Brand Reviews of e-wallet applications using Twitter sentiments

MSc Research Project
Msc Data Analytics

Snehal Ransing
Student ID: x19200714

School of Computing
National College of Ireland

Supervisor: Mr. Prashanth Nayak

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Snehal Ransing
Student ID:	x19200714
Programme:	Msc Data Analytics
Year:	2022
Module:	MSc Research Project
Supervisor:	Mr. Prashanth Nayak
Submission Due Date:	15/12/2022
Project Title:	Brand Reviews of e-wallet applications using Twitter sentiments
Word Count:	XXX
Page Count:	35

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Snehal Ransing
Date:	31st January 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Brand Reviews of e-wallet applications using Twitter sentiments

Snehal Ransing
x19200714

Abstract

Sentiment Analysis or opinion mining of the customer has become essential for a business to gain an upper hand in comparison to their competitors. To increase the footfalls of any product, it is important for a business to be aware of their current loopholes and the area they are lacking to provide services to the customer. Twitter being the most successful and popular platform of expression, users around the world grab the opportunity to comment and express their views. This study deals with capturing such tweets and analysing their expressions by judging the user sentiments regarding the product. Live twitter data is extracted for e-wallet apps used in India by making use of module with 35000 records, 7000 each for brand namely GooglePay, AmazonPay, Phonepe, Paytm, PayPal. Model development is done and evaluated using ML Models like Liner SVC (89% accurate), Logistic Regression (88% accurate), Random Forest (83% accurate), KNN-algorithm (56% accurate). Apart from ML, Deep Learning model namely BERT is used which provides an accuracy of 90% and Roberta provides 89% accuracy. With the model training and evaluation, Brand analysis is performed for negative areas where the outcomes are a brand must focus on marketing in certain regions, improve their interfaces and set up a proactive support team. Manual analysis also performed on the test data and predicted data to understand difference between accuracy.

Contents

1 INTRODUCTION

With everyday progression in the area of innovation, the present current time has moved from being exclusively dependent on paper to being dependent on the web-based implies. People have involved different payment techniques for buying items and administrations. Trading was quite possibly the earliest strategy; people traded labour and products as a trade-off for different labour and products Rampton (2016). The speed at which all the heritage frameworks are being modernized and improved has given a flood in the space of software development. There are n-number of programming now accessible to perform different exercises on the worldwide level, be it interacting with somebody or keeping a basic stockpiling area. The conventional technique for keeping up with and handling a financial balance has been moved to online means to some extent. Numerous on-paper undertakings have now been digitalized to stay aware of the speed of the developing world. Improvement of old software or development of new ones happens consistently, this shows that the advanced period is totally dependent on the innovation and online method for dealing with day-to-day tasks. According to Rampton (2016), in the year 1983, the philosophy of making payment a digital idea was proposed, which set apart as the start of electronic payment. Around a decade after the fact the first digital-based buy was made. Payment techniques have developed as a reaction for an expanded comfort interest; credit only payments satisfied this need from both client and shipper points of view Ondrus and Pigneur (2006)

Out of numerous Online-based uses, one mega utilization is the digitalized means of payment techniques. Not at all like conventional approach to making the payment, presently the greater part of the actually progressed nations has moved from making payment through money or coins to making payment through some internet banking or digital payment techniques. Web based banking or Internet banking is finished utilizing a completely tried and gotten sites created by the actual banks. On the other hand, in the event that a client wishes to proceed with the web-based banking through mobile, such facilities are likewise made accessible by the actual banks. There are some outsider applications which are dynamic for a huge scope to give the usefulness of digital payment, for example, Google Pay or Paytm. These applications are utilized to move the cash from a client to the banks, or from bank to a client, or from one client to another otherwise called peer-to-peer payments. There are various payment strategies utilized in the ongoing scene, for example, cash, cheque, Mastercard's, Debit cards, Electronic Payments, Internet Banking or Mobile Banking. Today every one of the previously mentioned payments techniques are well being used, yet relatively payment through cash is losing its ground slowly from the advanced world market. As per the research conducted by Singh and Rana (2017), most extreme clients of the digital payment applications are youngsters matured between 20-30 who are understudies or a functioning proficient in a confidential firm.

With such youthful age bunch clients as their customers who essentially favour performing out all kind of buys through web-based banking, the developers of the applications additionally get the chance to chip away at their constraints and upgrade their applications with added highlights. Digital Payment have been planned by the designers to make the payment interaction quicker and more secured. Some keys component of advanced payments or e-banking are: no lengthy cues, simple and quick cash move, less desk work, and no time utilization. With e-banking coming into picture, clients across the world have been sending cash as one companion to one more inside a negligible portion of seconds or minutes. The greater part of the global payment has become smooth because of the presence of web banking, for example when an international student wishes to pay the charge, the banks could utilize SWIFT exchange and make heavy payments conceivable.

In the year 2016, Indian Prime Minister Narendra Modi has brought vilification of 500-rupees and 1000-rupees notes in India with Immediate impact. This has given a dramatic ascent in acquiring Digital Payments on the roll in India. Presently, inside India, individuals utilizing multi-method for payments has upgraded by utilizing facilities like mobile, web, and smart watches. Transportation area in India have additionally moved from tolerating cash payments to acknowledges online payments. Significant vehicle organizations like UBER or OLA have inbuilt payments highlights checking out at the extraordinary possibility in future for online payment. Such transportation applications have additionally gotten themselves familiar with outsider applications like PAYTM, GOOGLE PAY, BHIM, PHONEPE, and so on.

Although since India is one of the fastest growing economies in the world today, along with dominating the technology and software sector, there are few set-backs the country experiences when rural regions are taken into consideration. According to the journal published by Cs (2017), problems faced by the rural India are quality of electricity or reach of digital cables in the rural regions. The journal also informs about other implications like accessibility to smart phones to every villager, and reach of ATM machines in the vicinity. As per the journal Cs (2017), there are nearly 712M debit cards in the village area which are used only 12 times in an ATM machine and just twice which swiping the card for making a sale.

1.1 Historical Context

In the radiance of controlling utilization and laundering of black money in the Indian currency market, the public authority of India has made a remarkable stride of cutting down the utilization of two highest denomination notes. This occasion occurred on 8th Nov, 2016 where the demonetization of 500 rupees notes and 1000-rupee notes were carried out with prompt impact. From that point forward, study led by Padiya and Bantwa (2018) gives the measurements of the use of e-wallets before the demonetization and post demonetization. According to Padiya and Bantwa (2018), by taking a gander at the govt. information there has been a critical surge in the administrations given by applications like MobiKwik, Paytm and Oxigen from 17 Lakhs on 8th Nov 2016 to 63 Lakhs as on 7th Dec 2016 (a leap of 271%). The value surge of such applications has additionally been huge enough with the leap of 267% from 52 crore to 191 crore INR

Post demonetization, the study directed by Narayanan (2020) uncovers those digital payments have arisen as a flat-out champ post demonetization wherein PAYTM has ended up being the best one acquiring traffic by 435% and expanding download by 200% in India. Individuals began accepting e-exchanges which prompted extent of e-banking through mobile banking and Immediate Payment System (IMPS) in India. According to the article, RBI determined the development of 175% and sum executed by 369%. Card payments have expanded by 133% with a day-to-day exchange count of 7 million by PAYTM worth 120 crores in a solitary day.

With the establishment and growth in mobile banking and e-wallet apps, the world was hit by COVID-19 epidemic turned pandemic. Cash transactions in mega countries like India who mostly makes transactions in cash were brought to stand still as everyone was asked to be quarantined. E-commerce sector of buying and selling goods and services have been the profitable ones for all sort of businesses like B2B (Business to Business), B2C (Business to Customers), C2B (Customers to Business) and C2C (Customer to Customer) as pandemic have encouraged people to buy and sell their stuffs more online rather in person Tripathi and Dave (2022).

1.2 Aim and Objective

The goal of this project is to analyse the data posted by users on social media platform like Twitter for payments methods using e-wallets, m-wallets, mobile banking, e-banking. Based on the posts written by the user on Twitter, business decisions will be taken to enhance the product by the makers to make it more user friendly and acceptable for all age group. This will enable customers to experience swift e-payment methods in India and shift from traditional cash payment methods to e-payment methods. With the aim of increasing more footfalls for using e-payments by downloading the required banking/third party m-payment apps from Play Store or App Store, the revenue of the organisation will be increased.

1.3 Research Question

- How well does Deep Learning dwell in analysing the customers' sentiment operating multiple digital payment applications on Twitter?
- Is Deep Learning a better option than Machine Learning to perform opinion mining on the review data obtained from Twitter for e-wallet users?
- Are Machine Learning and Deep Learning much more effective than IBM Watson for determining the sentiments of users regarding UPI payments?

1.4 Hypothesis

- Sentiment Analysis using Machine learning ends up with results having less accuracy than Deep learning.
- Deep learning can understand text data much better than Machine learning and hence will yield better results.
- Machine learning and Deep learning both performs better than IBM Watson for Sentiment Analysis.
- Deep learning can take better judgements than machine learning by reading and understanding the context of the text.

1.5 Scope

Since the study is based on the sentiment Analysis of the customers Twitter tweets of payment methods other than cash, the classification of sentiments will be carried out using users' opinion posted on Twitter. Different machine learning models will be utilized to perform the data analysis and the results will also be tallied with Deep learning frameworks. In scope Machine learning models for performing sentiment analysis is Naïve

Bayes, Support Vector Machines (SVMs), Logistic Regression, Random Forest algorithm. Models and frameworks to be used while performing analysis with Deep learning will include Distil BERT model using hugging face framework. Both the methodologies will be compared to check for a better accuracy and deduce the outcome as to ML or DL works well for opinion mining for driving to more accurate Business decisions. To dive deep with the motive of driving the business into a profitable direction, process of analysing the data with some graphs will be carried out.

1.6 Contribution to the body knowledge

As mentioned regarding the trend of e-wallet applications in the historical context section of the paper post demonetization in India, it can be learnt that previous work has been carried out on use of m-payments applications and its popularity among the younger generation. All scales of business have incorporated m-payment feature in their brand to deal with online monetary exchange. During the time of pandemic there were multiple downloads that underwent on play store for m-payment apps. But as the time passed and the effect of covid was reduced, higher age customers have started to step back and use the traditional method of monetary exchange i.e. cash payments This study will focus on the concerns that bothers the customers due to which they tend to leave the usage of these apps. Analysing all brands performances by judging the negative sentiments of the users on Twitter will bring a revolutionary growth in the business of the concerned m-payment brand, and also improve user experience. Taking into consideration the business objective of enhancing the growth and footfalls of the users, users with negative sentiments expressing their disagreement with the brand is where the focus area lies after having done with each brand analysis.

1.7 Outline

In Chapter 2, a brief history of the studies conducted by various author in the field of analysis is discussed. It covers significant topics like case study on digital payment, customer satisfaction and how Twitter is brought into use for making business decisions. Furthermore, studies which made use of Machine learning algorithms and Deep learning techniques have also been discussed focusing on digital payment. Chapter 3 informs about the CRISP-DM methodology utilized while conducting this study such as Business Understanding, Data Understanding and Data Preparation. Followed by Chapter4, which covers the modeling techniques and their evaluation used in Machine Learning and Deep Learning. In Chapter 5, the findings related to the analysis is depicted focusing on the weak areas of all the e-wallet brands which will help in boosting up the business. Chapter 6 concludes the final outcome of the study and briefly explains the understandings about the e-wallet brand in use today.

2 LITERATURE REVIEW

The following section outlined and reviewed the most popular methods used worldwide to analyse twitter sentiments, as well as the limitations of these methodologies.

2.1 Trend of Digital Payment

Post Demonization forced by Indian Prime Minister Narendra Modi, the country is pushing ahead with an expressed job of "Faceless, Paperless, Cashless" India Joshi (2017). From that point forward, there has been a surge in the utilization of digital payment through on the internet or mobile and has become extremely simple and helpful for individuals to use, in spite of which there has been sure obliviousness in embracing the payment method completely Nandal et al. (2020). Adjusting to the digital approaches to making an effective exchange, little nearby merchants and dealers have the talent of it however few are battling with adjusting the innovation with bother, for example, network issues [1]. Digitalization has likewise assumed a significant part in banking, where banks have thought of their own branding of m-payment, one of which is the State Bank of India (SBI), which has given YONO (You Only Need One) - mobile banking application impressively the most looked for applications by any Indian bank (MANDA, n.d.). The application has more than 50,000,000 downloads on play store. With tech becoming possibly the most important factor as the newbie in cash exchange area, there is likewise a gamble of fraud cases becoming an integral factor. Staying away from such fakes in the transaction is an extreme undertaking to achieve, the study has been conducted by Sarwan et al. (2021) where in the creator is utilizing BCT (Block Chain Technology) involving java for credit only economy. The creator has carried out BCT highlights like Decentralization, Visual cryptography, Hash algorithm, and encoded database, which will assist with following each transaction in BCT.

2.2 e-wallets/m-wallets/digital payment case studies

There have been numerous investigations performed to comprehend and improve the security of the payment methods. These examinations were completed to forestall, control and seize any fake occasions been polished for transaction. One such review was performed by Orche and Bahaj (2019) wherein the creators have utilized Machine Learning methods and joined with Ontology to forestall and control cheats. Creators have thought of a way to deal with foster a payment framework where in the commitment of ML will work out the score of the transaction to check whether the transaction made was a genuine one or a fraud one.

One more use of ML algorithm has been utilized for making examination on advanced digital payment applications by Chang et al. (2022). Here, the creators have broken down the steadiest model for fraud recognition in the digitalized payment strategies for Industry 4.0. Dataset utilized here was the genuine Mastercard dataset taken from Kaggle and examination was finished utilizing 5 ML models: Logistic Regression, Decision Tree, k-nearest neighbour, random forest and autoencoders. Their review has yielded extraordinary outcomes for Random Forest and Logistic Regression. Moreover, creators have additionally reasoned that performance improvement is likewise conceivable by accurately choosing the features in their examination utilizing PCA. The after effect of the test with original features and reduced features are looked at in the given picture Figure 1.

According to Hayashi (2012), users of the mobile applications do not have to incur additional charges for using the m-payment features as smartphones today are NFC (Near Field Communication) enabled. Despite of this, customers are even now more inclined towards the traditional payment methods using credit cards or cash. One among

[1]https://www.irjmets.com/uploadedfiles/paper/issue$_{5m}ay_2$022/25009/$final/fin_irjmets$1654353030.pdf

	Experiment 2	Accuracy	AUROC	Average Precision	Sensitivity	Specificity
Original features	LR	0.98	0.955	0.915	0.89	0.98
	KNN	0.99	0.905	0.824	0.85	1.00
	DT	0.97	0.952	0.859	0.87	0.98
	RF	0.98	0.957	0.906	0.84	1.00
PCA	LR	0.98	0.966	0.919	0.88	0.98
	KNN	0.98	0.933	0.865	0.84	1.00
	DT	0.98	0.939	0.870	0.84	0.99
	RF	0.98	0.969	0.925	0.85	1.00

Figure 1: Results of experimentation by (Chang, et al., 2022)

many reasons for this is security against fraudulent activities Hayashi (2012). Footfalls of the customers increases when the makes of the apps provide a guarantee of minimal fraudulent taking place on their medium only when a necessary action is taken in case of fraud pointing to risk-free usage of the product Mekovec and Hutinski (2012). Liébana-Cabanillas et al. (2014) have proposed two major reasons of changes in the behavioural acceptance of mobile payment by most of the customers which are security and privacy. According to Johnson et al. (2018), factors which enhances the perceived security in the product will increment the likelihood of adopting the product.

2.3 Customer Satisfaction

Various investigations on purchaser satisfaction have been led, with fluctuating outcomes. Consumer loyalty is viewed as a reaction that happens at a specific time and with a pre-determined concentration, despite the definitions' significant variations. Most of Authors like Cadotte et al. (1987) Westbrook and Reilly (1983) conceptualize happiness as a profound response. The level of emotional strength relies upon the circumstance; opinions related with delight can go areas of strength for from like happiness and energy to milder feelings like aloofness or unwinding Giese and Cote (2000). Seen risk influences adversely on apparent trust, satisfaction of the client and dependability with respect to mobile payment. Wang et al. (2015) states that trust is only a trade connection between a client and a dealer.

Wang et al. (2019) and others proposed number of Hypothesis and performed their analysis based on those hypotheses. The hypotheses are as follows: H1: "Security of the payment process impacts positively in customer satisfaction". H2: "Good value of service result in positive customer satisfaction".H3: "Considering customers' convenience results in positive customer satisfaction", H4: "Perceived usefulness has a positive customer satisfaction". H5: "Consumer satisfaction yields positive impact on consumer purchase" as shown in the figure Figure 2

2.4 Twitter data for Decision making

For a business to grow and flourish, marketing becomes the prime priority for any business in order to be successful. All sorts of small-cap to large cap e-commerce websites come up with a star rating for reviewing their product sold on their platform. To be more flexible, the e-commerce websites have in-text review systems other than star rating which gives their users the privilege of expressing themselves with the kind of experience they have after purchasing the product.

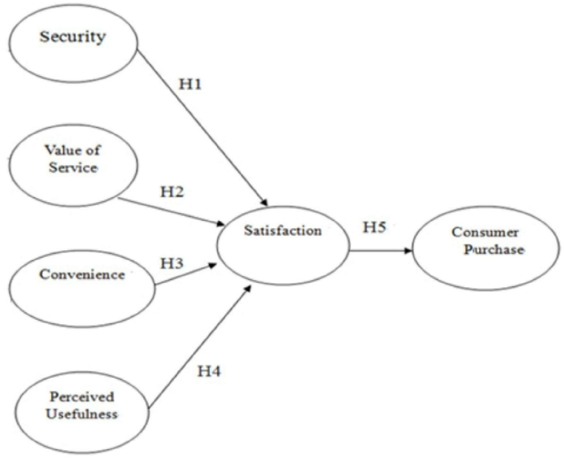

Figure 2: Impact on consumer purchase (Wang, et al., 2019)

One such analysis performed by Phan et al. (2021) where the authors have performed topic modelling on Twitter data to extract the tweets by passing the keywords by implementing LDA (latent Dirichlet allocation). Data cleaning and pre-processing is done and the final decision is made using satisfaction level using fuzzy decision tree. Another experiment is carried out by Gautam and Yadav (2014) wherein they have used labelled twitter data for the purpose of ease. After pre-processing and cleaning the dataset authors have used Naïve Bayes, Maximum Entropy, Support Vector Machine (SVM) and Semantic Analysis for analysing the data. The authors have used NLTK to train test the data of size 19340 out of which 18340 were used for training and 1000 for testing. Results of the analysis depicts the maximum accuracy for Semantic Analysis with the accurate rate of 89.9%

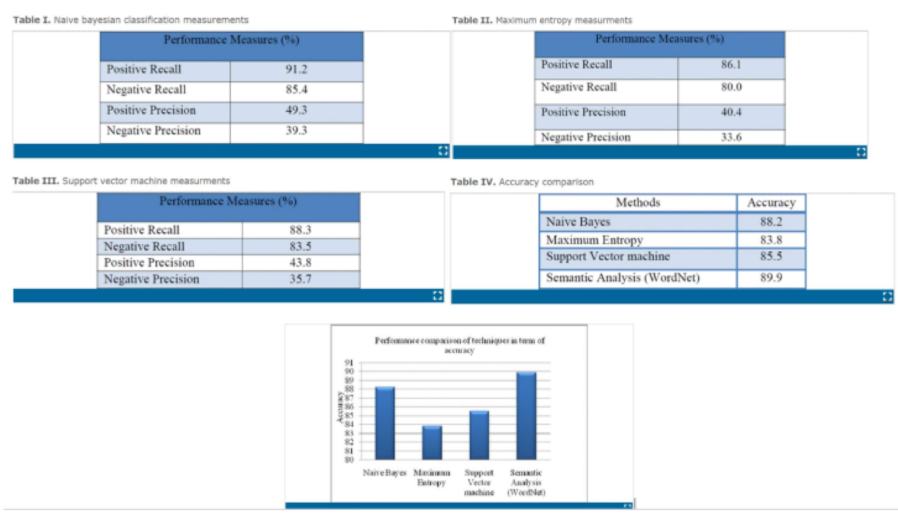
Table I. Naive bayesian classification measurements

Performance Measures (%)	
Positive Recall	91.2
Negative Recall	85.4
Positive Precision	49.3
Negative Precision	39.3

Table II. Maximum entropy measurments

Performance Measures (%)	
Positive Recall	86.1
Negative Recall	80.0
Positive Precision	40.4
Negative Precision	33.6

Table III. Support vector machine measurments

Performance Measures (%)	
Positive Recall	88.3
Negative Recall	83.5
Positive Precision	43.8
Negative Precision	35.7

Table IV. Accuracy comparison

Methods	Accuracy
Naive Bayes	88.2
Maximum Entropy	83.8
Support Vector machine	85.5
Semantic Analysis (WordNet)	89.9

Figure 3: (Phan, et al., 2021) analysis results for all ML algorithms

In other instance authors Wibawa et al. (2018) have extracted 2000 records of twitter data with the help of twitter official API to perform analysis on user satisfaction on service provided by their telecom operator in Indonesia. To carry out the experimentation, authors have used KNN technique with TF-IDF matrix and Part-of-Speech Tagging (POS). They have obtained the success with precision of 92.21% and recall of 93.74% and F1-score of 91.20% and Accuracy of 98.94%. The best value of K in K-NN is 20 with the accuracy of 97.65%.

Another such experiment is carried out by Mayasari and Hartanto (2019) where JNE twitter data is used for analysis purpose. JNE is a good delivery service organisation for which users have posted reviews regarding their services. Authors have used SVM to judge opinions of the users based on their tweets. Data cleaning process includes removal of hash tags and @ mentions. After the analysis 22.917% turns out to be positive tweets and 77.083% turns out to be negative tweets. The result is previously compared with the cosine similarity where closeness between the tweet and feature vector is determined, the result of which is depicted below after which evaluation of Support Vector Machine models that have been made using Kfold obtained an average value of 80% accuracy, 82.6% Precision, 80% Recall, and Fl-Score 80.1

Apart from Machine learning approach, research has been carried out by Liao et al. (2017) using Deep Learning method. Author is using DL in order to better understand the context underlying the tweet. Here, the appropriate reason for selecting the DL methodology of CNN – Convolutional Neural Network as it has a better ability to read the image and classify accordingly. Even though NLP extracts the text data features piece by piece, but due to the inability to correctly read the context of the text, the accuracy of the sentiment analysis isn't up to the mark. Here, the author has depicted that CNN performs better than SVM and Naïve Bayes. Dataset used in the experiment is MR (Movie Review) and STS Gold dataset which is the real Twitter dataset, post training the dataset, a No-SQL database name Mongo-DB is used to store the data.

Another application of Deep Learning for sentiment analysis on Twitter data is carried out by Hidayatullah et al. (2021) where user response on a politician is considered as a subject. Here Neural Network algorithms are used on the data to come up with the best algorithm. In this study the authors have used CNN (Convolutional Neural Network), Long short-term memory (LSTM), CNN-LSTM, Gated Recurrent Unit (GRU) -LSTM and Bidirectional LSTM. Apart from simply using DL algorithms, authors have also used and analysed the data using tradition machine learning algorithms like Linear Regression (LR), Support Vector Machine (SVM) and Multinomial Naïve Bayes (MNB). After comparison, the results depicted that Bidirectional LSTM has overshadowed other ML and DL algorithms with an accuracy of 84.6%

2.5 Sentiment Analysis on e-wallets using ML

Three different machine learning algorithms (Support Vector Machine, Logistic Regression, and K-Nearest Neighbor) and two vectorizing methods (Count vectorizer, Term Frequency Inverse Document Frequency (TFIDF)) have been compared and contrasted by Zahoor and Rohilla (2020) in order to analyze the tone of tweets about products. They've taken the necessary pre-processing measures to clean up the Twitter dataset. Logistic regression with the Count vectorizer proved to be the most effective algorithm/vectorizer combination. Through a modification to the vectorizer model, the SVM model attained the highest accuracy. Data accuracy may also be affected by how much information is used for training and how much is used for testing, according to studies. More train data compared to test data is used to improve study accuracy.

Support Vector Machine (SVM) analysis was used by Sulistyono et al. (2021) to examine Twitter data for Vaccine-19 sentiment. Crawling is employed to gather information for processing, and the text blob library is used for labeling the gathered information. A confusion matrix and K-fold cross-validation are used for the evaluation, followed by SVM with a linear kernel and RBF for the data analysis. The accuracy of the SVM- RBF kernel result was 85.88%, which was higher than the accuracy of the other kernel. They recommended using SVM's other kernel in addition to other techniques and labeling data to increase its performance matrix in future research.

Opinion or opinions of Jakarta's policies have been the subject of sentiment analysis by Saragih et al. (2021). Crawling is used to extract a dataset from Twitter, which is then used for analysis and Bag of Words and the TF-IDF method were used by the author for feature extraction. Support vector machine analyzes data that has been split into train and test sets using one of four possible splitting ratios. In this study, we evaluate the output of four alternative data splits by comparing the results obtained using various feature extraction methods. As a result, SVM with TF-IDF extraction plus the largest possible train data sample(90:10) achieved the highest accuracy (85.15%) of any combination tested. The findings would aid government data entry and the implementation of stricter regulations.

Nave Bayes, the Support vector Machine model (SVM), and long-term short Models for sentiment analysis on Twitter data that includes Indonesian government policies for Covid 19 have been presented Sujiwo et al. (2021). The classification and mining procedure has been applied to the lexical dataset (SentiWordNet). Data preparation is followed by data labeling and the separation of the dataset into train and test. Three SVM models (linear, RBF, and polynomial) with various kernels were trained using the SVM dataset. On the basis of a matrix that includes accuracy, precision, recall, and F1-score, all three models' results were compared. The SVM linear model was shown to perform better than both models, with 88.5% accuracy.

Kristiyanti et al. (2020) conducted study for the Indonesia zone in which they examined the opinions of various users of other UPI apps that were available on the Google Play store. Based on user reviews for the OVO and DANA applications, which totalled 2000 reviews in English language for both positive and negative ratings, the data was collected on the Google Play Store between March 2019 and January 2020. Naive Bayes algorithm and Support Vector Machine are two different types of approaches that have been used by researchers. The authors of the project have pre-processed the data using methods including tokenization, stemming, and the n-gram approach for text mining after extracting it via web scraping from the Google Play store. After conducting a careful analysis, the researchers came to the conclusion that the Naive Bayes algorithm did a good job of classifying the sentiments discovered from the data. The accuracy of Nave Bayes is around 94 percent, but the accuracy of SVM was only about 91%. Kristiyanti et al. (2020) has taken 500 positive and negative text from each OVA and DANA apps for classification purpose.

Using IBM Watson Analytics for social media, Maindola et al. (2018) conducted study on Indian UPI applications in the past. Data from a variety of online social media sites, including Facebook, Twitter, forums, videos, news, and blogs, has been utilised by the authors. They have examined the sentiments of numerous UPI payment applications, including BHIM, Oxigen, FreeCharge, Paytm, JioMoney, Phonpe, and Mobiwik. The primary goal of their research was to identify the best payment app that customers preferred after demonetization. For this analysis, the authors have taken data from 8 November 2016 to 7 November 2017 (a period of one year). The authors of the article came to the conclusion that BHIM and other applications like Mobikwik and Oxigen were more popular in India following the demonetization period than were attitudes for the Paytm app. The state with the most referrals was Maharashtra, which was then followed by Delhi, West Bengal, and Tamil Nadu.

2.6 Sentiment Analysis on e-wallets using Deep Learning

Tan et al. (2022) suggested a hybrid deep learning model that eliminates the sequence model's execution time constraints. The hybrid model consists of a robustly optimized BERT method and LSTM. IMDB, Twitter US Airline, and Sentiment140 datasets were utilized for analysis. The suggested model surpassed other cutting-edge approaches in terms of F1 scores and accuracy. To overcome problems like as lexical variety and an uneven dataset, the author applied the data-augmentation Glove word embedding pre-trained approach. According to the paper, data augmentation aided in improving the model's performance on the unbalanced dataset.

Aspect-based sentiment analysis on user comments from a variety of platforms, and it has been applied to four distinct domain datasets. Based on the feedback given by actual customers, Boumhidi et al. (2022) have calculated an objective numerical rating for the business in question. The study was broken down into four sections—collection of user feedback across platforms, span identification, elimination of spam data, use of an LCF-ATEPC model, and evaluation of the model's performance against state-of-the-art alternatives. Gain the worth of your reputation at last. The author has also done a system assessment to ensure the accuracy of the suggested technique, and user reviews show that it performs better than state-of-the-art models (AEN-BERT, BERT-BASE, BERT-SPC)

To determine the impact of the Covid19 epidemic on students and programs, Sandra et al. (2022) conducted a sentiment analysis using data from Twitter. Data was gathered with the aid of Tweepy and Pandas, both of which accessed the Twitter API. The data has been preprocessed in order to draw forth actionable insights. Based on the K-fold cross validation method with k set to 5, they suggest the RoBERTa (robustly optimized BERT approach) base model for evaluation. Additionally, the model has been adjusted for better performance. Ending up at 91.73% accurate with 83.33 percent precise. The suggested model received an F1 score of 86.42%, indicating widespread approval.

To analyze medicine evaluations and Twitter tweets sentiments, Sweidan et al. (2021) suggest a hybrid-Ontology XL-Net and Bi-directional Long short-term memory (Bi-LSTM) approach. They've utilized two algorithms and six datasets in their discussion. There has been feature extraction using ontology and the XLNet technique. The proposed XLNet approach achieved better results than the existing state-of-the-art techniques. They have employed Bi-LSTM for categorization. The ontology-XLNet combination outperforms the competition in terms of F-measures by 3.02% on the n2c2 2018 dataset and 2.6% on the WebMD dataset. It has been said that a multilingual model would be used in future work since XLNet has difficulties when it comes to analyzing extended contexts.

2.7 Summary

To briefly summarize the literature review, there have been various studies conducted in the field of data science and data analytics to understand the behaviour of e-wallet applications across the world. Many of the mentioned authors have made use of machine learning models. To analyse the sentiments of users there are multiple social media platforms available, the stated Literature review is focused on how Twitter has played a vital role in enhancing the business growth. As per the stated literature review, it's learnt that happiness of the customer impacts business development significantly. For studies related to e-wallet data the machine learning models that were mostly used were Naïve Bayes and SVM which gave the authors an accuracy of over 80-90%. Whereas, authors who conducted their study using Deep learning models like CNN, LSTM, GRU-LSTM and Bidirectional LSTM have got an accuracy rate of around 84%

3 METHODOLOGY

Referring to the standard IT solutions, the trickiest bit in a data science project is structuring the data and maintaining a flow of data. For a data science project to be termed as a successful experimentation, one of the many challenging factors are its productionization and maintainability. Instead of considering and working on the overall Data Science life cycle, the researchers today are more focused on improvising the results that are achieved from machine learning models nowadays. By analysing the advancements made by automation sector today, industries nowadays are more focused on improvising the outcomes of already existing models as per business needs. In order to organize and perform data mining projects, the widely accepted methodology is CRoss-Industry Standard Process for Data Mining, also popularly abbreviated as CRISP-DM. It is profoundly adopted and considered to be the top Data mining method for a complicated Data Science project Danmon Projects (2022)

The prominent focus point area of the researchers is to enhance the effectiveness in the project management and comprehensive documentation of the driving factors to attain success in Knowledge Discovery in Database (KDD) projects Tan et al. (2000). Even though there were several efforts made to bring in multiple techniques to manage data mining projects, there are multiple lures happening in a Data Mining project which indicated the lack of methodology for project development Moss and Atre (2003), and to overcome this lack in development technique, KDD approach is still preferred for unstructured data on an ad hoc basis Becker and Ghedini (2005) Nadali et al. (2011).

The official document of CRISP-DM is published by the CRISP-DM consortium which provides a step-by-step data mining guide Chapman et al. (2000). To proceed with the proposed agenda of the said thesis, CRISP DM methodology is used to carry out the analysis. To process business data and perform analysis, CRISP-DM is widely recognized as the prominent technique to help businesses excel and achieve business goals.

The steps involved in CRISP-DM methodology are as followed:

- Business Understanding
- Data Understanding
- Data Preparation
- Modeling

- Evaluation
- Deployment

Here, in this Chapter, the key points related to Business Understanding, Data Understanding and Data Preparation is discussed. The latter phase in methodology is discussed in the following chapters.

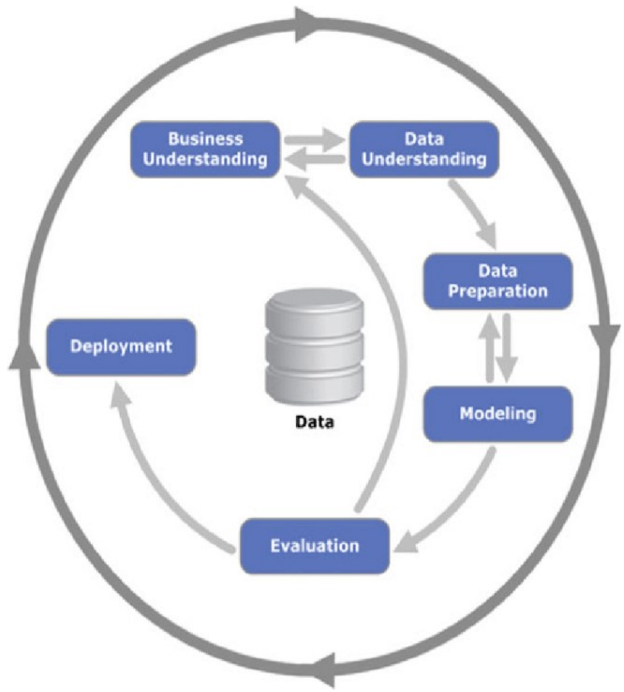

Figure 4: CRISP-DM methodology (Chapman, et al., 2000)

3.1 Business Understanding

For all Data analytics or a data science project, defining the business goal is the first and foremost step in the process. This process starts with identifying the existing business problems and finding an effective solution to them. Identifying the business problem becomes important as to when the complications in the business is not clearly defined at an earlier stage then that may cause influential complexity and challenges to problem resolutions.

The process of identifying an authentic business issue resembles to the well-known "divide-andconquer" algorithm in computer science. The prerequisites examination expected to do such a cross-hierarchical plan not just comprises of an objective investigation, wherein the pertinent vital objectives of the taking part organizations are adjusted, yet in addition of a worth investigation, in which the business sup-portability of the star grouping is investigated Gordijn et al. (2006).

It becomes essential to align with the business stakeholders during the process of gathering business requirements and working on possible resolutions to curb all the potential bottlenecks that may lead to an absolute project failure. The business stakeholders can be a direct end consumer, customer, or a user when the business model is a B2C (Business to Customer) business model. The other business model is B2B (Business to Business) model where the business will be aligned with the consumer indirectly. Their main agenda turns out to provide a better business outcome by aiming at the optimization of technical process such as increasing the cost efficiency, satisfaction of the user or enhancements.

The later significant aspect of business understanding is to inform the end users regarding the success criteria of the proposed business model. The business model needs to be tracked precisely during its execution depending on the countable criteria, which are also mostly known as Key Performance Indicators (KPIs).

3.2 Business Understanding of this project

The business goal of this project is to come up with a business decision using the sentiments of the users using their reviews posted on the social media. For Mobile payment apps used in India or other parts of the world, users around the world have mixed sentiments regarding their use and what the product has to offer. Depending on the user experience, the objective of this study will be to check whether the mobile payment experience was satisfying or dissatisfying. The research is focused on clarifying the ambiguity in the users' reviews and focus on the negative or dissatisfying experience to come up with a logical solution to increase the footfalls of their mobile apps in the market. This study is conducted based on multiple insights obtained from the twitter users posting their reviews regarding the products good points and the bad points.

The business model of the conducted study is depicted in the below image Figure 5.

3.3 Data Gathering

Working as per the business requirement, a very large set of data was scraped through social media, specifically Twitter, for this study. Scraping data from twitter can be performed in multiple ways, for instance, using an API (Application Programming Interface) or directly web scraping. In many cases, twitter data is also obtained from a pre-stored database repository. Specially for this study, live data was scraped from twitter since the mobile payment applications that are used into analysis are widely spread across India. To put lights on the horizons of the usability of mobile payment apps or e-wallets, many of these apps are also used outside India to achieve its worldwide distribution such as GooglePay.

It becomes essential to collect the live tweets since mobile apps are enormously used on a daily basis. Due to this, these apps have a decent amount of user base across the country and world, which in turn helps in gathering data that are related to currently existing issues in those mobile apps. Instead of working on the old dataset for analysis, the project focuses on fresh dataset that are posted a few couple of days ago on twitter.

By extracting live tweets, the project is focusing more on analysing the current bottleneck in the business of these mobile payment apps. For this study, e-wallet applications like GooglePay, AmazonPay, Paytm, BharatPay, Phonepe are used. Here, the technique used to extract application specific records or tweets is by looking for @mentions in the tweet.

There are a few packages accessible to get tweets from Twitter. A couple of them are recorded underneath:

- Twint

- Tweepy
- Snscrape
- Octoparse
- GetOldTweet3

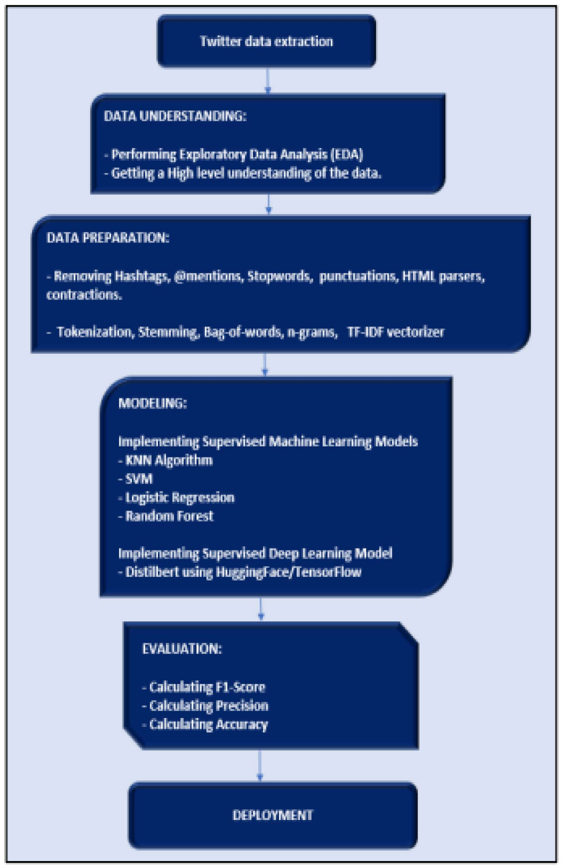

Figure 5: Flow Diagram

Out of the above-recorded packages, Twint is a tweet scraping tool that is openly open for use. The essential benefit of using Twint is that it requires no Twitter login. The computer programmer using the pack isn't supposed to make any record on Twitter to play out the examination. Involving Twint accompanies specific limits too, for example, Twint is challenging to install for everybody and has a specific measure of count limitation with regards to querying a catchphrase[1].

Talking about different libraries like GetOldTweet3, there have been restricted updates accomplished for quite a while and has become less pertinent for the scrubbers to utilize this library. As most recent as 27th Nov 2019, this library hasn't been refreshed in any event, for complying with Twitter rules, this makes GOT pretty less significant for this venture[1].

[1] https://medium.com/dataseries/how-to-scrape-millions-of-tweets-using-snscrape-195ee3594721

However long Octoparse is viewed as utilized as a medium for scraping information for an analytical reason, it should be bought as Octoparse is not an open-source library available for all. In addition, Octoparse is extensively delayed in scratching the web data [1].

Another choice is by involving Tweepy as the library which utilizes Twitter API. Be that as it may, add to its restrictions of separating information just until the previous week. Besides, it likewise has a cut-off count of almost 3200 tweets on a single scrape. To gain admittance to Tweepy as a library, one priority a designer account made to accept their entrance token and security key. These snippets of data become vital for utilizing tweepy and accordingly give limitations in the event that the engineer account demand is dismissed from the Twitter end[1].

Taking into account the above restriction in utilizing different libraries, this study will adhere to the Snscraper library for getting the tweets. Utilizing Snscraper it turns out to be exceptionally simple to scrape the information with no limitations as far as possible and it doesn't need a software engineer to make a developer account. Utilizing snscraper, almost 5000 records are extracted for each application specifically GooglePay, AmazonPay, Paytm, Phonepe, BharatPay. The information is removed in all the languages without giving any regional language bar at the time of query. All the data gets stored into a csv file named 'ewallet$_t weets.csv'$.

3.4 Data Understanding

There exist various ways of investigating the data for factual and ML procedures that are to be applied to the current issue that should be tended to. For example, OntoLearn is a text mining device that yields higher efficiency during ontology development Velardi et al. (2001). ML for NLP and text examination includes a bunch of tangible methods for distinguishing grammatical characteristics, elements, opinions, and different parts of the text. Thus, with regards to NLP, the most wellknown NLP ML procedures and calculations are introduced in figure Figure 6., this study is completed utilizing a portion of the beneath recorded algorithms which is momentarily examined going further.

Supervised Learning methodology and models
CLASSIFICATION
1. 1. Linear SVC
2. 2. Logistic Regression
3. 3. k-NN algorithm
4. 4. Random Forest
DEEP LEARNING
1. LSTM (Long Short-Term Memory)
2. BERT (Bi-directional encoder Representations from Transformers
3. RoBERTa

Figure 6: Models

There are 30000 records fetched from twitter using 17 distinct columns which contains information related to the user such as their ID, location and the tweets. Some system

[1]https://medium.com/dataseries/how-to-scrape-millions-of-tweets-using-snscrape-195ee3594721

information is also extracted such as number of likes, retweets, etc. There is total 50 distinct languages of tweets, out of which English, Hindi are few languages among top 10.

Figure 7: Trending Hashtags

3.5 Data Preparation

This stage chiefly centres around progress in the nature of data determined to guarantee the improvement of results in the next stage. This is experimentally viewed as the lengthiest stage in the whole course of model execution. Setting up the data almost adds to 50-70% of complete efforts Chapman et al. (2000), numerous multiple times it could likewise take almost 80% of absolute planning time Duhamel et al. (2003) or even above 90% in some cases. The sole justification for effective and significant model execution time is the rationale of getting perfect and predictable data Zhang et al. (2003).

There are numerous moves toward pre-processing the data prior to making that data accessible for a model to process and act. One of the main advances is to deal with NULL values or missing values. Missing values in the dataset enhance the volume and nothing better could be derived from them. More often than not when there are occasions of data missing from any column then, at that point, taking care of those missing records are finished by taking the mean or mode of the whole segment. Occasions where there are categorical values present in the data, for the most part when the information comprises of text rather than numbers, a few explicit words like "unavailable" could be utilized to top off those holes. There are a few tools accessible that automates this cycle, yet this interaction should be performed with absolute attention to detail, since erasing or skipping the data might prompt potential data loss moreover. A sure round of EDA should be performed in the wake of cleaning the information to keep up with and check for its relevancy.

The additional technique that is utilized during data cleaning stage is feature engineering where - in a greater number of columns are gotten from the all-existing ones. Hardly any models could be changing the example of the information or arrangement of the information. Occasion of removing a specific line of data from a generally existing information and stacking it into an alternate segment could be considered as other illustration of feature engineering. There are examples where the dataset contains large number of columns which aren't important to direct data analysis, during such cases a strategy called dimensionality reduction can be performed to dispose of the superfluous sections and simply keep the significant ones.

Typically for NLP project, a text dataset can exist in a form of structured, semi-structured or unstructured state. Interpreting such data becomes challenging. The data cleaning approach followed during this project is listed below:

```
print(cust_data.columns)
print('''
''')
cust_data.info()

print('''
Rows and columns length:''',
cust_data.shape)
```

```
Index(['application', 'date', 'content', 'userid', 'username', 'displayname',
       'followersCount', 'friendsCount', 'location', 'replycount', 'likecount',
       'retweetcount', 'language', 'source', 'mentionedusers',
       'retweetedtweet', 'hashtags'],
      dtype='object')


<class 'pandas.core.frame.DataFrame'>
RangeIndex: 30000 entries, 0 to 29999
Data columns (total 17 columns):
 #   Column          Non-Null Count  Dtype 
---  ------          --------------  ----- 
 0   application     30000 non-null  object
 1   date            30000 non-null  object
 2   content         30000 non-null  object
 3   userid          30000 non-null  object
 4   username        30000 non-null  object
 5   displayname     30000 non-null  object
 6   followersCount  30000 non-null  object
 7   friendsCount    30000 non-null  object
 8   location        30000 non-null  object
 9   replycount      30000 non-null  object
 10  likecount       30000 non-null  object
 11  retweetcount    30000 non-null  object
 12  language        30000 non-null  object
 13  source          30000 non-null  object
 14  mentionedusers  30000 non-null  object
 15  retweetedtweet  30000 non-null  object
 16  hashtags        30000 non-null  object
dtypes: object(17)
memory usage: 3.9+ MB

Rows and columns length: (30000, 17)
```

Figure 8: Data Understanding A

3.5.1 LANGUAGE CONVERSION

For this dataset obtained, there are 50 different languages of data extracted. For Vader sentiment to classify the text, it requires English language dataset, hence the languages of most of the vernacular medium is converted into English. Example of code is shown below:

```
print("Number of dictinct langauge tweets present in dataset:",len(pd.unique(cust_data['language'])))

print(''' -------------------------
Top 10 langauge tweet count:
 ''',
cust_data.groupby(['language']).language.count().sort_values(ascending=False).head(10))
```

```
Number of dictinct langauge tweets present in dataset: 50
 -------------------------
Top 10 langauge tweet count:
  language
en     22901
qme     1681
hi      1620
und      861
qam      478
in       416
tl       320
es       315
ja       304
et       183
Name: language, dtype: int64
```

Figure 9: Data Understanding B

3.5.2 CASE-CONVERSION

The data extracted in the text format contains various formats. Mostly all the text in the tweets follow camel casing in general. Some texts could be written in complete capital letters to express and emphasise on some events. These kinds of data need to be handled immediately after extraction. To deal with case handling of the dataset, the entire text can be converted into upper case or lower case. In this study, the entire tweet is read one at a time using a lambda function and the string is converted from multiple case to lower case.

Case Conversion	Extracted Data	Converted Data
Camel Case	"Extracting Tweets Using Python Is Fun"	"extracting tweets using python is fun"
Upper Case	"EXTRACTING TWEET USING PYTHON IS FUN"	"extracting tweets using python is fun"

Figure 10: Case Conversion

3.5.3 REMOVING PUNCTUATIONS, HASHTAGS AND @MENTIONS

To express emotions in a text, the text has to be written in a certain way which includes a use of punctuations. For instance, to express a query "?" is used, and to express anger or Joy an exclamation "!" is used. But such punctuations are not understood by any machine learning model, so to make the text much understandable to the NLP algorithm, all sorts of punctuations are removed and the text is converted into a clean and plain text.

Extracted Data	Converted Data
"Extracting tweets in it's raw form using Python, Snscrape is FUN!!"	extracting tweets in its raw form using python snscrape is fun

Figure 11: without Punctuation

To extract data from Twitter there are several ways. The ones which is used in this study is by utilizing @mentions in the text. The value for @mentions is provided in a list of string which is iterated and the data is extracted. Those extracted tweets consists of @mentions which were used in the query. Similar to the concept of punctuations, hashtags and @mentions are also considered to be punctuations but could not be cleaned during punctuation cleaning. Since hashtags '' and '@' mentions could not be treated as punctuations, these are taken up as a part of string handling or string replacement. But since in this study, the payment app name become essential to store, and @mentions are included as a part of punctuation handling itself.

3.5.4 REMOVING STOPWORDS

Stopwords are the words that are very frequently used in a common text. Those word which does not add any value to the statement and are only used for a good grammar are treated as stopwords. For any text to be understandable by a human, stopwords are necessary to be included that helps in framing the sentence. But while analysing the sentiments in a text, only those words are treated considerable from which any meaning could be derived. Words like ' if, but, the, you, me, is, etc. adds no value to the text in analysing the sentiments. So these words are removed at the time of data cleaning.

3.5.5 TOKENIZATION

To read the data and check for the occurrence of a particular word in the document, the entire text is converted into number of tokens. Usually, tokenization is done to achieve the most important words that are used in a corpus of text. It is beneficial to tokenize the string after removal of stopwords. In this study the data is tokenized using the tokenization technique and iterated over a loop in the function. Using the lambda function tokenization is achieved for each document in the corpus. The tokenized sentence looks like as follows:

3.5.6 DATA VECTORIZATION

Data vectorization is a cycle where the text is changed over into numbers. The motive to vectorize the Data is to make the text more reasonable to the machines. Giving contribution as text can't be treated as a coherent boundary for the machine to deal with it and subsequently, data vectorization comes in to picture. There are numerous approaches to vectorizing the information which is completed in this review.

- Count vectorization
- TF-IDF vectorization
- n-gram countvectorizer

Count vectorizer - Count vectorizer is a strategy where the vectorizer will attempt to change over each words in the text data into a number. Here, the value is N = 1 as a matter of course. Count vectorizer is unequipped for demonstrating the significance of the word in a document or in a corpus according to a more noteworthy point of view. Each word in the message is tagged as 1 on the off chance that the message is available in the document. At the point when the message isn't present in the document then it is tagged as 0.

n-gram countvectorizer approach is precisely like count vectorization. The main contrast between count vectorization and n-gram approach is the value of N. Like the value of N in Count vectorization is utilized as 1 of course, the incentive for N in N-Gram approach can be changed according to the requirement. Here,N signifies the gathering of words and regarding the assembled word as one token. In this review, n-gram technique utilized is bigram, for example the value for n=2. This value can likewise be changed to 3 and termed as tri-gram. For n=1 the term is uni-gram, n=2 the term is bi-gram and n=3 the term is tri-gram.

TF-IDF vectorization strategy is utilized to detect the significance of the word by calculating weight of each word in the document or corpora. The more the times the word shows up in the section, the higher value gets appointed to the word. The word with higher value means the word is more significant than other present words.

3.6 Sentiment Classifier

The course of sentiment characterization of a word can be done in more ways than one. There are numerous vocabularies used to label a message sentence and judge its opinion. Those are:

- Vader Sentiment
- TextBlob Sentiment
- Flair

Here, in this study, VADER sentiment analyser is utilized to isolate the vocabularies in view of the text. The manner in which Vader sentiment characterization works is by checking the and matching each word referenced in the document by the word put away in its word reference. Vader sentiment in the python for a sentence goes from - 1 to +1 [1]. - 1 being the negative opinion and +1 being the positive sentence. The most fascinating component of Vader Sentiment is it can decipher the upper casing of the words utilized in the sentence to underline an inclination. Some of the time to show exceptionally serious articulation, users regularly utilize accentuations like interjections imprint or two-fold statements, these varieties in the sentences are additionally perceived by the Vader vocabulary. For this review, the texts are characterized into three vocabularies named NEGATIVE, NEUTRAL, POSITIVE so at the hour of text cleaning, punctuations and case transformations are completed and the classification is exclusively finished based on texts or words.

Polarity	Most Negative Replies on googlepay Handle
negative	googlepay today pm scammed poor street vendor rupee customer messaged vendor said transferred money showing message quite big street vendor day income
negative	phonepe people misappropriated rupee small googlepay problem
negative	worst experience googlepayindia time need transfer large urgently googlepay fails deliver pleasant experience money deducted bank reach anothe gpay account time phonepesupport work
negative	pain family try coordinate bank responds bank feeling helpless help googlepay icicibank axisbank
negative	surprised srapper complained brought attention googlepay axis upi blocked reported googlepay axisbank

Polarity	Most Positive Replies on phonepe Handle
positive	saurabh duas paytm paytmcare phonepe phonepesupport phonepe safety upi transaction limit hour
positive	phonepe phonepesupport hello want say thank using app year really like app ui mean user friendly request add dark mode feature important
positive	yusufzaizahid phonepe googlepay matter
positive	lot respect phonepe particularly love phonepe pulse launched recently kudos phonepe
positive	making transaction phonepe july partial got deducted flipkart supercoins r deducted supercoins option select deselect payment mode supercoin phonepesupport giving irrelevant reason

Figure 12: Classified Sentiments

[1]https://medium.com/@piocalderon/vader-sentiment-analysis-explained-f1c4f9101cd9

3.7 Summary

To Summarize, it is prominent to have a clear understanding of the Business prior initiating any work on it. To avoid any sort of end-moment mishappening with the project, the developers are expected to timely connect and update the Business on progress. For any business problem to be solved, the understanding of data is required which is done by analysing the data by performing exploratory data analysis. Post understating the data, it has be to pre-processed for bringing the data into best usable state for modeling process. Once the data is processed and cleaned, using Vader sentiment and labelling the data to be fed to Machine Learning and Deep Learning models.

4 DESIGN

For the research project we have followed below methodology approach.

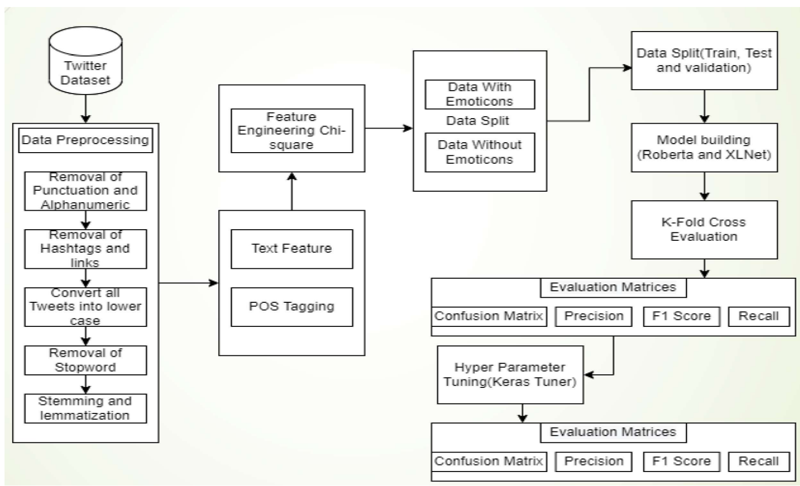

Figure 13: Methodology

5 MODELING AND EVALUATION

Post-fulfilment of the past stage comes to the model building stage. In this stage, model plans are executed in view of the necessity. The point of carrying out the model is to upgrade the business by foreseeing the most dependable outcomes. There are various kinds of models utilized in this study to examine the opinions and concoct the best accuracy rate.

Those models are as per the following:

- KNN Algorithm
- Logistic Regression
- Linear SVC
- Random Forest

- Bert
- RoBerta

5.1 knn-algorithm

As per Kumar (2020), KNN is one of the least complex algorithms that can be utilized for taking care of the classification problems. It isolates the data in light of how its neighbours are classified. The data with comparable sort of values are isolated together and named as one, and remaining information focuses are similarly isolated and marked. Here, in this study, the value for k is taken as [3,5,7]. The justification for picking all the odd adjoining point is fundamentally to try not to any kind of get between two distinct classes incorrect.

```
def  classification_algorithm(X_train,X_test,y_train,y_test,algorithm_type,vectorizer_type):
  if algorithm_type == 'knn':
    tuning_parameter  = [3,5,7] #knn tuning parameter in number of neighbours
    for value in tuning_parameter:
      knn = KNeighborsClassifier(n_neighbors= value)
      knn.fit(X_train, y_train)
      knn_pred = knn.predict(X_test)
      key = 'knn with ' + str(value) + ' neighbours ' + '(' + vectorizer_type + ')'
      value = accuracy_score(y_test, knn_pred)
      accuracy_list.append({'algorithm':key,'accuracy': value})
      print("Classification report for K-nearest neighbour model with tuning parameter",value,
            "- \n{}:\n{}\n".format(knn, classification_report(y_test, knn_pred)))
```

Figure 14: Knn Algorithm

Post evaluating knn – algorithm, it is learnt that the accuracy for knn is maximum (57%) when the nearest neighbour considered as 3 for count vectorizer. Whereas for Tf-idf the maximum accuracy for knn achieved is 45% for 3 nearest neighbours and n-gram count vectorizer attains 47% for 3 nearest neighbours.

```
feature extraction is completed using countvectorizer
Algorithm is trained using countvectorizer
model is training with knn algorithm
Classification report for K-nearest neighbour model with tuning parameter 0.5688333333333333 -
KNeighborsClassifier(n_neighbors=3):
              precision    recall  f1-score   support

          -1       0.59      0.27      0.37      1425
           0       0.46      0.96      0.62      1860
           1       0.87      0.46      0.60      2715

    accuracy                           0.57      6000
   macro avg       0.64      0.56      0.53      6000
weighted avg       0.68      0.57      0.55      6000


Classification report for K-nearest neighbour model with tuning parameter 0.5466666666666666 -
KNeighborsClassifier():
              precision    recall  f1-score   support

          -1       0.62      0.22      0.33      1425
           0       0.43      0.98      0.60      1860
           1       0.88      0.42      0.57      2715

    accuracy                           0.55      6000
   macro avg       0.64      0.54      0.50      6000
weighted avg       0.68      0.55      0.52      6000


Classification report for K-nearest neighbour model with tuning parameter 0.5231666666666667 -
KNeighborsClassifier(n_neighbors=7):
              precision    recall  f1-score   support

          -1       0.64      0.18      0.28      1425
           0       0.42      0.98      0.58      1860
           1       0.87      0.39      0.54      2715

    accuracy                           0.52      6000
   macro avg       0.64      0.52      0.47      6000
weighted avg       0.67      0.52      0.49      6000
```

Figure 15: Classification Result KNN

5.2 Logistic Regression

One more classification algorithm utilized in this study is Logistic regression. The reason for utilizing Logistic regression is because of the benefits presented by the algorithm while model execution. A few of the benefits as recorded by (Rout, 2022) are the convenience, and the speed at which the classification is finished by this model. When the dataset comprises of huge number of dimensions, then there are chances of over-fitting occurring in the model, which can be tried to be avoided by utilizing L1 and L2 regularization techniques, in any case when the dataset has negligible aspects then Logistic regression is more unlikely to over-fit and yield improved result (Rout, 2022). To stay away from the gamble of over fitting the model, parameter tuning is performed by passing the value of C as [0.01, 0.05, 0.25, 0.5, 1]. By Iterating over these values during the model execution, the best-fit model will be accomplished.

Post evaluating Logistic regression for all three vectorizations, it is learnt that count vectorizer attains a higher accuracy of 91% for C=1. Whereas for tf-idf and n-gram count vectorizer the accuracy is 89% and 90% respectively.

```
model is training with logistic_regression algorithm
Classification report for logistic regression model with tuning parameter 0.787 -
LogisticRegression(C=0.01):
              precision    recall  f1-score   support

          -1       0.87      0.52      0.65      1425
           0       0.73      0.90      0.80      1860
           1       0.81      0.85      0.83      2715

    accuracy                           0.79      6000
   macro avg       0.80      0.76      0.76      6000
weighted avg       0.80      0.79      0.78      6000

Classification report for logistic regression model with tuning parameter 0.8551666666666666 -
LogisticRegression(C=0.05):
              precision    recall  f1-score   support

          -1       0.90      0.67      0.77      1425
           0       0.80      0.95      0.87      1860
           1       0.89      0.89      0.89      2715

    accuracy                           0.86      6000
   macro avg       0.86      0.84      0.84      6000
weighted avg       0.86      0.86      0.85      6000

Classification report for logistic regression model with tuning parameter 0.8925 -
LogisticRegression(C=0.25):
              precision    recall  f1-score   support

          -1       0.91      0.76      0.83      1425
           0       0.85      0.96      0.90      1860
           1       0.92      0.91      0.92      2715

    accuracy                           0.89      6000
   macro avg       0.89      0.88      0.88      6000
weighted avg       0.89      0.89      0.89      6000

Classification report for logistic regression model with tuning parameter 0.9046666666666666 -
LogisticRegression(C=0.5):
              precision    recall  f1-score   support

          -1       0.92      0.80      0.85      1425
           0       0.87      0.96      0.91      1860
           1       0.93      0.92      0.92      2715

    accuracy                           0.90      6000
   macro avg       0.90      0.89      0.90      6000
weighted avg       0.91      0.90      0.90      6000

Classification report for logistic regression model with tuning parameter 0.9146666666666666 -
LogisticRegression(C=1):
              precision    recall  f1-score   support

          -1       0.93      0.81      0.87      1425
           0       0.88      0.96      0.92      1860
           1       0.93      0.93      0.93      2715

    accuracy                           0.91      6000
   macro avg       0.91      0.90      0.91      6000
weighted avg       0.92      0.91      0.91      6000
```

Figure 16: Classification Result Logistic Regression

5.3 Linear SVC

The goal of a Linear SVC (Support Vector Classifier) is to fit to the information you give, returning a "best fit" hyperplane that partitions, or sorts, your information. From that point, subsequent to getting the hyperplane, you can then take care of certain highlights to your classifier to see what the "anticipated" class is. This makes this particular calculation somewhat reasonable for our purposes, however you can involve this for some circumstances.

Post Evaluating Linear SVC model, the results obtained were quite remarkable in accuracy. For Count vectorizer the accuracy is maximum approaching at 93% for C=1 and maximum iteration set to 100. For tf—idf also the accuracy is 92% for C=1 and for n-gram count vectorizer the accuracy is 91%.

5.4 Random Forest

The following classification algorithm utilized in this study is Random Forest. Random Forest is additionally viewed as one of the most amazing algorithms with regards to characterizing texts. Random forest capabilities in a way by producing results for different number of trees. Here, in this review, there are different trees made and that multitude of values are passed as a parameter tuning for Random Forest. The values used are [5, 10, 15, 20], the values in the list signifies the number of trees the model is considering at the time of single execution.

Post evaluating the Random Forest model, the accuracy for 20 estimators is 84% for count vectorizer, 84% for tf-idf vectorizer and 83% for n-gram count vectorizer.

```
model is training with linearsvc algorithm
Classification report for Linear SVC model with tuning parameter 0.8613333333333333 -
LinearSVC(C=0.01, max_iter=100):
              precision    recall  f1-score   support

          -1       0.90      0.69      0.78      1425
           0       0.80      0.96      0.87      1860
           1       0.90      0.89      0.89      2715

    accuracy                           0.86      6000
   macro avg       0.87      0.84      0.85      6000
weighted avg       0.87      0.86      0.86      6000


Classification report for Linear SVC model with tuning parameter 0.9003333333333333 -
LinearSVC(C=0.05, max_iter=100):
              precision    recall  f1-score   support

          -1       0.92      0.78      0.85      1425
           0       0.86      0.97      0.91      1860
           1       0.92      0.92      0.92      2715

    accuracy                           0.90      6000
   macro avg       0.90      0.89      0.89      6000
weighted avg       0.90      0.90      0.90      6000


Classification report for Linear SVC model with tuning parameter 0.925 -
LinearSVC(C=0.25, max_iter=100):
              precision    recall  f1-score   support

          -1       0.93      0.83      0.88      1425
           0       0.90      0.97      0.93      1860
           1       0.94      0.95      0.94      2715

    accuracy                           0.93      6000
   macro avg       0.92      0.91      0.92      6000
weighted avg       0.93      0.93      0.92      6000


Classification report for Linear SVC model with tuning parameter 0.9308333333333333 -
LinearSVC(C=0.5, max_iter=100):
              precision    recall  f1-score   support

          -1       0.94      0.85      0.89      1425
           0       0.91      0.96      0.94      1860
           1       0.94      0.95      0.95      2715

    accuracy                           0.93      6000
   macro avg       0.93      0.92      0.92      6000
weighted avg       0.93      0.93      0.93      6000


Classification report for Linear SVC model with tuning parameter 0.933 -
LinearSVC(C=1, max_iter=100):
              precision    recall  f1-score   support

          -1       0.93      0.86      0.90      1425
           0       0.92      0.96      0.94      1860
           1       0.95      0.95      0.95      2715

    accuracy                           0.93      6000
   macro avg       0.93      0.92      0.93      6000
weighted avg       0.93      0.93      0.93      6000
```

Figure 17: LinearSVC Evaluation

5.5 BERT

The next model that is utilized is from the group of Deep Learning named Bert. This model is acquainted in this study validate the data and break down the way of behaving of a Deep Learning model contrasted and a Machine Learning model. The reason to involve this model in the trial and error is to get the comprehension whether a profound learning model like Bert will figure out the context inside the text while perusing the text. The speculation of this task is likewise to check whether a DL algorithm yields a superior precision when contrasted with a ML model. Since a profound learning is equipped for grasping the unique circumstance and characterize the text, it is expected

```
model is training with random_forest algorithm
Classification report for Random forest model with tuning parameter 0.7843333333333333 -
RandomForestClassifier(n_estimators=5):
              precision    recall  f1-score   support

          -1       0.72      0.62      0.67      1425
           0       0.75      0.91      0.82      1860
           1       0.84      0.78      0.81      2715

    accuracy                           0.78      6000
   macro avg       0.77      0.77      0.77      6000
weighted avg       0.79      0.78      0.78      6000


Classification report for Random forest model with tuning parameter 0.8308333333333333 -
RandomForestClassifier(n_estimators=10):
              precision    recall  f1-score   support

          -1       0.81      0.64      0.72      1425
           0       0.81      0.94      0.87      1860
           1       0.86      0.85      0.86      2715

    accuracy                           0.83      6000
   macro avg       0.83      0.81      0.81      6000
weighted avg       0.83      0.83      0.83      6000


Classification report for Random forest model with tuning parameter 0.8401666666666666 -
RandomForestClassifier(n_estimators=15):
              precision    recall  f1-score   support

          -1       0.82      0.67      0.74      1425
           0       0.82      0.94      0.87      1860
           1       0.87      0.86      0.87      2715

    accuracy                           0.84      6000
   macro avg       0.83      0.82      0.82      6000
weighted avg       0.84      0.84      0.84      6000


Classification report for Random forest model with tuning parameter 0.8446666666666667 -
RandomForestClassifier(n_estimators=20):
              precision    recall  f1-score   support

          -1       0.85      0.65      0.74      1425
           0       0.83      0.94      0.88      1860
           1       0.86      0.88      0.87      2715

    accuracy                           0.84      6000
   macro avg       0.85      0.82      0.83      6000
weighted avg       0.85      0.84      0.84      6000
```

Figure 18: Evaluation of Random Forest Model

that the exactness of a Bert algorithm ought to be more prominent than some other ML algorithms, which likewise adds to the hypothesis.

Having finished with the model development, the precision of the model is good compared with others. Contingent upon the opinions of the users from the Tweets are then used to perform brand investigation for each brand. Performing Brand investigation will help in investigating the negative region of all the e-wallet applications where lies the work zone an association ought to be finding certain ways to limit the business bottlenecks.

After performing some tests, by using one hot encoding on the target variable we achieved higher accuracy. For this reason have chosen one hot encoding over label encoding and resulted in better accuracy. Then, have created a custom function to host the pre trained BERT model, and attach to it a 3 neurons output layer, necessary to perform the classification of the 3 different classes of the dataset (the 3 emotions).

Classification Report for BERT:

	precision	recall	f1-score	support
Negative	0.86	0.90	0.88	1468
Neutral	0.93	0.88	0.90	638
Positive	0.91	0.90	0.91	1969
micro avg	0.90	0.90	0.90	4075
macro avg	0.90	0.89	0.90	4075
weighted avg	0.90	0.90	0.90	4075
samples avg	0.90	0.90	0.90	4075

Figure 19: Evaluation of BERT Model

After evaluating the BERT model it's observed that the accuracy for the BERT validation dataset is 90%. This result is good accurate as depicted by other machine learning models.

5.6 RoBERTa

With the use of a dynamic masking technique called RoBERTa, the BERT pre-trained model's next sentence prediction is removed. The RoBERTa model, is an advance on the BERT model masking method.

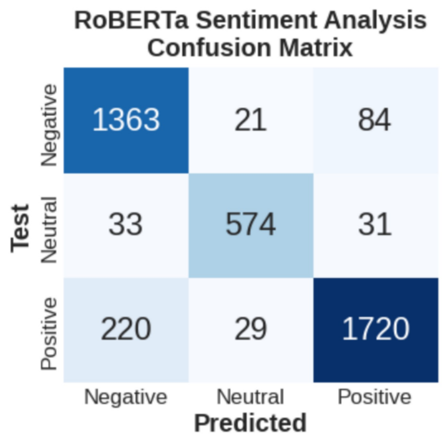

Figure 20: Confusion Matrix RoBERTa

Classification Report for RoBERTa:

	precision	recall	f1-score	support
Negative	0.84	0.93	0.88	1468
Neutral	0.92	0.90	0.91	638
Positive	0.94	0.87	0.90	1969
micro avg	0.90	0.90	0.90	4075
macro avg	0.90	0.90	0.90	4075
weighted avg	0.90	0.90	0.90	4075
samples avg	0.90	0.90	0.90	4075

Figure 21: Evaluation of RoBERTa Model

After evaluating the RoBERTa model it's also observed that the accuracy for the validation dataset is 90%. This result is good accurate.

5.7 Summary

Modeling phase is the prominent phase in CRISP-DM methodology which is done for making predictions on the previously cleaned dataset. To Summarize this work, there are 7 models used, out of which 4 are machine learning models and 3 is a deep learning model. All the models are predicting the sentiments classified into three classes namely NEGATIVE, POSITVE and NEUTRAL. Based on the accuracy of the prediction, the accurate model will be used to make sentiment analysis on the twitter data for e-wallet applications.

5.8 Manual Analysis Random Forest

After creating dataframe with Random Foest test Tweet, test Sentiments and Predicted Sentiments data, we have done manual analysis and analysed that total 1278 tweets are not correctly predicted.Few sample tweets shown in below figure. It shows that when two tweets are connected then prediction result considering only first half and ignoring second half.

	A	B	C
1	X_Test(Tweets)	Y_Test	Pred
4	reinforces how much i hate	-1	0
5	but today when i was debited with same amount rs39970 then i called and now bank telling this below story to me how is it possible to credit and debit my account without asking or informing	1	0
6	till yesterday time provided by your team for 22 now its extended again and again till 30th nov and no proper details when ill get my	-1	-1
7	very bad experience i use the phonepe app but money not transfering why the problem	-1	-1
8	poor service i have successfully made the payment to merchant but they didnt recieved and my money has debited from my account	1	-1
12	20days over but paid bill not shown no amount refunds no complaints solution what can i do will one bill have to be deposited again and again	1	-1
14	no one is responding on the given number please expedite	1	-1
19	your dumb sms login system doesnt work	-1	0
21	plz do something i completed a payment through phone pe for train ticket to irctc app but after payment this app misbehaving and after that i didnt get any confirmation message for ticket my money didnt return after that plz help	1	-1

Figure 22: Manual evaluation of Random Forest Model

5.9 Manual Analysis BERT

After creating dataframe with BERT test Tweet, test Sentiments and Predicted Sentiments data, we have done manual analysis and analysed that total 445 tweets are not correctly predicted.Few sample tweets shown in below figure. In this also, it shows that when two tweets are connected then prediction result considering only first half and ignoring second half. most of the tweets also were part of random forest manual analysis.

	A	B	C
1	X_Test	Y_Test	Pred
16	so youre saying one of the ultimate goals of your fellow marxist friends and idols is not to confiscate peoples firearms you better check back with your fake latino pendejo beto	positive	negative
20	till now the money is not been updated in pay later or refunded now telling me to wait till 13 th as there is some technical issues the money which paid on 2nd thru is stuck somewhere till 13 th will give interest for my money stuck with them	negative	positive
48	deducted money but not recharged my number	negative	positive
53	how can transactions can happen without any authorisation i havent used my debit card at marvacose barksdale but money have deducted automatically i even tried paytm 24 x 7 help but there was no correct answer from them	positive	neutral
54	sir i am biswajit mondal belong from kushiyara city mihijam district jamtara jharkhand my area everyones paytm account is blocked in my city pin 815354 you are requested to solve this problem soon as possible and active upi translation for future transactions	negative	neutral

Figure 23: Manual evaluation of BERT Model

5.10 Discussion

As stated in the Literature Review, there were few works carried out using twitter data to analyse the sentiments of the users. Moreover, a few of the authors have performed some work to indicate the growth of e-wallet applications in India after Demonetization was brought into immediate effect and post covid era. This study has made a brief analysis on the post covid dataset extracted from twitter directly to judge the opinions of the users after the growth and when the covid restrictions were taken down. Using some of the essential ML models, this project has determined the way to work on the live tweets or Raw data for sentiment analysis and by using Vader sentiment which gives 96% of accurate text classification results. The results obtained from the best accurate model can be used for future dataset which could be supervised or unsupervised.

Furthermore, the research also focuses on getting a business solution for all the individual payment apps by looking at all the negative comments posted by the users. This is achieved by plotting word clouds of the negative tweets. As the revenue of each brand is directly dependent on the number of users on the app, to increase the footfalls of users on their apps, the effort making domain is determined by this study. With the help of graphs related to location, the marketing team could put more efforts in proper exposure of their product and features in the highly negative locations.

The Null hypothesis of this study stated that Deep Leaning models will perform better as compared to machine learning models as it is possible to understand the context of the tweets for few of the terms which have multiple meaning. After the experimentation and analysis, it is learnt that the Stated Null Hypothesis does hold true for this dataset. As per the obtained results, accuracy is achieved by LinearSVC is 91% for count vectorizer with C=1 and max iteration = 100 and with BERT achieved accuracy is 90%.

6 Conclusion and Future Work

In this project, data analysis is performed using python for multiple e-wallet brands that are used in India. By analysing the user experience, it is determined that the users are unhappy with the performance of applications due to various issues like transaction failure, non-responsive support team during the need of an hour. It is also learnt that applications like Paytm whose interface allows to purchase movie tickets are being criticized for hosting the movies that have negative impact on user sentiments, and hence there are many boycott hashtags in trend for those apps.

By performing model development and evaluation, it can be concluded that Linear SVC can be utilized for sentiment analysis for raw and live twitter data classification for sentiment analysis, as the accuracy for Linear SVC is achieved to be 91% at the C value of 1 for maximum 100 iterations. Linear SVC has also performed well for TF-IDF and n-gram Count vectorizer with an accuracy rate of 91% and 89% respectively. The second-best performing model is Logistic Regression with count vectorization method having an accuracy of 91%. The poorest performance is achieved by KNN algorithm having a maximum accuracy of 56%.

6.1 Limitations

- Since the study is conducted on a raw dataset, there isn't any solid base to which the model can be compared to get the accuracy since the ground truth for the live dataset will always be missing.

- As the data is raw, cleaning the raw dataset becomes a tedious task and a 100% cleaned dataset cannot be achieved.

- Since the data is extracted using snscraper module where filtering the English records goes out of option, data gets extracted into multiple languages, and Vader Sentiment analyser can only understand English language.

- There could be instances where the users' tweets are written using English alphabets but the actual language of the word is other than English, so the conversion of such data into English language become tedious.

6.2 Future work

To make the model classification more reliable, in future the study can be extended using labelled dataset which is marked professionally by someone. To get a better understanding of the Deep Learning, other models can be tested other than BERT model and Roberta as utilized in the current study. To extend further on the analysis for e-wallets, two algorithms could be clubbed together to gain an even accurate model for classification. For instance, the currently best performing model turned out to be BERT, output from this model can be treated as an input to another Machine learning or Deep Learning model. The above suggested combinational model could result in better accuracy provided we have the ground truth to compare the classification of the models.

References

Becker, K. and Ghedini, C. G. (2005). A documentation infrastructure for the management of data mining projects, *Inf. Softw. Technol.* **47**: 95–111.

Boumhidi, A., Benlahbib, A. and Nfaoui, E. H. (2022). Cross-platform reputation generation system based on aspect-based sentiment analysis, *IEEE Access* **10**: 2515–2531.

Cadotte, E. R., Woodruff, R. B. and Jenkins, R. L. (1987). Expectations and norms in models of consumer satisfaction, *Journal of Marketing Research* **24**: 305 – 314.

Chang, V., Doan, L. M. T., Di Stefano, A., Sun, Z. and Fortino, G. (2022). Digital payment fraud detection methods in digital ages and industry 4.0, *Computers Electrical Engineering* **100**: 107734.

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T. P., Shearer, C. and Wirth, R. (2000). Crisp-dm 1.0: Step-by-step data mining guide.

Cs, R. (2017). Digital payments system and rural india: A review of transaction to cashless economy, *International Journal of Commerce and Management Research* **3**: 169–173.

Danmon Projects (2022). *Hot girls coloring book*, Independently Published.

Duhamel, A., Nuttens, M. C., Devos, P., Picavet, M. and Beuscart, R. (2003). A preprocessing method for improving data mining techniques. application to a large medical diabetes database, *Studies in health technology and informatics* **95**: 269–74.

Gautam, G. and Yadav, D. (2014). Sentiment analysis of twitter data using machine learning approaches and semantic analysis, *2014 Seventh International Conference on Contemporary Computing (IC3)* pp. 437–442.

Giese, J. and Cote, J. (2000). Defining consumer satisfaction, *Academy of Marketing Science Review* **4**: 1–24.

Gordijn, J., Petit, M. and Wieringa, R. (2006). Understanding business strategies of networked value constellations using goal- and value modeling, *14th IEEE International Requirements Engineering Conference (RE'06)* pp. 129–138.

Hayashi, F. (2012). Mobile payments: What's in it for consumers?, *Economic Review* pp. 35–66.

Hidayatullah, A. F., Cahyaningtyas, S. and Hakim, A. M. (2021). Sentiment analysis on twitter using neural network: Indonesian presidential election 2019 dataset, *IOP Conference Series: Materials Science and Engineering* **1077**.

Johnson, V. L., Kiser, A. I. T., Washington, R. and Torres, R. (2018). Limitations to the rapid adoption of m-payment services: Understanding the impact of privacy risk on m-payment services, *Comput. Hum. Behav.* **79**: 111–122.

Joshi, M. (2017). *Digital Payment System: A Feat Forward of India*.

Kristiyanti, D., Putri, D., Indrayuni, E., Nurhadi, A. and Hairul Umam, A. (2020). E-wallet sentiment analysis using naïve bayes and support vector machine algorithm, *Journal of Physics: Conference Series* **1641**: 012079.

Kumar, A. (2020). KNN algorithm: When? why? how?, `https://towardsdatascience.com/knn-algorithm-what-when-why-how-41405c16c36f`. Accessed: 2022-12-13.

Liao, S., Wang, J., Yu, R., Sato, K. and Cheng, Z. (2017). Cnn for situations understanding based on sentiment analysis of twitter data, *Procedia Computer Science* **111**: 376–381. The 8th International Conference on Advances in Information Technology.
URL: *https://www.sciencedirect.com/science/article/pii/S1877050917312103*

Liébana-Cabanillas, F. J., Sánchez-Fernández, J. and Muñoz-Leiva, F. (2014). Antecedents of the adoption of the new mobile payment systems: The moderating effect of age, *Comput. Hum. Behav.* **35**: 464–478.

Maindola, P., Singhal, N. and Dubey, A. D. (2018). Sentiment analysis of digital wallets and upi systems in india post demonetization using ibm watson, *2018 International Conference on Computer Communication and Informatics (ICCCI)* pp. 1–6.

Mayasari, A. E. and Hartanto, A. D. (2019). User satisfaction levels sentiment analysis toward goods delivery service on twitter using support vector machine algorithm (SVM), *2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, IEEE.

Mekovec, R. and Hutinski, Z. (2012). The role of perceived privacy and perceived security in online market, *2012 Proceedings of the 35th International Convention MIPRO* pp. 1549–1554.

Moss, L. T. and Atre, S. (2003). Business intelligence roadmap: The complete project lifecycle for decision-support applications.

Nadali, A., Kakhky, E. N. and Nosratabadi, H. E. (2011). Evaluating the success level of data mining projects based on crisp-dm methodology by a fuzzy expert system, *2011 3rd International Conference on Electronics Computer Technology* **6**: 161–165.

Nandal, N., Nandal, N. and Malik, R. (2020). Is loyalty program as a marketing tool effective?

Narayanan, R. (2020). Cashless transactions and its impact of salaried class, *Test Engineering and Management* **83**: 2834 – 2842.

Ondrus, J. and Pigneur, Y. (2006). Towards a holistic analysis of mobile payments: A multiple perspectives approach, *Electronic Commerce Research and Applications* **5**(3): 246–257. Mobile technology and services.
URL: *https://www.sciencedirect.com/science/article/pii/S1567422306000159*

Orche, A. E. and Bahaj, M. (2019). Approach to use ontology based on electronic payment system and machine learning to prevent fraud, *NISS19*.

Padiya, J. and Bantwa, A. (2018). *Adoption of E-wallets: A Post Demonetisation Study in Ahmedabad City*.

Phan, H. T., Nguyen, N. T., Tran, V. C. and Hwang, D. (2021). An approach for a decision-making support system based on measuring the user satisfaction level on twitter, *Inf. Sci. (Ny)* **561**: 243–273.

Rampton, J. (2016). BWorld Robot Control Software, `https://techcrunch.com/2016/06/17/the-evolution-of-the-mobile-payment/?guccounter=2`. [Online; accessed 17-November-2022].

Sandra, L., Marcel, Gunarso, G., Fredicia and Riruma, O. W. (2022). Are university students independent: Twitter sentiment analysis of independent learning in independent campus using roberta base indolem sentiment classifier model, *2021 International Seminar on Machine Learning, Optimization, and Data Science (ISMODE)*, pp. 249–253.

Saragih, P. S., Witarsyah, D., Hamami, F. and Machado, J. M. (2021). Sentiment analysis of social media twitter with case of large scale social restriction in jakarta using support vector machine algorithm, *2021 International Conference Advancement in Data Science, E-learning and Information Systems (ICADEIS)*, pp. 1–6.

Sarwan, K., Thore, S., Satav, N., Kamble, P. and Mandal, D. (2021). A secure e-wallet system using block chain, *International Journal of Scientific Research in Computer Science, Engineering and Information Technology* pp. 454–466.

Singh, S. and Rana, R. (2017). Study of consumer perception of digital payment mode, *The Journal of Internet Banking and Commerce* **22**: 1–14.

Sujiwo, B., Wibowo, A. and Saputro, D. R. S. (2021). Sentiment analysis of indonesian government policies in handling covid 19 through twitter data, *2021 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, pp. 453–457.

Sulistyono, A., Mulyani, S., Yossy, E. H. and Khalida, R. (2021). Sentiment analysis on social media (twitter) about vaccine-19 using support vector machine algorithm, *2021 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, pp. 62–67.

Sweidan, A. H., El-Bendary, N. and Al-Feel, H. (2021). Sentence-level aspect-based sentiment analysis for classifying adverse drug reactions (adrs) using hybrid ontology-xlnet transfer learning, *IEEE Access* **9**: 90828–90846.

Tan, A.-H., Ridge, K., Labs, D. and Terrace, H. (2000). Text mining: The state of the art and the challenges.

Tan, K. L., Lee, C. P., Anbananthen, K. S. M. and Lim, K. M. (2022). Roberta-lstm: A hybrid model for sentiment analysis with transformer and recurrent neural network, *IEEE Access* **10**: 21517–21525.

Tripathi, S. and Dave, N. (2022). Cashless transactions through e-commerce platforms in post-covid-19, *International Journal of Management, Public Policy and Research* .

Velardi, P., Fabriani, P. and Missikoff, M. (2001). Using text processing techniques to automatically enrich a domain ontology, *Formal Ontology in Information Systems*.

Wang, S. W., Ngamsiriudom, W. and Hsieh, C.-H. (2015). Trust disposition, trust antecedents, trust, and behavioral intention, *The Service Industries Journal* **35**(10): 555–572.
URL: *https://doi.org/10.1080/02642069.2015.1047827*

Wang, Y., and (2019). Factors that influence consumer satisfaction with mobile payment : The china mobile payment market, *Informs Journal on Computing* **15**: 82–83.

Westbrook, R. A. and Reilly, M. D. (1983). Value-percept disparity: an alternative to the disconfirmation of expectations theory of consumer satisfaction, *ACR North American Advances* .

Wibawa, D. W., Nasrun, M. and Setianingsih, C. (2018). Sentiment analysis on user satisfaction level of cellular data service using the k-nearest neighbor (K-NN) algorithm, *2018 International Conference on Control, Electronics, Renewable Energy and Communications (ICCEREC)*, IEEE.

Zahoor, S. and Rohilla, R. (2020). Twitter sentiment analysis using machine learning algorithms: A case study, *2020 International Conference on Advances in Computing, Communication Materials (ICACCM)*, pp. 194–199.

Zhang, S., Zhang, C. and Yang, Q. (2003). Data preparation for data mining, *Applied Artificial Intelligence* **17**: 375 – 381.