# A Comparison of Sentiment Analysis Algorithms Using Twitter Feedback on a Local Bank: Developing Marketing Strategies for Philippine Banks

## Alexander Manuel Dean T. Ferrazzini [1], Leo Neil P. Flores [2], Justin Jay L. Joaquin [3], Cherry B. Lisondra [4]

[1, 2, 3] Student, [4] Faculty,
College of Computer and Information Science, Mapua Malayan Colleges Mindanao, Philippines.

## Abstract

A breach in BDO's online security has escalated into clients losing their money, and BDO is faced with the need to provide assurance that strict cyber security protocols are being followed and that the risk of outsider hacking is negligible. Therefore, sentiment analysis is an appealing tool to facilitate the collection of feedback to improve banking processes. To perform the analysis, several Python Libraries and the Naïve-Bayes Algorithm were used. For data mining, Apify was used via its Twitter Scraper feature to gather data for the dataset used to train the algorithm. The individual accuracy of the program created was then compared to existing studies that used the Semantic Approach and J48 Algorithm to determine which algorithm is the most accurate and superior. Compared to J48 and Semantic Approach, J48 is the most accurate with Semantic Approach being the least accurate. The accuracy of the Naive Bayes algorithm is in the middle at 86.82%. While not as accurate as J48, this is compensated by its faster processing time and faster learning capabilities. From the results of the sentiment analysis, BDO must improve in various facets of their service such as cybersecurity, customer service, and privacy. The website and app must also be optimized for better online banking. For future research, it is recommended that the dataset is processed in other algorithms for a more accurate comparison. Front-end integration should also be considered for non-technology adept users and the code must also be optimized for efficient and more accurate data processing.

**Keywords: Bank, Sentiment Analysis, Marketing, Feedback**

## 1. Introduction

Customer reviews and feedback are the lifelines of any service. Past customer sentiments often back decisions regarding where to avail of certain services–new customers make intelligent inferences about the quality of a business based on the reviews they read online (Thorson & Duffy, 2011; as cited in Terkan, 2014). Due to the vast number of online discussions reflecting consumers' thoughts, feelings, and opinions about products and brands, social media platforms have become popular vehicles for studying consumer sentiment on a wide scale and in a natural setting. Organizations are increasingly interested in analyzing the public's view of their services using social media - including blogs, online forums, comments, tweets, and product evaluations (Dwivedi et al., 2020).

Banks and other similar services are no exception to the growing population of online consumer reviews. As banks serve as one of the key players in the financial growth of an individual, it also encounters feedback on the relative smoothness and efficiency of their functions. The rapid advancement of technology has become a threat to existing banks, as new generation private sector banks are emerging and can easily create a niche with innovation (Kaura, 2013). Therefore, customer satisfaction serves a vital role in the relative survival of banks - customer retention requires that services are delivered to their satisfaction (Kaura, 2013).

However, it is not that banks can reach consumers' expectations at all times. In the case of Banco de Oro, commonly known as BDO Unibank, Inc, a recent breach in their online security has escalated into clients losing their money. Several victims have stated that they have lost substantial funds even without providing third-party access to their bank accounts and not clicking any phishing links. Cepeda (2021) recounts that numerous BDO clients were surprised to receive emails and text messages from the bank notifying them that a bank transfer involving thousands of pesos was successfully processed.

The Bangko Sentral ng Pilipinas (BSP) states that the recent hacking incident may be a case of an inside job. However, they are leaving it to the bank to investigate while they remain in close coordination (Fernandez, 2021). This circumstance led to an online uproar, damaging the bank's reputation and the trust and satisfaction of its current customers. The problem now remains: how can BDO regain the trust of their consumers and prevent existing clients from switching banks.

BDO is now faced with the need to provide more rigid marketing and assurance to clients that strict cyber security protocols are being followed. The risk of outsider hacking is almost zero - the backbone of this is customer feedback. Following the incident, the BSP has continued monitoring the surge in complaints across social media platforms (Luna, 2021). BDO also collects feedback via forms posted on its official website, but this mode of manual monitoring and feedback collection could be time-consuming and labor-intensive. This circumstance will inevitably cost the additional bank capital, which is potentially an expense they would rather minimize considering the losses from the hacking incident.

(Hasson et al., 2019). This notion opens an avenue to consider automatic and automated methods of collecting and analyzing feedback. Social media platforms tend to be saturated, making it difficult to navigate for BDO-specific reviews. As such, sentiment analysis becomes an appealing tool to simplify and facilitate the overall feedbacking process. Simple sentiment annotation tasks, in which annotators

must assess whether a sentence is positive, negative, or neutral, are frequently used to analyze sentiment in textual content (Medhat et al., 2014).

Sentiment analysis has shown its potential in terms of feedback evaluation. The system was able to identify features regarding a service that were previously not taken into consideration in an existing survey questionnaire (Kumar & Jain, 2015). However, sentiment analysis is an umbrella term encompassing various algorithms and methods. It is essential to explore a method that is accurate yet easy to apply and possibly even cost-free, mainly when banks such as BDO who have losses aim to retain and attract customers while improving their services. Two contenders for this are Machine Learning and Lexicon-based Approaches.

It is said that one approach is more accurate than the other. Results may vary depending on the training of the algorithms. Therefore, both approaches will be compared to determine which one will provide a more accurate picture of consumer sentiments to serve as the backbone of new marketing strategies.

The surge of online complaints following the unfortunate hacking incident from BDO has emphasized the relative importance of feedback monitoring and evaluation. To regain client trust and recover from its losses, the bank would have to resort to more aggressive marketing and client assurance that the same incident will no longer occur in the foreseeable future. The backbone of this is consumer feedback, but given the manual methods the bank and even the BSP use, feedback collection will become time-consuming. Therefore, sentiment analysis provides an automated, automatic feedback collection and evaluation solution, saving the bank time and effort and its services and reputation.

## 1.1. Statement of the Problem
In regaining customer trust and improving customer retention and loyalty, obtaining an accurate picture of consumer sentiment is prioritized. For banks such as BDO, which operate in a saturated market and have experienced losses due to the recent hacking incident, this portrayal of sentiments is pivotal to its marketing strategies.

Therefore, sentiment analysis is key to summarizing the surge of online complaints and general feedback of the bank. However, which sentiment analysis method is most accurate and relatively easy to apply remains a concern. As such, the research aims to answer the following problems:
**(1)** Which among Machine Learning and Semantic Analysis provides a more accurate picture of consumer sentiments?
**(2)** How can the chosen method impact the feedback collection and evaluation of businesses?
**(3)** How can the analysis results - whether it be positive, negative, or neutral - shape the marketing strategies of businesses.

## 1.2. Objectives of the Study
The main objective of this paper was to develop and optimize a program that collects customer feedback from Twitter modeled after a Machine Learning algorithm, specifically the Naïve Bayes classifier, applied in Python. Additionally, this inquiry also aimed to compare the results of the developed program to existing studies related to the Semantic Approach and recommend an algorithm for potential use by banks. Specifically,

**(1)** To utilize data mining techniques to collect BDO-related customer sentiments in Twitter effectively,

**(2)** To compare the individual accuracy of the developed program (Machine Learning) to the results of existing Semantic Approach studies and;

**(3)** To determine which among the algorithms is most accurate and superior.

## 1.3. Scope and Limitation

The scope of the study was hinged on the creation of a program to collect and analyze BDO-related comments and feedback on Twitter. Twitter was chosen over other social media platforms because of the absence of a strict Data Privacy Policy on their platform. This feature required the consent of individuals for their respective feedback to be used for the dataset. Well known and renowned social media platforms like Facebook and Instagram implement this policy, which hinders the process of creating the dataset to be used in the analysis.

The proponents used the Naïve Bayes classifier algorithm & J48 for Machine Learning, written in Python: a high-level, general-purpose programming language. After the program using the Machine Learning approach was created, the results were compared to results from existing studies related to the Semantic Approach.

Moreover, the paper only covered Twitter posts surrounding the bank of choice, BDO. The volume of tweets that encompass the bank affected the data gathering and the significance of the data set to be used for the analysis. Furthermore, creating the dataset for analysis may be time and labor-intensive.

## 1.4. Significance of the Study

The study is set to not only accomplish its objectives but also to gain social relevance, especially to the following entities:

**(1) Banks:** This study may determine if the Machine Learning and Sentiment Analysis methods to be tested are accurate and significant in perceiving consumer sentiment. As such, the data analyzed and collected from the platform by the program may be used by BDO and other Philippine banks alike to determine the right marketing strategy to implement to address concerns such as cybersecurity, customer retention, and customer loyalty. This inquiry can also help banks shift from the manual mode of feedback collection into an automated and automatic approach, improving costs moving forward.

**(2) Future Researchers:** The literature already exists comparing the latent accuracy of semantic analysis methods and machine learning techniques. However, literature regarding using any of these two methods to improve marketing strategies is little to none. As such, this study may be used as a reference for future researchers who plan to follow the same rule of thought that is expressed by this paper.

**(3) Fellow Students:** The general notion and ideas expressed in this paper can help Computer Science students who also wish to venture into the field of sentiment analysis. Given that these two methods can be quite cumbersome to implement, referencing this study may aid them in a flawless implementation of sentiment analysis. Furthermore, they may also become more familiar with other use-cases of sentiment analysis, especially in the banking industry.

**(4) Community:** Through the results of this study, the general community may become more knowledgeable about sentiment analysis and how it also has real-world uses and implications. By shedding light on how simple it can be to provide a general sentiment on a particular topic or field, the community may also apply this knowledge in several disciplines of their interest.

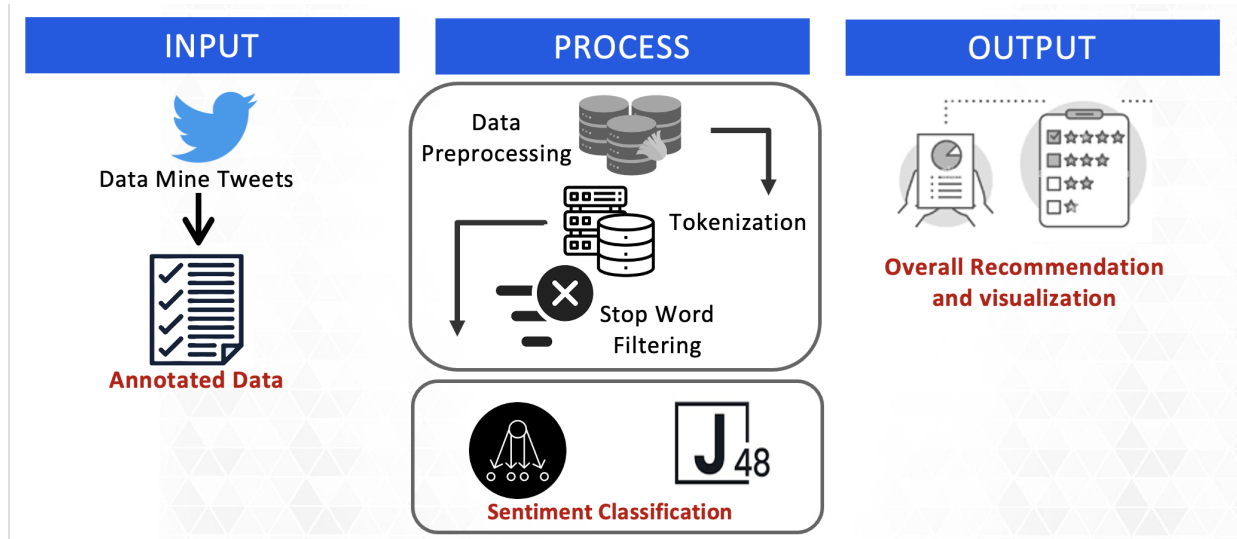## 1.5. Conceptual Framework



**Figure 1: Conceptual Framework**

The conceptual framework outlined above provides an overview of the process that this study had undergone to fulfill its primary objectives. The procedure begins with Data Collection, followed by Data Preprocessing, Tokenization, Stop Word Filtering, Sentiment Classification, and finally, the generation of the Sentiment Score. The processes are further explained in the succeeding paragraphs.

Beginning with Data Collection, this involves the initial collection of data using data mining techniques. The data then enters into the preprocessing stage, where the data collected can be annotated and classified between positive and negative to assist in training the program. Afterward, the text enters the Tokenization phase, breaking the raw text into small chunks. The sentences collected are broken into words that are coined as tokens. These assist in understanding the context and in developing the model.

Once the Tokenization phase is concluded, Stop Word Filtering begins, where stop words such as *if, but, we, he, and she* are removed and filtered. These words can be safely removed without changing the semantics of a text. In doing so, it improves the performance of a model. After the tokenization and stop word filtering processes, Sentiment Classification takes place. The program labels the texts as positive, negative, or neutral, based on data collected and the model developed. The program defines the orientation of the sentiment, known as Sentiment Polarity. Finally, a Sentiment Score is assigned, ranging from zero to ten–from most negative to most positive.

## 2.   Literature Review

This chapter aims to provide an in-depth discussion on the investigated phenomenon. In this chapter, a review of related articles, publications, and researches about marketing in the digital age, competitive

marketing of banks, the hacking case of BDO, customer satisfaction, value of feedbacking, and the usage of sentiment analysis in feedbacking.

## 2.1. Marketing in the Digital Ages

To recognize the increasing influence of globalization and the interdependence of countries, nations, and communities would recognize the increased use of technology in many mediums. This era of globalization has introduced a plethora of technological advances, leading to the spread of various business ideas, knowledge, and information amongst individuals located around the globe (Terkan, 2014). With the increasing pervasiveness of technology in people's daily lives, it is slowly becoming the center of worldwide transactions and communication.

This technological age has marked a shift in the way businesses operate. Customers have now taken to the online means, disregarding print media and traditional advertising. They are now in control of what they wish to see, hear, and buy (Thorson & Duffy, 2011; as cited in Terkan, 2014). As consumers continue to penetrate the digital space, businesses are now urged to transform the way they interact with them and assert their business in methods that were previously unthinkable in the traditional times. This innovation brings forward newfound importance for marketing. The general marketing process involves building profitable, value-laden exchange relationships with customers (Sheldrake, 2011). This notion would mean engaging with a target market or group of consumers strategically to sell better the business's products and services (Gleeson, 2019). Under the traditional setting, marketing would generally involve producing advertising materials distributed to passersby, booking a timeslot for a radio commercial, or pasting posters on sidewalks and public areas. However, the digital age has marked the onslaught of "digital saturation" - a state where the digital space has become filled with roughly two billion websites, with an internet-viewing audience that does not have the time to view more digital content (Mandr Group, 2020).

The question now remains: how do businesses stay relevant in this era of digital saturation? With the vast online choices and how consumers have become reliant on online marketing, it has become a requirement now for businesses to keep in touch with their consumers, whether through surveys, quick ratings, or asking for short comments. However, this process is complex and may require some feedbacking mechanism between the consumer and the business.

## 2.2. Technology and Competitive Marketing by Banks

Bank marketing is defined as that part of management activity that seems to direct the flow of banking services profitably to selected customers (Meidan, 1984). Traditionally, banks did not struggle to attain profitability, as readily available customers would generate enough profits (Masocha et al., 2011). However, the radical shift in technology has transformed the customer experience, banks' back-end operations, and the marketing function of banks (Loftin, 2020). The disconnect between the bank management's strategic focus and the marketing activities employed could potentially damage the bank's overall performance and service delivery and remain a point that must be critically addressed (Loftin, 2020).

Other than the technological revolution, intense competition from other banks, financial corporations, and insurance companies has caused banks to prioritize marketing (Meidan, 1984). This is one of the

significant changes outlined in the so-called "retail banking revolution," which was initially seen in 1958 up to the last fifteen years. This revolution has brought about a more precise outline of a bank's typical marketing approach, which Meidan (1984) outlines as follows: (i) research to determine customers' financial requirements, (ii) design new services or innovate old ones according to the findings, (iii) market services to the customer for whom they were researched and designed at a profit, and (iv) satisfy the customers' financial needs. Like any other business, banks remain concerned with relevance and service delivery.

Of course, banks should not stop at competitive marketing alone. While it is essential to seek new customers actively, this could potentially be more expensive than retaining the current customer base (Kaura, 2013). It is often stressed that customer retention is addressed by relationship marketing - a marketing strategy for customer-building and retention (Rootman et al., 2012). Superior relationship marketing translates to maintaining profitable banking clients alongside client trust and customer satisfaction (Info-Electronics Systems 2004; Swartz & Iacobucci 2000: 96, 328; as cited in Rootman et al., 2012).

While the reality is that banks remain non-negotiable in the financial growth and maintenance of an individual, the onset of technological advancements has caused banks to innovate. The banking industry has capitalized on the growth of technology and now has a sophisticated infrastructure designed to cope with constant changes (Aldaihani & Ali, 2019). Technology has also driven relatively newer banks to appeal to the general consumer. The rising figure of players in the banking sector has somewhat increased customer needs and shifted their desires (Aldaihani & Ali, 2019), leading to a preference for one bank over another. This feature is a potential driver of customer satisfaction, and the survival of banks remains on the line (Kaura, 2013).

### 2.3. Impaired Customer Trust: The Case of BDO

Banco de Oro Unibank, Inc., abbreviated as BDO Unibank, Inc., is a full-service universal bank in the Philippines, providing an array of financial products and services, including but not limited to the following: Lending, Deposit-taking, Foreign Exchange, Brokering, Trust and Investments, Credit Cards, Retail Cash Cards, and Corporate Cash Management and Remittances (BDO, n.d.). It is one of the most popular banks in the Philippines. It is arguably the most awarded bank globally, garnering awards from various institutional bodies in Asia and Europe.

Despite being an award-winning bank that also boasts its excellence in cybersecurity, the onslaught of the recent online bank account hacking has caused BDO to lose the trust of its most important constituents. Nearly 700 clients reported to have had their funds - usually in thousands of pesos - maliciously withdrawn without their consent and deposited to a UnionBank account of a certain "Mark Nagoyo" (Cepeda, 2021; Dumlao-Abadilla, 2021). Clients could only be surprised that the BDO text service messages notified them of the withdrawal of their funds and the subsequent successful transfer of said funds (Cepeda, 2021). Despite not having clicked any malicious links, clients remained victims of the unfortunate incident.

The case had further escalated into the hands of The Bangko Sentral ng Pilipinas (BSP), the overall regulatory body for banks and financial institutions alike. BSP Governor Benjamin Diokno asserts that

the incident may have been a case of an inside job, but such has yet to be confirmed. They continue to coordinate with the two banks involved in resolving the incident (Fernandez, 2021). Admittedly, this has caused even more reluctance among many Filipinos and has slashed their trust in online bank payments and money transfers (Gamboa, 2021).

Interestingly, this is not the first case of a banking hack encountered by BDO. Even dating back to 2018, the universal bank had experienced an extraordinary rise in unauthorized transactions (Dumlao-Abadilla, 2018). Dumlao-Abadilla (2018) further recounts the incident, stating that this was under the assumption that crime syndicates were placing substantial efforts into obtaining individual card and account details, possibly through illegal means such as the black market or social engineering means. BDO had assured its clients that they were investigating the matter thoroughly and devoted significant human resources to fraud investigation, but this did not stop the surge of complaints on social media.

In both cases, BDO had been subjected to an online uproar fueled by clients who expressed their frustrations and disappointments towards the bank (Chipongian, 2021; Dumlao-Abadilla, 2018). To remedy the more recent case, BDO promised to reimburse the almost 700 clients affected by the fraud attack to keep a good customer relationship despite not being legally liable (Chipongian, 2021). In a statement issued by BDO on December 21, non-liability in the case of fraudulent or unauthorized utilization of the accounts due to theft is average compliance among the banking industry. The bank had just made an exception and shouldered the losses from the incident (Luna, 2021).

Having lost its clients' trust and reputation, BDO took measures to recover, such as the absorption mentioned above of losses and reimbursement of clients affected. Both the bank and the BSP remain in coordination and actively monitor numerous social media channels for complaints and feedback (Luna, 2021). Similarly, BDO already has a feedback-collection system via a feedback form posted on their website. While these are steps to regaining the trust and loyalty of their clients, they need to emphasize that excellent firm-client relationships remain important as these influence satisfactions, support, and the retention of banking clients (Rootman et al., 2012).

## 2.4. Customer Satisfaction, Social Media, and Feedbacking

The marketing process is heavily centered on the consumer. Intricate marketing disciplines have been crafted to provide a more effective analysis of the behavior and preferences of the consumer, such as Consumer Market Research (Lumen Learning, n.d.). The success of marketing strategies would lead to business success and considering that marketing strategies are geared towards wooing the consumer, customer satisfaction becomes an important metric.

Milner & Furnham (2017) briefly define customer satisfaction as an assessment of how a business's products or services measure the customer's expectations. They further expound that it evaluates the customer's experience and is more likely to predict customer retention, loyalty, and product repurchase. The most common metrics for customer satisfaction are gathered via a traditional survey, but this generally involves added costs and requires active customer participation, which are areas of potential risk to a business (Hasson et al., 2019). However, with the emergence of social media platforms such as Facebook and Twitter, businesses can view customer satisfaction through their online written feedback organically.

The creation of social media platforms has allowed internet users to have a medium for expressing and communicating their thoughts with thousands of other users alike (Bhatia et al., 2013). Part of their expression involves sharing opinions on products and services, which influences the buying decision of other consumers in the market for similar products. This streamlined consumer-to-consumer communication facilitated by social media platforms can impact a company's reputation and sales (Kietzmann et al., 2011; as cited in Bhatia et al., 2013), providing helpful feedback to companies and businesses alike.

As hinted in the previous paragraphs, customer feedback is defined as customer communication regarding a product or service (Erickson & Eckrich, 2001; as cited in Nasr et al., 2014). The importance of feedback cannot be stressed enough - many successful businesses and companies rely on customer feedback to improve their current products and services and determine what is most important to their customers (Chron, 2021). Consumers' major problems become well highlighted, and businesses can take appropriate corrective measures to remedy the common issues faced (Bhatia et al., 2013).

With social media platforms becoming well-received and incorporated into consumers' daily lives, there is an opportunity to monitor feedback communicated within such platforms actively. In contrast to the more traditional print and radio advertising, where it was almost impossible to determine consumer feedback unless survey questionnaires were manually printed and handed out to passersby, consumer reviews on the internet are almost instantaneous and can be viewed by the public. This notion presents an opportunity to perform feedback evaluation, but it can be tedious and time-consuming without the proper means and suitable plan. However, as the deployment of survey questionnaires requires more labor capital and cost, it would not hurt to venture into this option.

## 2.5. Sentiment Analysis and Feedback Evaluation

Feedback does not only end with its collection, but its evaluation is also necessary for improving a business's operations and marketing strategies. A questionnaire-based system is often used to evaluate performance (Kumar & Jain, 2015). However, this is costly and imposes more outstanding labor capital on a business. Therefore, an automated or automatic feedback evaluation system could be a reasonable consideration.

Beigi et al. (2016) define Sentiment Analysis as a category of computational and natural language processing techniques. These are heavily based on language processing used to identify, extract, or characterize personal information conveyed in a text, such as opinions. The fundamental goal of sentiment analysis is to categorize a writer's attitude toward diverse issues into three categories: positive, negative, and neutral. Sentiment analysis has a wide range of applications in various fields, including corporate intelligence, politics, sociology, and more.

Sentiment Analysis is an umbrella program that encompasses multiple algorithms. These include Machine-learning, Lexicon-based, and hybrid approaches, the most common. Statistical, knowledge-based, and hybrid techniques are also under Sentiment Analysis but are less commonly used. According to Hassan et al. (2013), Sentiment Analysis is considered to be a steady practice of extracting information from available data on social networks for election prediction, educational purposes, business, communication, and marketing purposes. Due to the difficulty in computationally evaluating

views and attitudes, Sentiment Analysis has been widely connoted in multiple studies concerning view evaluation.

A typical implementation of Sentiment Analysis is through feedback evaluation. In reference to Kumar & Jain (2015), who used sentiment analysis to facilitate feedback evaluation, they concluded that their system could identify various features that were previously not stated in their control questionnaire. Similarly, a study by Altrabsheh et al. (2014) also investigated the use-case of different machine learning techniques in analyzing real-time student feedback. This study found that the different models they tested, specifically the SVM and CNB models, were accurate and only had approximately 1.5 percent accuracy loss.

These suggest that sentiment analysis is a gateway to a more efficient feedbacking process. It opens the door to the importance and use-case of automated and automatic feedback evaluation, especially as it generally remains accurate once trained and optimized. This could change the way businesses collect customer feedback because the result is almost instantaneous and real-time, but it is more cost-effective and easier to implement. The contrast between the labor-intensive deployment of survey questionnaires and simply running a sentiment analysis program is enough for businesses to consider venturing and exploring this field.

## 2.6. Sentiment Analysis Algorithm: Semantic Analysis

Under the umbrella of sentiment analysis, one of the more popular algorithms is semantic analysis. This method was originally coined by Lewis (1990), a popular figure in the late 1900s. It was initially described as a lexical approach to teaching foreign languages. In recent years, academics have been experimenting with lexicon-based approaches for sentiment analysis. The underlying premise behind this technique is that understanding and producing lexical phrases as chunks is a crucial element of learning a language. When students are educated in this fashion, they are supposedly able to discern language patterns (grammar) better and have meaningful predefined uses of words at their disposal.

As per Saif et al. (2014), conceptual semantic sentiment approaches combine syntactic and semantic processing techniques to capture the latent conceptual semantic relationships in the text that communicate sentiment implicitly. Sentic Computing, for example, is a sentiment analysis paradigm in which common sense ideas (e.g., "happy birthday," "simple life") are collected from texts and assigned to respective sentiment orientations utilizing semantic parsing and effective common-sense knowledge sources. Gangemi et al. (2014), as cited by Saif et al. (2014), looked into the syntactic structure of sentences further to see whether there were any finer-grained relationships between the various semantic pieces inside it. This method can identify not just text sentiment but also opinionated themes, subtopics, opinion holders, and their sentiments. As a result, semantic approaches are more sensitive than syntactic methods to the hidden semantic links between words in texts. However, Saif et al. further explain that neither syntactic nor semantic approaches are adapted to Twitter in the prior studies due to the absence of linguistic formality and well-structured sentences in tweets. Furthermore, Semantic techniques are typically confined to the extent of their underlying knowledge bases, which is particularly difficult for analyzing generic Twitter streams, which are characterized by fast semiotic change and linguistic deformations.

Despite doubts regarding the accuracy of semantic algorithms, it nevertheless provides good performance and accuracy when trained. Referencing the study of Saif et al. (2014), the classifiers trained from their Semantic Sentiment Patterns (SS-Patterns) showed an overall consistent and superior performance over the other sets that they tested concurrently. This led the study to conclude that this approach can stay strongly consistent with the sentiment of the terms and can derive patterns of entities of controversial sentiments in tweets. Similarly, Gautam & Yadam (2014) revealed that semantic analysis using the WordNet database displayed an accuracy rating of 88.2 percent, which can be lifted to 89.9 percent when subjected to a unigram model.

Semantic analysis displays great accuracy and potential in analyzing user sentiments over an electronic, social platform where conventional sentence structures are not followed, such as Twitter. Compared to other methods such as Machine Learning, it presents the greatest accuracy, perhaps due to its non-conventional structure. The Semantic approach does not only take into account the arrangement of words and the structure of a particular statement, but it can also identify opinionated themes and various subtopics. However, other sentiment analysis approaches should not be disregarded because they may still provide considerable insights that the Semantic approach cannot convey.

## 2.7. Sentiment Analysis Algorithm: Machine Learning

Machine learning techniques is a general term that encompasses the use of a training set and a test set for classification. Neethu & Rajasree (2013) further expound that the training set is composed of input feature vectors and their corresponding class labels, used to develop a classification model that attempts to clarify the vectors into their corresponding class labels. The test set was utilized to validate the crafted model by predicting the class labels of unseen feature vectors.

Similar to the semantic analysis approach, one of the common use-cases of machine learning sentiment analysis is evaluating feedback on social media platforms. In this case, a study provided by Gautam & Yadav (2014) asserts that upon the conclusion of their study, the Naïve Bayes technique, a technique under machine learning, was the second most precise in analyzing Twitter data, with only a slight difference of 1.7 percent compared to the WordNet database. Similarly, Neethu & Rajasree (2013) study posed almost consistent results with the previous study, whereby the Naïve Bayes technique provided up to 79 percent accuracy compared to other classifiers.

It is also interesting to note that when compared to a human judgment, lexical-based techniques are generally 80.02 percent more accurate in identifying status signals (Ortigosa et al., 2014). However, Pratma & Sarno (2015) provided a contrasting result: the Naïve Bayes only provided an accuracy of 63 percent, but it slightly outperformed other methods such as the K-Nearest Neighbors (KNN) and the Support Vector Machine (SVM). They had accuracies of 60 percent and 61 percent, respectively. Nevertheless, other factors may have contributed to the low accuracy, such as the latent accuracy of the dataset itself. In contrast, the results of a study provided by Vijayarani & Dhayanand (2015), which determined whether SVM and Naïve Bayes could accurately predict liver disease, suggest that SVM is more accurate in terms of its classifying power, being able to classify 79.66 percent of the instances correctly. Compared to the Naïve Bayes, which could only correctly classify 61.28 percent of instances. However, the execution time of the Naïve Bayes algorithm was exceptionally lower, providing an almost instantaneous result while remaining reasonably accurate.

The Machine Learning method, specifically the Naïve Bayes algorithm, can remain reasonably accurate. However, previous studies would suggest that its results are not as accurate as envisioned, highlighting that the Naïve Bayes algorithm is potentially inferior to other methods. However, its efficacy should not be discredited - the execution time needed to run the Naïve Bayes algorithm is significantly lower than other methods. This approach, thus, is ideal for disciplines and especially businesses who wish to have an overview of the general community sentiment instantaneously. Perhaps providing it with a more accurate training data set could improve its accuracy and performance. The goal of this study is to optimize a program using the Machine Learning approach and compare it to other approaches, such as the Semantic approach.

### 2.8. Synthesis
Social media platforms have become the prime facilitator of connecting individuals and businesses. With the freedom of communication present in social media, businesses can use this to determine feedback and comments from their respective consumers. The same is applicable for banks, as they actively seek new customers and work to retain their current customer base. This is most significant in the case of BDO, whereby efforts to remedy the broken customer trust from the recent hacking incident via monitoring social media feedback and responding to them are being observed. However, traditional data and feedback gathering are costly and labor-intensive. This would mean added costs to the bank's back-end operations, which is undesirable considering its losses following the hacking incident. Therefore, automatic and automated feedback collection and evaluation are necessary, such as sentiment analysis. The umbrella of sentiment analysis is vast. However, both of its major approaches—Semantic Analysis or Lexicon-based and Machine Learning can present exceptional accuracy in determining polarity and sentiment. Both methods present exceptional accuracy, especially when combined with other methods. However, the WordNet classifier compared to the Naïve Bayes classifier presented more accuracy, confirming that the Semantic or Lexicon-based Approach was slightly superior in performance. Nevertheless, this does not discredit the Naïve Bayes, as it was still accurate, if only slightly more accurate than the other mentioned approach.

The literature referenced in this paper was intended to serve as a guide for the performance of the study. Further, these related articles provide a deeper understanding of how the techniques employed in this study have been used and their results. Given that the literature cited has generally followed a profound understanding of the disparity concerning the accuracy between the semantic and machine learning approaches, the study followed a general direction in its data analysis; however, results may still vary—these articles aided in analyzing the data.

### 3. Methodology
### 3.1. Research Design
This study's primary goal was to compare the Semantic, and Machine Learning approaches to Sentiment Analysis in terms of accuracy and implement the method with the highest accuracy in crafting improved marketing strategies for banks, specifically Banco De Oro (BDO) and hopefully other Philippine banks.

The researchers created a program modeled after the Machine Learning approach and compared its results with existing studies related to the Semantic approach. The program was trained using datasets mined from Twitter to improve its accuracy and provide points of comparison in terms of latent

performance. After the training, the program was compared between the accuracy of the Semantic approach, Naive Bayes, and J48. The one with the highest accuracy and best performance was recommended to serve as a guide in crafting improved marketing strategies for banks and their decision-making.

## 3.2. Research Participants, Respondents, and Local

The study required the collection of thousands of data from a random sample of users on Twitter. Given that the proposed software was trained to increase its accuracy in determining whether a particular sentiment is negative or not, there was no limit to the number of participants involved in this study. However, the feedback that was analyzed from the participants shall pertain to tweets directed to BDO. Thus, the participants were engaged with the said bank and had transacted with them before. This was regardless of their demographic profile. The participants were randomly selected via data mining techniques chosen by the proponents.

## 3.3. Data Collection

The researchers conducted their study using social network data collected through data mining techniques. However, social media can be very disorganized, and due to the freedom of communication, there is no specific structure or medium for collecting specific data. Therefore, the study collected data from tweets related to BDO due to its diverse set of feedback caused by the hacking incident. The inputs serve as the data that was inputted in the program. Data gathering was conducted via data mining Twitter's dataset.

## 3.4. Data Analysis

In connection with the data collection procedure, a data-mining program was used to extract datasets from Twitter. Once this data was extracted, it was inputted into the program to be developed that used the Machine Learning approach. This program classified the data gathered and trained the model for better accuracy in identifying sentiments. After the program finished running, its sentiment score and accuracy were compared to results from the Semantic approach and other machine language algorithms chosen.

Following the above-planned data analysis procedure, the following algorithms are relevant:

### 3.4.1. Naïve Bayes

Singh et al. (2017) previously conducted a study concerning optimizing sentiment analysis using machine learning classifiers. Under the said study, the Naïve Bayes classifier was defined as a widely used supervised classifier that allows users to express positive, negative, and neutral sentiments in online content. Conditional probability is used to categorize words into their appropriate groups, and its primary advantage is that it requires minimal training for the dataset. The raw web data is preprocessed, with numeric, foreign terms, HTML elements, and special symbols removed, providing a list of words. Human specialists execute the manual tagging of words with labels of positive, negative, and neutral tags. For the training set, this preprocessing generates word-category pairings. Consider the word 'y' from the test set (unlabeled word set) and a document window of n-words (x1, x2,. xn).

### 3.4.2. Semantic Analysis

In reference to the study of Geetika et al. (2014), a lexical database was employed under the Semantic Analysis approach. Specifically, the WordNet database was used. According to Geetika et al. (2014), this database is made up of related English words. When two words are semantically similar, they are considered close to one another. The main focus of this method was looking through the stored texts for terms. Afterwhich, it compared them to the words that the user put in their statements for any semantic similarities. This allowed the algorithm to determine the polarity of the statement. For example, in the statement "I am pleased," the adjective "happy" is chosen and compared to the stored feature vector for synonyms. Assume two words: 'glad' and 'satisfied', both are highly similar to the word 'happy'. Following the semantic analysis, the word 'glad' has replaced the word 'happy', resulting in a positive polarity.

### 3.4.3. J48

As cited by Singh et al. (2017), J48 is a decision tree-based classifier used to generate rules to predict target terms. Because of this, it can work with larger training datasets than the mentioned classifiers. It works because the word features it finds on the sentences taken from the training set are represented on leaf nodes of a decision tree. As a near feature qualifies the label condition of the internal feature node, its level is now raised up in the same branch of the decision tree. As the assignment of labels to the word features goes on, it creates two branches in the decision tree. The algorithm uses the Entropy function for testing the classification of terms from the training set. An example of this is that Bigrams like "Horrible acting", "Bad writing", and "very misleading" are labeled as negative terms. In contrast, "more enjoyable" is labeled as positive sentiment in the movie.

### 3.5. Trustworthiness of the Study

The reliability of the study was sufficiently followed, considering the availability of the datasets to be used for processing. The researchers have discovered that creating a dataset from BDO's page on the social media platform Facebook required time and effort. This was due to Facebook's Data Privacy Policy requiring the researchers to request consent from each related to the data needed to be compiled for the dataset. All of which were used to process results considering that there were thousands of consents needed to be approved for the researchers to be able to use the data legally. However, the researchers have thought of another way to legally collect the data necessary, which is through changing the social media platform for data collection that does not have a Data Privacy Policy in effect. Twitter, another social media platform being used today, is one of the platforms being considered by the researchers for data collection. They do not have the said policy in effect, which sped up the process of creating the datasets to be used. Aside from this matter, the reliability of the study was adequate for the researchers to work.

### 3.6. Ethical Considerations

The researchers considered the confidentiality of the individual participants of the research and their respective feedback and comments on the bank's page. An unbiased dataset was needed to garner the best results. With this in mind, the researchers only needed the participants' feedback and comments without their identities being revealed. This step was crucial to processing the sentiment analysis system to determine the results. This is to maintain the integrity of the dataset by impartiality on the different opinions of the individuals caused by different factors, such as individuals' relationship to the business

and public opinion. Keeping the individual's identity anonymous avoids sudden changes of opinion and protects the individual's identity and freedom of speech.

## 4.    Results and Analysis

### 4.1.  Program Specifications

The sentiment analysis program was developed using Jupyter Notebook and hard-coded via Python. Jupyter Notebook allows for developing open-source software and various services to facilitate interactive computing across multiple programming languages. The said language and program were used for their relative ease of use and simplicity. The table below summarizes the hardware and software requirements of the Jupyter notebook.

Table 1: Software Requirements

| | |
|---|---|
| **Operating System** | Windows |
| **Tools** | Python with Jupyter Notebook |

Table 2: Hardware Requirements

| | |
|---|---|
| **Processor** | Dual Core |
| **RAM** | 4 GB |

### 4.2.    Materials and Models

### 4.2.1.  Seaborn, pandas, RegEx, NLTK

Seaborn is a Python data visualization library based on matplotlib. The tool provides a high-level interface for crafting attractive and informative statistical graphics. Pandas is a software library written for the Python programming language for data manipulation and analysis. In particular, it presents data structures and operations for manipulating numerical tables and time series. The RegEx (Regular Expression) is a series of characters that forms a search pattern. The Natural Language Toolkit (NLTK) is a suite of libraries and programs for symbolic and statistical natural language processing for English written in the Python programming language.

### 4.2.2.  Multinomial Naive Bayes

The proponents utilized multinomial Naive Bayes to create the program for sentiment analysis. Multinomial NB employs the Naive Bayes algorithm for multinomially distributed data. The said component was where the data are commonly presented as word vector counts, although tf-idf vectors are also known to work well in practice.

### 4.2.3.  Apify

Apify is a software platform that offers the largest source of information ever created, including web scraping, web automation, and web integration. The proponents chose Apify for the data collection due to its Twitter Scraper, which enables the proponents to extract tweets with no Twitter API limits, download the data on an array of file types, and use it for training and testing immediately. Apify was chosen over RapidMiner, as stated in previous chapters because the proponents found that RapidMiner does not recognize tweets related to the BDO Hacking incident due to unknown causes. It may be asked to be hidden by the company, but there is no reliable or concrete evidence.

**Figure 2: Apify Twitter Scraper Actor Search Terms for Data Mining**

The figure above shows Apify's Twitter Scraper Actor interface and the following search terms the proponents used in determining what the actor bases its search on through the terms the proponents provide to the actor. The following search terms provided by the proponents are listed in the figure above.

**Figure 3: Apify Twitter Scraper Actor Settings**

The above figure shows the configuration of the Apify Twitter Scraper Actor. This allowed the proponents to gather only the necessary Tweets to perform sentiment analysis. This included settings for showing tweets only or if it includes replies and the maximum number of tweets gathered.
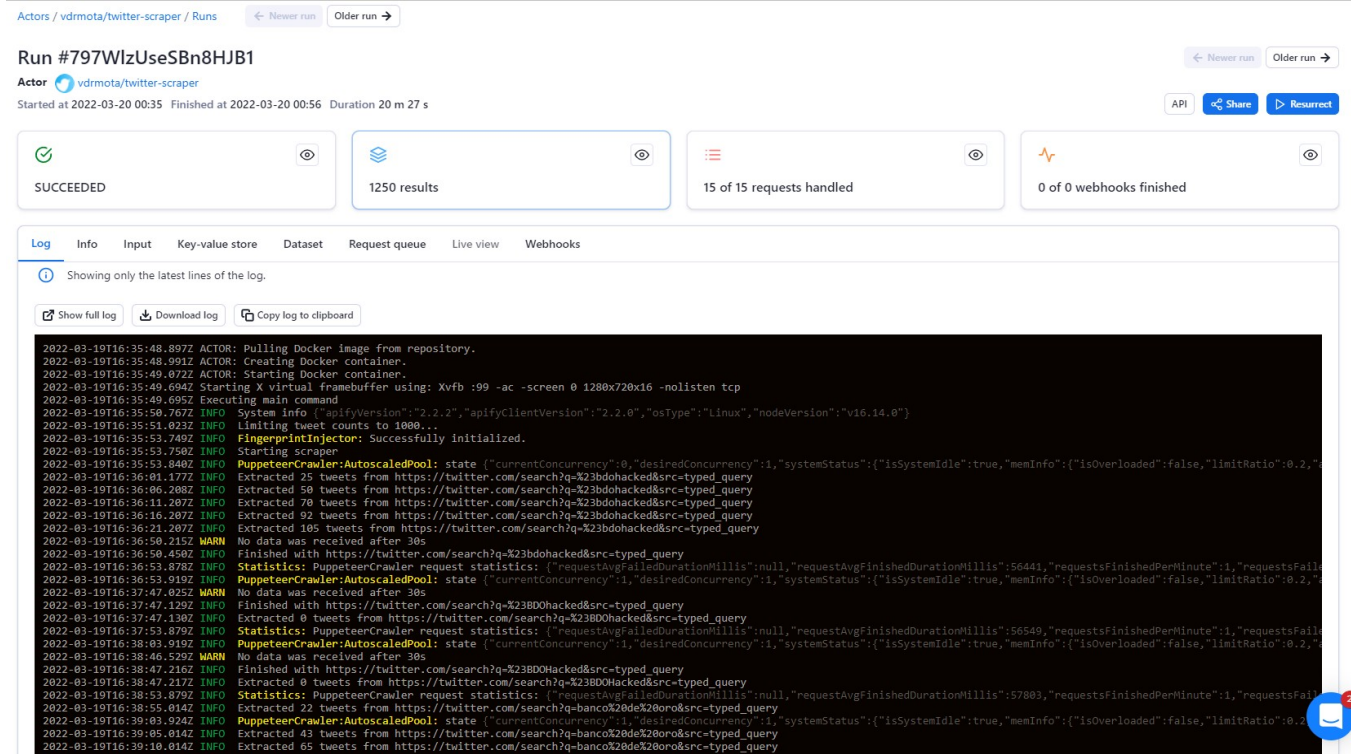
**Figure 4: Apify Twitter Scraper Data Mining Results**

The results of the Data Mining performed by the actor are shown in the above figure. A total of 1,250 Tweets or results have been gathered, which were subsequently reviewed and consolidated via Microsoft Excel.
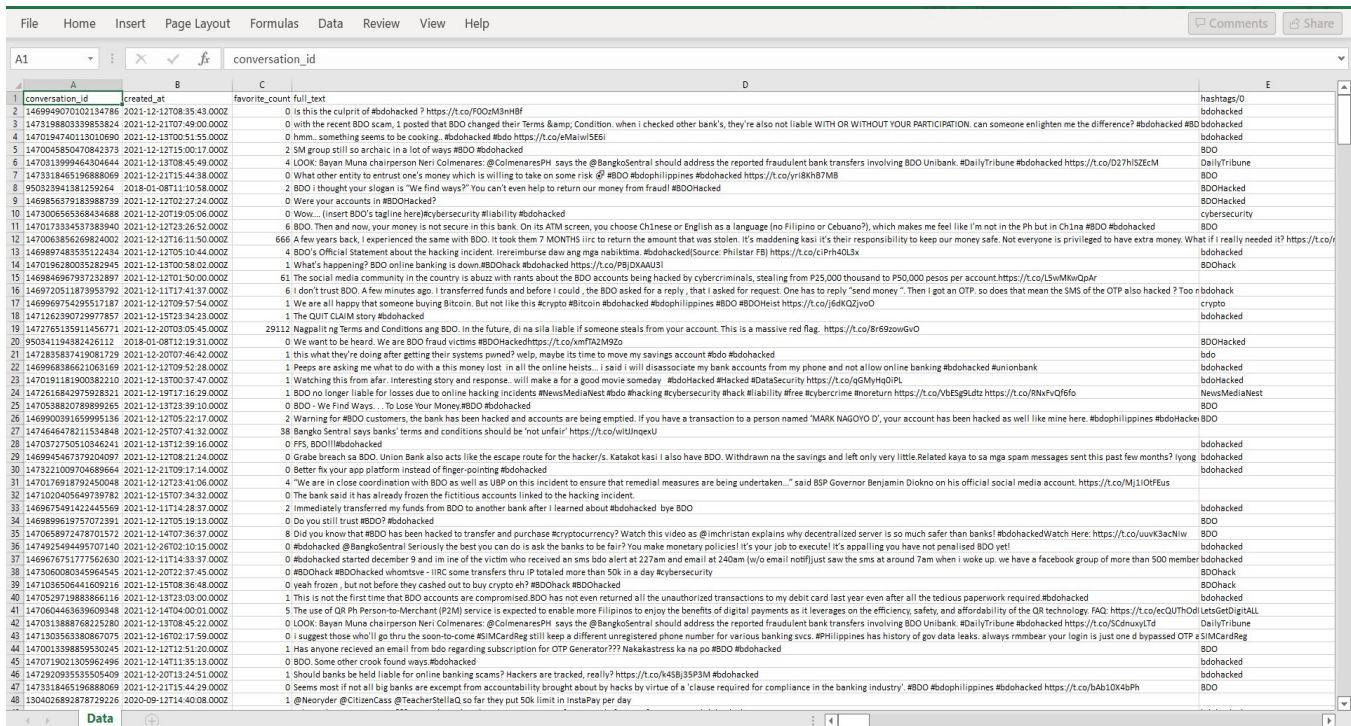


**Figure 5: Microsoft Excel Viewing Data Mining Results**

The Tweets returned by the actor were further consolidated and reviewed via Microsoft Excel, as portrayed in the figure above. The proponents manually removed Tweets that were not relevant to the subject matter (Banco De Oro or BDO) and Tweets that did not comment or say anything about the bank to properly create an accurate dataset training and testing phases of the program.

## 4.3. Model Testing

After the manual review and cleaning performed by the proponents, the final dataset is ready to be converted into CSV and imported into the program for training and testing. Figure 8 shows the clean dataset as a result of the review.

```
In [65]: tokenized_tweet = combine['tidy_tweet'].apply(lambda x:x.split()) #it will split all words by whitespace
         tokenized_tweet.head()

Out[65]: 0      [group, still, archaic, lot, ways, #BDO, #bdohacked]
         1                                                          [BDO, thought, your, slogan, find, ways, You, can, even, help, return, our, money, from, fraud, #BDOHacked]
         2      [BDO, Then, and, now, your, money, not, secure, this, bank, its, ATM, screen, you, choose, nese, English, language, Filipino, Cebuano, which, makes, feel, like, not, the, but, #BDO, #bdohacked]
         3      [few, years, back, experienced, the, same, with, BDO, took, them, MONTHS, iirc, return, the, amount, that, was, stolen, maddening, kasi, their, responsibility, keep, our, money, safe]
         4      [don, trust, BDO, few, minutes, ago, transferred, funds, and, before, could, the, BDO, asked, for, reply, that, asked, for, request, One, has, reply, send, money, Then, got, OTP, does, that, mean,...
         Name: tidy_tweet, dtype: object
```

**Figure 6: Sample of Tokenization**

**Removing short words (a,is,so etc..)**

```
In [63]: combine['tidy_tweet'] = combine['tidy_tweet'].apply(lambda x: ' '.join([w for w in x.split() if len(w)>2])) #removing words whose

In [64]: combine.head()
```

| | id | label | tweet | tidy_tweet |
|---|---|---|---|---|
| 0 | 1 | 1.0 | SM group still so archaic in a lot of ways #BDO #bdohacked | group still archaic lot ways #BDO #bdohacked |
| 1 | 2 | 1.0 | BDO i thought your slogan is □We find ways?□ You can□t even help to return our money from fraud! #BDOHacked | BDO thought your slogan find ways You can even help return our money from fraud #BDOHacked |
| 2 | 3 | 1.0 | BDO. Then and now, your money is not secure in this bank. On its ATM screen, you choose Ch1nese or English as a language (no Filipino or Cebuano?), which makes me feel like I'm not in the Ph but i... | BDO Then and now your money not secure this bank its ATM screen you choose nese English language Filipino Cebuano which makes feel like not the but #BDO #bdohacked |
| 3 | 4 | 1.0 | A few years back, I experienced the same with BDO. It took them 7 MONTHS iirc to return the amount that was stolen. It□s maddening kasi it□s their responsibility to keep our money safe. | few years back experienced the same with BDO took them MONTHS iirc return the amount that was stolen maddening kasi their responsibility keep our money safe |
| 4 | 5 | 1.0 | I don□t trust BDO. A few minutes ago. I transferred funds and before I could , the BDO asked for a reply , that I asked for request. One has to reply □send money □. Then I got an OTP. so does that... | don trust BDO few minutes ago transferred funds and before could the BDO asked for reply that asked for request One has reply send money Then got OTP does that mean the SMS the OTP also hacked Too... |

**Figure 7: Sample of Removing Short Words**

**Removing twitter handles**

```
In [61]: combine['tidy_tweet'] = np.vectorize(remove_pattern)(combine['tweet'].apply(str),"@[\w]*")
         combine.head()
```

| | id | label | tweet | tidy_tweet |
|---|---|---|---|---|
| 0 | 1 | 1.0 | SM group still so archaic in a lot of ways #BDO #bdohacked | SM group still so archaic in a lot of ways #BDO #bdohacked |
| 1 | 2 | 1.0 | BDO i thought your slogan is □We find ways?□ You can□t even help to return our money from fraud! #BDOHacked | BDO i thought your slogan is □We find ways?□ You can□t even help to return our money from fraud! #BDOHacked |
| 2 | 3 | 1.0 | BDO. Then and now, your money is not secure in this bank. On its ATM screen, you choose Ch1nese or English as a language (no Filipino or Cebuano?), which makes me feel like I'm not in the Ph but i... | BDO. Then and now, your money is not secure in this bank. On its ATM screen, you choose Ch1nese or English as a language (no Filipino or Cebuano?), which makes me feel like I'm not in the Ph but i... |
| 3 | 4 | 1.0 | A few years back, I experienced the same with BDO. It took them 7 MONTHS iirc to return the amount that was stolen. It□s maddening kasi it□s their responsibility to keep our money safe. | A few years back, I experienced the same with BDO. It took them 7 MONTHS iirc to return the amount that was stolen. It□s maddening kasi it□s their responsibility to keep our money safe. |
| 4 | 5 | 1.0 | I don□t trust BDO. A few minutes ago. I transferred funds and before I could , the BDO asked for a reply , that I asked for request. One has to reply □send money □. Then I got an OTP. so does that... | I don□t trust BDO. A few minutes ago. I transferred funds and before I could , the BDO asked for a reply , that I asked for request. One has to reply □send money □. Then I got an OTP. so does that... |

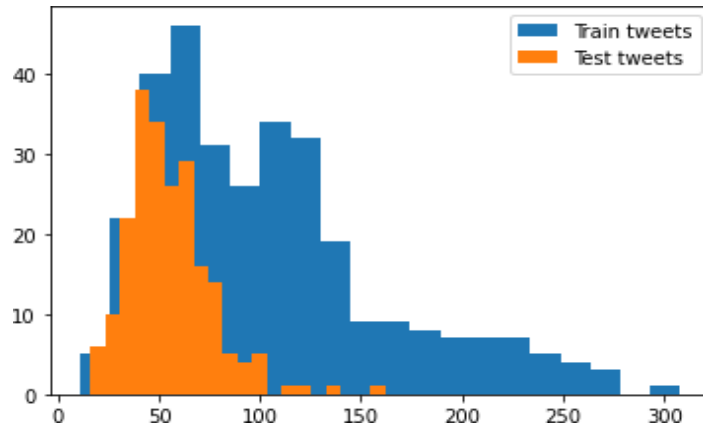**Figure 8: Sample of Removing Twitter Handles**

**Figure 9: Total Tweets for Training and Testing**

Following the creation of the final dataset, Figure 9 presents a graph of the total amount of tweets used for the training and testing of the developed model. The number of tweets used subsequently affects the runtime and accuracy results of the program.
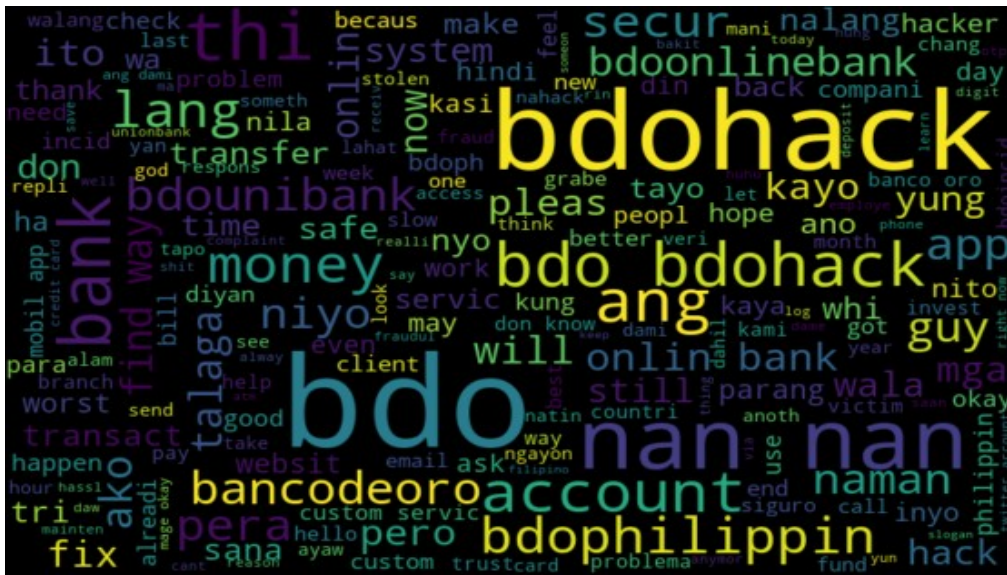


**Figure 10: Overall Word Cloud for the Dataset**

According to the collected Tweets, a word cloud was developed to visualize the saturation of Tweets better, as shown in the above figure. The larger the word and more visible, the more it is repeated in the dataset. Examples are bdohack and bdo.
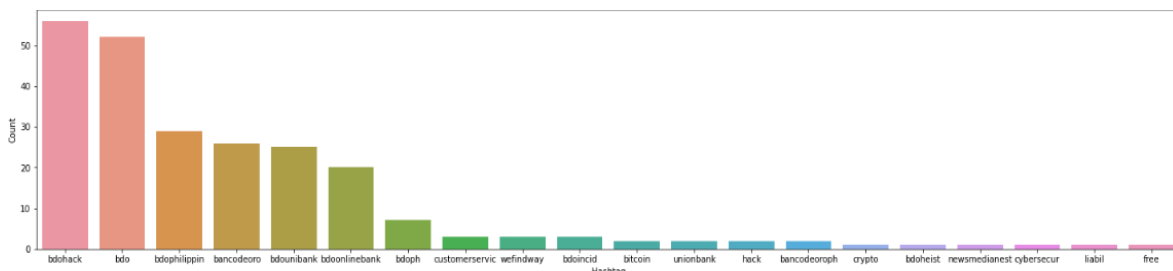


**Figure 11: Negative Tweets**

After importing the CSV file into the program, the above graph shows the disaggregation of the negative Tweets received by the dataset. It also portrays the number of times a particular word is repeated in the dataset.

**Table 3: Summary of Test Results**

| Test Number | Training Data | Testing Data | Accuracy |
|:---:|:---:|:---:|:---:|
| Test 1 | 101 | 428 | 63.21% |
| Test 2 | 213 | 330 | 86.82% |
| Test 3 | 213 | 428 | 87.60% |

A total of three tests were performed to measure the accuracy of the developed program. The results of each test are further discussed in the following paragraphs, however, the table alone indicates that Test 3 garnered the most significant accuracy, with 87.60%, as opposed to the two previous tests performed.

| Accuracy with Naive-Bayes: | 0.6320754716981132 |
|---|---|

**Figure 12: 1st Result with <u>101</u> testing Data sets and <u>428</u> training data**

In the first testing of the program, the first result utilized 101 testing data and 428 training data from the dataset. From the imported dataset, the program showed an accuracy of 63.21%.

| Accuracy with Naive-Bayes: | 0.8682170542635659 |
|---|---|

**Figure 13: 2nd Result With <u>213</u> Test Data and <u>330</u> training data**

In the second testing of the program, the accuracy of 86.82% was obtained from the following two primary phases: training and testing. The training phase was conducted using a training data set consisting of 330 data, while the testing phase was conducted using a testing data set of 213.

| Accuracy with Naive-Bayes: | 0.875968992248062 |
|---|---|

**Figure 14: 3rd Result With <u>213</u> Test Data and <u>428</u> training data**

In the third testing of the program, the accuracy of 87.59% was obtained from the following two primary phases: training and testing. The training phase was conducted using a training data set consisting of 428 data, while the testing phase was conducted using a testing data set of 213.

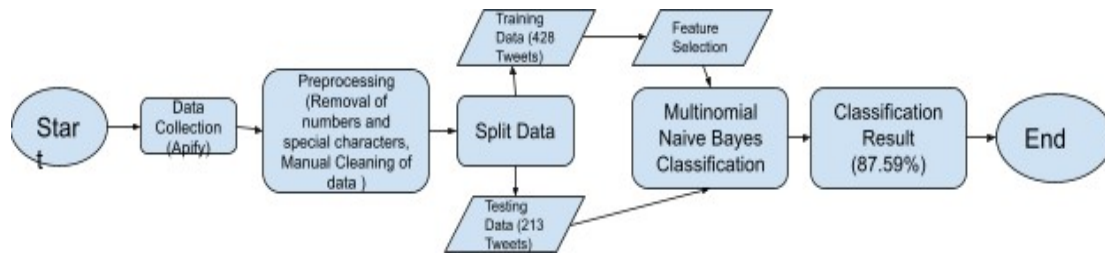# 5. System and Discussion

## 5.1. System Design



**Figure 15: System Flow Chart**

Figure 15 displays the flowchart of the system. The process starts when collected data is preprocessed and divided into Training Data and Testing Data. It starts using the Multinomial Naive Bayes for the classification. After the accuracy score has been given, the process ends.
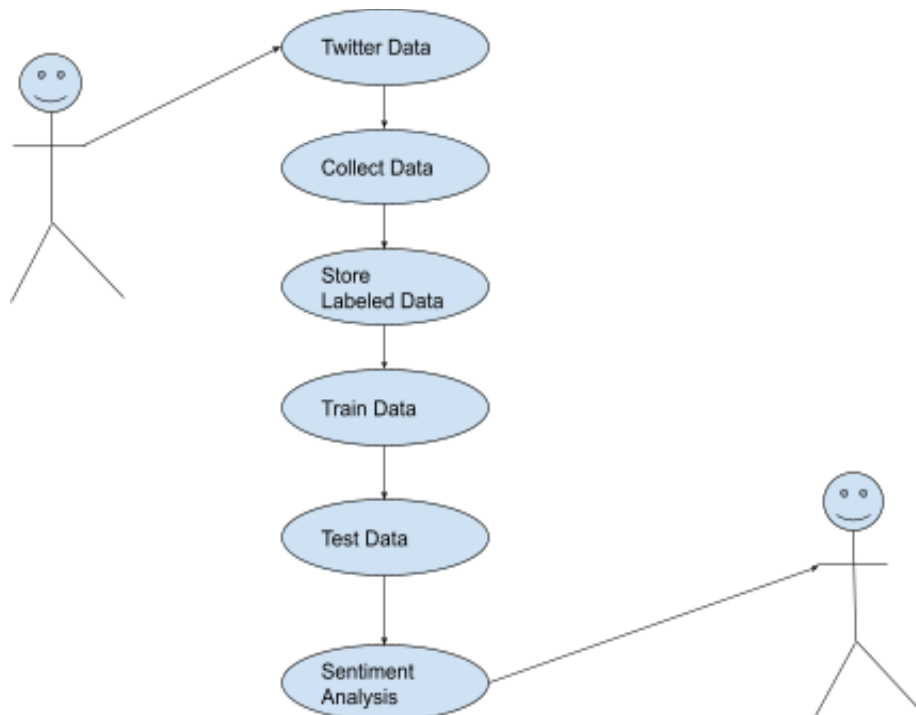
## 5.2. Use Case Diagram



**Figure 16: System Use Case Diagram**

Figure 16 displays the System Use case Diagram. The target is mainly Twitter data; after the data is collected, it is then stored and labeled if positive or negative. After that, the data is used for training and testing to get the sentiment score.
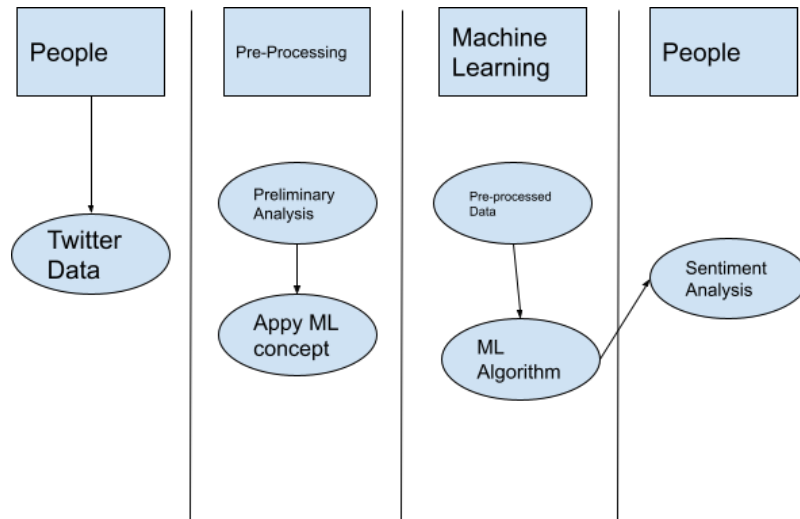
**5.3. Activity Diagram**



**Figure 17: System Activity Case Diagram**

Figure 17 displays the System Activity Case Diagram. The proceeding diagram illustrates the flowchart for both user and system. After the data gets collected from Twitter, it is pre-processed. After that, when the sentiment score is given, the person can now see if the accuracy is high enough to be significant.
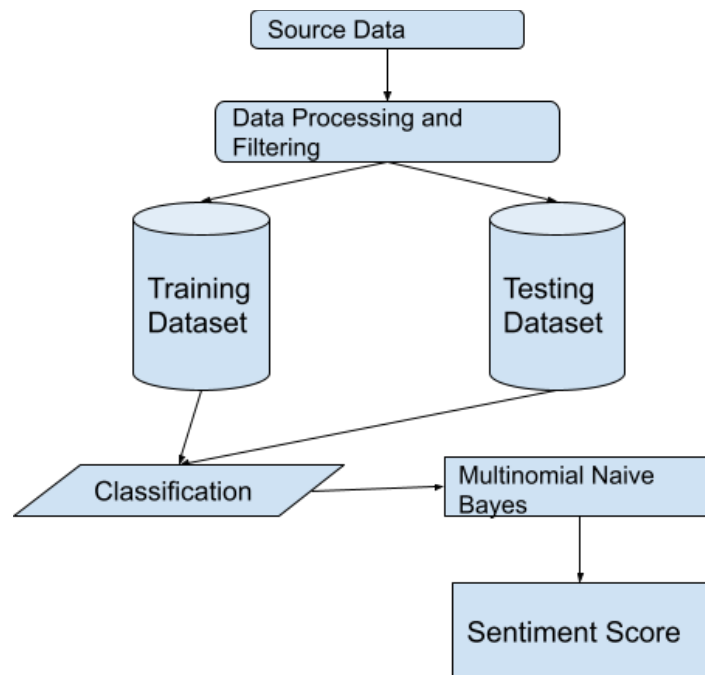
**5.4. Workflow Diagram**



**Figure 18: Workflow Diagram**

Figure 18 displays the System Work Flow Diagram. The proceeding diagram illustrates the flowchart for the system. This shows how the data is processed and filtered before the classification and sentiment score are applied.

The low amount of data mined was caused by the lack of tweets on the subject aimed by the proponents. The proponents also realized that the bank chosen had no official account on the chosen social media platform. This limited the interactions of the people with the bank and made the data more complex because it was disorganized in the platform, making it hard for the proponents to gather the data for the dataset.

The proponents decided to choose Apify's Twitter Scraper as the actor for the Data mining process instead of RapidMiner due to RapidMiner not being able to find any tweets related to the BDO Hacking incident. The proponents do not have any idea nor any concrete evidence as to what happened to the data and why the Data Mining program has chosen was not able to gather the data that the proponents asked for through the search terms provided.

## 5.5.  Analysis of Naïve-Bayes Accuracy

In the preceding chapter, two datasets were prepared for the sentiment analysis program to run: the first set of datasets included 101 testing data and 428 training data, while the second set of datasets included a much higher 213 testing data and a lower 330 training data. The first set introduced a lower accuracy of roughly 63.21 percent, which is understandable considering that only a tiny portion of the data collected was used to test the model's accuracy. Ideally, larger test datasets provide a more accurate calculation of model performance than if the test dataset contained a lower number of data. This is true in the second dataset, as the accuracy of the program increased by around 23 percent. It further reached new lengths with a new accuracy rate of 86.82 percent.

Several factors could be attributed to the change inaccuracy. First, the training data used in the first dataset is greater than the training data of the second dataset. While a larger training data is often advised to let the algorithm learn better, there is also the risk of overfitting the data to only that particular dataset. Hence it is preferred to record two instances of the model using different quantities for both testing and training datasets. Moving forward to the second dataset, more testing data is used with lower training data. This provided greater accuracy than the first dataset, by which the program was able to correctly identify more Tweets when a larger testing dataset was used.

Considering the factors stated above, the third set of data was used to test the model. Following 213 test data and 428 training data, the program presented an increased accuracy of 87.60 percent, which is 0.78 percent higher than the previously recorded 86.82 percent accuracy. This cements that while retaining the same test data but increasing the training data, the model becomes more accurate in performing sentiment analysis. In this case, sentiment analysis is best performed with extensive test and training datasets to generate the highest possible accuracy for the algorithm or model in question.

## 5.6.  J48 Algorithm

The J48 algorithm is explained by Singh, et al. (2017) as a decision tree-based classifier that is utilized to generate rules for the prediction of target terms. Compared to the Naïve Bayes algorithm, which requires the pre-processing of data, i.e., the removal of numbers, special symbols, and annotation of words, the J48 algorithm moves the test set along the branches of the decision tree once it satisfies the label conditions set in the internal feature node. However, the J48 algorithm is less used for sentiment prediction and more often applied in emotion recognition from text strings.

The J48 algorithm is remarkably slower in data processing compared to other known algorithms – however, it has shown good numbers in its accuracy. In the study of Singh et al. (2017), whereby they sought to optimize sentiment analysis using machine learning classifiers, J-48 took 49.73 seconds to render the data set, which was the slowest among the four algorithms they had tested. However, it was the third most accurate algorithm with an accuracy of 87.62 percent, just falling short of the OneR algorithm and BFTree algorithm. The Naïve Bayes algorithm was also tested in the same study. While the duration of its data processing was almost instantaneous, it was the least accurate among the four algorithms, with only 85.24 percent correctly classified instances compared to the others that showed an accuracy of more than 90 percent.

The comparison between Naïve Bayes and J48 in Singh et al. (2017) study is parallel to this study's data processing results. The data set was rendered line by line via Jupyter Notebook and took approximately 20 seconds to finish, with an accuracy of 87.60 percent. In comparison, this is still short of the accuracy shown by the J48 algorithm. The processing time was nearly half of what was recorded in the mentioned study, and there is only a $\pm 2$ percent difference in accuracy. The Naïve Bayes is still a unique algorithm, especially for larger data sets that may require a longer processing time. In conjunction with this, Singh, Singh, & Singh (2017) also mentioned that the J48 algorithm performs relatively better with smaller datasets and has a steeper learning curve than the Naïve Bayes, which showed faster learning capabilities.

## 5.7. Semantic Approach

The Semantic Approach incorporates semantic concepts (e.g., person, company, city) to represent more specific entities. In the study of Saif et al. (2012), which attempted to incorporate semantic features into their sentiment analysis of specific tweets, they assert that introducing these features can provide a more consistent correlation with positive or negative sentiment. This, in return, increases the accuracy of sentiment analysis by determining the sentiment of semantically relevant or similar entities.

The results of Saif et al. (2017) study reveal that the semantic approach outperforms the other three algorithms tested. It averaged 83.90 percent accuracy considering its 84.25 percent accuracy in positive sentiment identification and 83.80 percent accuracy in negative sentiment identification. The authors further explain that better Recall and F score when classifying negative sentiments was prevalent when semantic features were incorporated into the data set, and better Precision with lower Recall and F score in positive sentiment classification. (Saif et al., 2017).

The 83.90% accuracy of the Semantic Approach is less than the accuracy of this study's Naïve Bayes algorithm, which peaked at 87.60 percent. While the aforementioned semantic approach study claims that integrating semantic features in sentiment analysis leads to better Precision and accuracy, its accuracy was still less than that of the Naïve Bayes algorithm that was used. However, this could be because semantic features improve sentiment analysis accuracy for certain concept types while reducing accuracy in other concept types (Saif et al., 2017). Regardless, the Naïve Bayes algorithm that was used in performing this study's data processing is still superior in terms of raw accuracy.

### 5.8. Comparison of Algorithms

**Table 4: Comparison of Accuracy between Algorithms**

| Algorithm | Test No. | Accuracy | Implication |
|---|---|---|---|
| J48 Algorithm | Test 3<br>Test 2<br>Test 1 | 87.62%<br>76.86%<br>74.86% | Most accurate but has the greatest time delay in the data processing. |
| Naïve-Bayes | Test 3<br>Test 2<br>Test 1 | 87.59%<br>86.82%<br>63.20% | Sacrifices a small percentage of accuracy compared to the J48 Algorithm however its faster processing time and ability to learn quickly made it attractive. |
| Semantic Approach | Average | 83.90% | Similar processing time with Naïve Bayes but less accurate. |

Numerous algorithms are available for sentiment analysis use; however, it is essential to compare the latent accuracy of these algorithms and how they outperform other algorithms or how their underperformance could lead to some possible issues.

Recalling the prior discussion on the J48 algorithm, out of the three algorithms that have been compared (Naïve Bayes, J48, and the Semantic Approach), it presented the biggest accuracy at 87.62 percent. Naïve Bayes and the Semantic Approach, on the other hand, have accuracies of 87.60 percent and 83.90 percent, respectively. The J48 algorithm is an excellent tool for sentiment analysis–however, considering the time delay that occurs during its data processing and the fact that the time taken is exponentially larger than the two other algorithms combined. It may pose a problem when dealing with larger data sets. When considering the possible action that can be taken after analyzing the sentiment analysis results, J48 may introduce a time lag between the identification of issues and the resolution of the same.

On the other hand, the Semantic Approach is relatively near in processing time to the Naïve Bayes. However, its accuracy is much lower than the latter. Although the difference in accuracy between the two algorithms could be dubbed as almost negligible, ideally, sentiment analysis should be performed in the most accurate way possible. Furthermore, the ambiguity in accuracy when incorporating semantic features in certain concept types can also be a concern.

When directly compared to the two algorithms presented, the Naïve Bayes holds a good position of having decent accuracy with a shallower learning curve. It sacrifices a small percentage of accuracy compared to the J48 algorithm, and neither does it consider semantic features such as the Semantic Approach. However, its faster processing time and ability to learn quickly make it attractive for sentiment analysis. The same is also true for other studies that used the Naïve Bayes for sentiment analysis, such as the study of Neethu & Rajasree (2013). Although the Naïve Bayes, compared to the other algorithms they have tested, was not the most accurate and most precise, it still presented a good accuracy of 89.50 percent. As such, the Naïve Bayes algorithm presents an attractive measure in more practical, time-bounded situations.

## 6.    Conclusion, Recommendation, and Implications

## 6.1.  Conclusion

This paper aimed to develop and optimize a program that collects customer feedback from Twitter modeled after a Machine Learning algorithm, specifically the Naïve Bayes classifier, applied in Python. Further, this inquiry desired to compare the results of the developed program to existing studies related to the Semantic Approach. Additionally, the proponents aimed to recommend an algorithm for potential use by BDO and other Philippine banks. The researchers further utilized data mining techniques to effectively collect BDO-related customer sentiments on Twitter to achieve those goals. They probed deeper by comparing the individual accuracy of the developed program to the results of existing Semantic Approach studies and determining which among the algorithms is most accurate and superior.

The proponents were able to develop and optimize a program that is modeled after the Naïve Bayes classifier Machine Learning algorithm. However, the proponents manually fed the data collected from Twitter, specifically customer feedback related to BDO, to the program, which answers the first objective of the thesis.

The Data Mining techniques to collect BDO-related customer sentiments on Twitter were successful with Apify's Twitter Scraper by feeding the actor the proper search terms to effectively collect data related to the bank chosen. The first chosen Data miner, RapidMiner, was replaced because it could not scrape tweets related to the BDO Hacking incident. In this incident, the proponents found it odd that the said platform could not mine data, especially with the tags #BDOHacked and other related search terms. Apify, on the other hand, successfully collected several tweets related to the subject, although including plenty of unrelated tweets about the subject, which left the proponents to manually remove the unrelated tweets to properly create the dataset to be used for training and testing the program developed.

Despite the results of the data set, the proponents were able to achieve a score of 63.21 percent accuracy score using 101 testing data and 428 training data. The second test achieved a more impressive accuracy score of 86.82 percent accuracy score using 213 testing data and 330 training data with the program created utilizing the Naive-Bayes Classifier. Using 213 test data and 428 training data, the third test generated the highest accuracy of 87.60 percent. In this case, a higher number of testing and training data affects the accuracy score, given that the data collected by the proponents was much more robust than what they had to use.

Lastly, for the Algorithm Comparison, the proponents found out that the J48 algorithm presented the most significant accuracy at 87.62 percent out of the three algorithms compared. Naïve Bayes and the Semantic Approach, on the other hand, have accuracies of 87.60 percent and 83.90 percent, respectively, following the research found by the proponents for the J48 Algorithm and Semantic approach. The accuracy score may not be fair to compare due to the related studies not having the same dataset being used by the proponents. However, the datasets being used by the studies found are much more robust than the proponents were using.

The J48 algorithm is an excellent tool for sentiment analysis. Compared to the other two algorithms, the said tool had a longer time delay due to data processing. The Semantic Approach, on the other hand, is relatively near in processing time to the Naïve Bayes. However, its accuracy is much lower than the

latter. The Naïve Bayes holds a good position of having decent accuracy with a shallower learning curve. It sacrifices a small percentage of accuracy compared to the J48 algorithm, and neither does it consider semantic features such as the Semantic Approach. However, its faster processing time and ability to learn quickly make it attractive for sentiment analysis. Thus, this tool is the superior algorithm in terms of sentiment analysis. Further, it is the proper algorithm recommended by the proponents for the banks to use in creating their business strategies based on the sentiments gathered from the users of the bank.

## 6.2. Recommendation

One of the objectives of this study is to provide recommendations for BDO according to the recent hacking of their system, which caused thousands of losses from clients. The general sentiment was negative and driven by anger from the bank's slow interception of the hacking incident added to other existing concerns. As such, the following are recommendations for BDO after the conduct of sentiment analysis of Twitter users in response to the December 2021 incident and other problematic matters:

**(1) Improve Cybersecurity:** Because BDO and their online banking channel had been vulnerable to hacking, there is an imminent need to improve and invest in better cybersecurity. This suggestion could include more stringent protection over a user's online bank account or constant monitoring of transactions that seem to be malicious. As users constantly fear whether the incident may occur again, it is vital to keep them reassured that their funds are safe in BDO's online portal.

**(2) Improve Customer Service:** From the sentiment analysis conducted, several tweets mentioned that BDO was slow to respond to urgent issues. Users have mentioned that the bank is unresponsive, and the waiting times are long. Customer service portals such as 24/7 chat support could be made available to users, especially those who cannot visit the bank locally. This suggestion could improve response time and the general sentiment of the customers.

**(3) Implement Stricter Privacy Settings for Clients:** Many users have complained regarding the scam messages they receive that allegedly pretend to be BDO. Some of the messages include links that redirect to phishing sites, and some ask for OTPs of the users. In this case, BDO could invest in technology that can amplify protection for users' sensitive data and prevent them from receiving scam messages. They may also opt to issue official text messages reminding users to never click on links from messages that pretend to be from BDO, nor will BDO ever ask for their OTP.

**(4) Optimize Website and Application For Ease of Use:** Under BDO's slow response time is also their slow website and mobile application. Both must be optimized, primarily as these are used for client transfer of funds. The risk of the website and mobile application failing is that clients may lose their funds simply due to a technical glitch.

The sentiment analysis performed was only through the lens of the Naïve Bayes algorithm. While this study was still successfully able to compare its results with existing studies, it is best if future research would include the use of other algorithms that will process the same data set for a more accurate comparison. More accurate results could also be presented if J48 was integrated into the sentiment analysis. However, because of dataset limitations, the study could not pursue the J48 integration.

Therefore, a better dataset that is larger would be pertinent to improve the accuracy of the Naïve Bayes algorithm used.

The program was run via a console; however, this would be difficult to use for individuals who are not familiar with code. Given that the objective of this study is to be able to provide strategies for BDO and other banks alike based on the results of their sentiment analysis, it is to be considered that non-technology adept users could potentially use the program. The proponents recommend a potential front-end integration to the program for better ease of use, especially from an employee or management-level perspective. Finally, future research can look into better optimization of the code to make the program run efficiently and smoothly. This suggestion can also produce better results and more accurate processing of the datasets.

### 6.3. Implications

This study provides an alternative solution to BDO's pertinent issues using sentiment analysis and providing recommendations to BDO and other banks based on user sentiment. The study is rooted according to the proponent's objective of wanting banks to address their system defects. This can be achieved through understanding the sentiments of their user base. The researchers also wanted to integrate the J48 decision tree into the program; however, the dataset composed of data-mined information from Twitter was insufficient for the successful integration of J48 in classifying the data.

### Authors' Biography

Alexander Manuel Dean T. Ferrazzini, Leo Neil P. Flores Justin Jay L. Joaquin are 4th year BS in Computer Science Students.

Cherry B. Lisondra is a faculty member and a Program Coordinator of BSIS Program in the College of Computer and Information Science at the Mapua Malayan Colleges Mindanao. A degree holder of BS in Information Technology, with a Masters in Information Technology, and completed my academic requirements in Doctor in Information Technology.

## References

[1] Aldaihani, F.M.F. & Ali, N.A.B. (2019). Impact of relationship marketing on customers loyalty of Islamic banks in the state of Kuwait. International Journal of Scientific & Technology Research, 8(11), 788.

[2] Altrabsheh, N., Cocea, M., & Fallahkhair, S. (2014). Sentiment analysis: Towards a tool for analyzing real-Time students feedback. 2014 IEEE 26th International Conference on Tools with Artificial Intelligence. https://doi.org/10.1109/ictai.2014.70

[3] Akter, S., & Aziz, M. T. (2016). Sentiment analysis on facebook group using lexicon based approach. 2016 3rd International Conference on Electrical Engineering and Information Communication Technology (ICEEICT). https://doi.org/10.1109/ceeict.2016.7873080

[4] Bhatia, S., Li, J., Peng, W., & Sun, T. (2013). Monitoring and analyzing customer feedback through social media platforms for identifying and remedying customer problems. Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. https://doi.org/10.1145/2492517.2500287

[5] Cepeda, M. (2021, December 12). BDO clients lose money due to alleged online banking hack. Rappler. https://www.rappler.com/business/bdo-clients-lose-money-due-alleged-online-banking-hack/

[6] Chipongian, L.C. (2021, December 12). BDO to reimburse affected clients soon; bankers group call for vigilance against cybercrimes. Manila Bulletin. https://mb.com.ph/2021/12/12/bdo-to-reimburse-affected-clients-soon-bankers-group-call-for-vigilance-against-cyber-crimes/

[7] Corporate Profile. BDO. https://www.bdo.com.ph/about-bdo/corporate-profile-032112

[8] Dhaoui, C., Webster, C., & Peng Tan, L. (2017). Social media sentiment analysis: lexicon versus machine learning. Journal of Consumer Marketing. https://doi.org/10.1108/JCM-03-2017-214

[9] Dumlao-Abadilla, D. (2021, December 14). 700 BDO clients assured of reimbursement. Philippine Daily Inquirer. https://business.inquirer.net/336683/700-bdo-clients-assured-of-reimbursement

[10] Dumlao-Abadilla, D. (2018, January 10). BDO apologizes for fraud incidents, vows to intensify investigation. Philippine Daily Inquirer. https://business.inquirer.net/243881/bdo-apologizes-account-breaches-vows-intensify-fraud-investigation

[11] Dwivedi, Y. K., Ismagilova, E., Hughes, D. L., Carlson, J., Filieri, R., Jacobson, J., Jain, V., Karjaluoto, H., Kefi, H., Krishen, A.S., Kumar, V., Rahman, M.M., Raman, R., Rauschnabel, P.A., Rowley, J., Salo, J., Tran, G.A., Wang, Y. (2020). Setting the future of digital and social media marketing research: Perspectives and research propositions. International Journal of Information Management, 1-37. https://doi.org/10.1016/j.ijinfomgt.2020.102168

[12] Fernandez, D. (2021, December 14). Inside job a probability in recent hacking of bank accounts, says BSP. Philippine Daily Inquirer. https://business.inquirer.net/336650/inside-job-a-probability-in-recent-hacking-of-bank-a ccounts-says-bsp

[13]  Gautam, G., Yadav, D. (2014). Sentiment analysis of twitter data using machine learning approaches and semantic analysis. Proceedings of the 2014 Seventh International Conference on Contemporary Computing (IC3). 437–442. https://doi.org/10.1109/IC3.2014.6897213

[14]  Hasan, A., Moin, S., Karim, A., & Shamshirband, S. (2018). Machine learning-Based sentiment analysis for Twitter accounts. Mathematical and Computational Applications, 23(1), 11. https://doi.org/10.3390/mca23010011

[15]  Hasson, S. G., Piorkowski, J., & McCulloh, I. (2019). Social media as a main source of customer feedback. Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Network Introduction to consumers. (n.d.). Lumen Learning. (Retrieved on December 5, 2021) https://courses.lumenlearning.com/boundless-marketing/chapter/introduction-to-consumers/

[16]  Is the digital market too saturated to stand out? (2020, February 4). M&R Group. (Retrieved on December 4, 2021) https://www.mandr-group.com/is-the-digital-market-too-saturated-to-stand-out/

[17]  Kaura, V. (2013). Antecedents of customer satisfaction: A study of Indian public and private sector banks. International Journal of Bank Marketing, 31(3), 167–186. https://doi.org/10.1108/02652321311315285

[18]  Kumar, A., & Jain, R. (2015). Sentiment analysis and feedback evaluation. 2015 IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE). https://doi.org/10.1109/mite.2015.7375359

[19]  Loftin, S. (2020, June 30). The rising role of marketing in banking. ABA Bank Marketing. https://bankingjournal.aba.com/2020/06/the-rising-role-of-marketing-in-banking/

[20]  Luna, F. (2021, December 21). Fact check: Did BDO change terms and conditions to remove liability? The Philippine Star. https://www.philstar.com/headlines/2021/12/21/2149309/fact-check-did-bdo-change-term s-and-conditions-remove-liability

[21]  Masocha, R., Chiliya, N., & Zindiye, S. (2011). The impact of technology on competitive marketing by banks: A case study approach. African Journal of Marketing Management, 3(3), 68-77.

[22]  Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal, 5(4), 1093–1113. https://doi.org/10.1016/j.asej.2014.04.011

[23]  Meidan, A. (1984). Bank Marketing Management. Macmillan Publishers Limited 1984.

[24]  Milner, R., & Furnham, A. (2017). Measuring Customer Feedback, Response and Satisfaction. Psychology, 8, 350-362. https://doi.org/10.4236/psych.2017.83021

[25]  Misra, S., & Li, Hao. (2020). Noninvasive fracture characterization based on the classification of sonic wave travel times. In W. Zgallai (Ed.), Biomedical signal processing and artificial intelligence in healthcare. Elsevier. https://doi.org/10.1016/C2018-0-04775-1

[26]  Nasr, L., Burton, J., Gruber, T., & Kitshoff, J. (2014). Exploring the impact of customer feedback on the well-being of service entities. Journal of Service Management, 25(4), 531–555. https://doi.org/10.1108/josm-01-2014-0022

[27]  Neethu, M. S., & Rajasree, R. (2013). Sentiment analysis in twitter using machine learning techniques. 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT). https://doi.org/10.1109/icccnt.2013.6726818

[28] Pratama, B.Y., & Sarno, R. (2015). Personality classification based on Twitter text using Naïve Bayes, KNN and SVM. 2015 International Conference on Data and Software Engineering (ICoOSE), 170-174. https://doi.org/10.1109/ICODSE.2015.7436992

[29] Rootman, C., Tait, M., & Sharp, G. (2012). Relationship marketing and customer retention lessons for South African banks. African Journals Online, 15(3), 184-185.

[30] Sheldrake, P. (2014). The Business of Influence: Reframing Marketing and PR for the Digital Age. John Wiley & Sons.

[31] Saif, H., He, Y., Fernandez, M., & Alani, H. (2014). Semantic patterns for sentiment analysis of Twitter. Lecture Notes in Computer Science, 324–340. https://doi.org/10.1007/978-3-319-11915-1_21

[32] Singh, J., Singh, G., Singh, R. (2017). Optimization of sentiment analysis using machine learning classifiers. Human-centric Computing and Information Sciences, 7(1), 32. https://doi.org/10.1186/s13673-017-0116-3

[33] Terkan, R. (2014). Importance of creative advertising and marketing according to university students' perspective. International Review of Management and Marketing, 4(3), 239-246. Retrieved from https://dergipark.org.tr/en/download/article-file/366692

[34] The importance of customer feedback. (2021, October 5). Chron. (Retrieved on December 5, 2021) https://smallbusiness.chron.com/importance-customer-feedback-2089.html

[35] Vijayarani, S., & Dhayanand, S. (2015). Liver disease prediction using SVM and Naïve Bayes algorithms. International Journal of Science, Engineering and Technology Research (IJSETR), 4(4), 816-820.