24] F. A. Naim, "Bangla aspect-based sentiment analysis based on corresponding term extraction," in *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, IEEE, 2021.

[25] A. Pal, T. Herath, R. Dé, and H. R. Rao, "Why do people use mobile payment technologies and why would they continue? an examination and implications from india," *Research Policy*, vol. 50, no. 6, p. 104 228, Jul. 2021. DOI: 10.1016/j.respol.2021.104228.

[26] S. Shree, B. Pratap, R. Saroy, and S. Dhal, "Digital payments and consumer experience in india: A survey based empirical study," *Journal of Banking and Financial Technology*, vol. 5, pp. 1–20, 2021.

[27] S. Afroj, M. M. U. Hasan, and N. Fuad, "The who, when and why of uber trips in dhaka: A study from users' perspective," en, *Transp. Dev. Econ.*, vol. 8, no. 2, 2022.

[28] Á. Aldunate, S. Maldonado, C. Vairetti, and G. Armelini, "Understanding customer satisfaction via deep learning and natural language processing," en, *Expert Syst. Appl.*, vol. 209, no. 118309, p. 118 309, 2022.

[29] L. Kaewkitipong, C. Chen, J. Han, and P. Ractham, "Human–computer interaction (hci) and trust factors for the continuance intention of mobile payment services," *Sustainability*, vol. 14, no. 21, p. 14 546, 2022.

[30] M. S. Mahmud, A. Jaman Bonny, U. Saha, M. Jahan, Z. F. Tuna, and A. Al Marouf, "Sentiment analysis from user-generated reviews of ride-sharing mobile applications," in *2022 6th International Conference on Computing Methodologies and Communication (ICCMC)*, IEEE, 2022.

[31] K. Tabianan, S. Velu, and V. Ravi, "K-Means clustering approach for intelligent customer segmentation using customer purchase behavior data," en, *Sustainability*, vol. 14, no. 12, p. 7243, 2022.

[32] M. G. Albrecht, M. Green, and L. Hoffman, *Principles of Marketing*. 2023.

[33] A. B. Alok, H. Sakib, S. A. Ullah, *et al.*, "'khep': Exploring factors that influence the preference of contractual rides to ride-sharing apps in bangladesh," in *Proceedings of the 6th ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies*, 2023, pp. 43–53.

[34] A. Alsayat, "Customer decision-making analysis based on big social data using machine learning: A case study of hotels in mecca," *Neural Computing and Applications*, vol. 35, no. 6, pp. 4701–4722, 2023.

[35] C. Natuhamya, "Level weights for modeling with complex survey data," 2023.

[36] Statista. "Digital payments - bangladesh — statista market forecast." (2023), [Online]. Available: https://www.statista.com/outlook/dmo/fintech/digital-payments/bangladesh. [Accessed: 03-Jan-2024].

[37] C. Zhou, Y. Chen, R. Wattenhofer, and Y. Wang, "Print your money: Cash-like experiences with digital money," in *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023, pp. 1–7.

Appendix

GitHub Repository of All the Graphs:
https://github.com/Samonto-Karmaker/Undergrad-Thesis

Survey Questionnaire:
https://forms.gle/C7g7UY4ESECntAD27