Sentiment Analysis of Filipinos About Online Classes in the Post-Pandemic Era using Machine Learning and Code Switching

Abstract: Understanding the issue faced during and after the pandemic is a concern of the educational system in the country, particularly the feelings of teachers and students regarding the matter. In this post-pandemic era of the years 2021-present, online learning continues to be a norm and its implementation still causes issues for some students and teachers alike. Student satisfaction and Instructor preparedness are both key factors to effective delivery of education through online learning. This study aimed to develop a model that analyzes the Filipino public sentiment on the implementation of online classes using Machine Learning. Result revealed that the Filipinos sentiment towards online classes are predominantly positive. The output of the study will be used for better and improved delivery of online classes in the Philippines in the post covid-19 pandemic era.

Keywords: Machine Learning, Naïve Bayes, Online Class, Sentiment Analysis.

I. INTRODUCTION

Covid-19, a highly contagious and deadly virus which sprung from its country of origin, spread all over the world, not sparing the Philippines and its people. Philippine society was experiencing the alarming effects of the pandemic. Life in all its aspects was affected, impacting tremendously schools and education. According to the United Nations Educational, Scientific and Cultural Organization (UNESCO), over 800 million learners from around the world have been affected, 1 in 5 learners cannot attend school, 1 in 4 cannot attend higher education classes, and over 102 countries have ordered nationwide school closures while 11 have implemented localized school closure [29]. In March 2020, the number of positive cases was increasing which compelled the Philippine Government to mandate the temporary closure of all school activities due to the coronavirus infection. After the DOH reported the first COVID-19 case in the country, acquired through local transmission, the Philippine President Rodrigo R. Duterte signed Proclamation No. 922 on 8 March 2020, declaring the country under a state of public health emergency. This was to prevent the further spread of the disease and mitigate its effects on communities [26]. This prevented teachers and students from having classes in the traditional way, face-to-face interaction. As an emergency alternative for education to continue, online classes were adopted.

With the start of online classes as the new mode of teaching-learning, various difficulties were encountered by both teachers and students, among which, are learners’ motivation, satisfaction, and interaction. Tria [27], pointed out that changes on the grading system, assessment and evaluation of students’ performance also became a challenge to every administrator. However, in the course of time, the teaching-learning community was able to adjust and achieve results using this new form of instructional delivery. Rodriguez [24], mentions in his article that Filipinos have mixed experiences with online classes. Undoubtedly, online learning has its downsides, but its benefits cannot be disregarded as well. To fully understand its effects, the study must consider that the Philippines is a multilingual country, and its national language is Filipino also known as Tagalog, a language used by the entire nation even in casual conversation. Hence, in order to account for its use, the study employed code switching techniques to gather the sentiments of those using the Filipino language in their tweets. This is a reality of our post pandemic society which brings an issue that needs to be addressed: how do the teachers and students feel about the implementation of online classes and its probable continuous implementation?

The research limitations include the following: (1) The research only focused on the sentiments of Filipinos who have posted tweets regarding online classes, which may not represent the entire population who are going to school in the Philippines. (2) The research findings are limited to the sentiments expressed which may not accurately
represent the tweets’ true attitude and perception towards online learning. (3) The research also focused on two of the most common languages in the Philippines which are English and Filipino. Lastly, the study did not examine the factors that may affect efficiency of teaching and learning online classes compared to face-to-face classes, such as the quality of instructional materials and teaching methods that teachers use. Despite these limitations, the research findings may provide insights into the sentiments of Filipinos towards online classes, which can be useful for educational institutions to determine whether this method of teaching can provide quality education for students living in the Philippines.

II. RELATED LITERATURE

In order to establish the importance of the study, this chapter looked into related studies that support the understanding of existing research and looked into the gaps and recommendations of these research that may be useful in strengthening the present study.

A. Online Class

Online Education has been used as a modality even prior to the Covid-19 pandemic. However, with the pandemic, online education was not an option, it became a necessity to ensure continuity of learning. Noteworthy is the important role played by the Learning Management System in the delivery of instruction and the student’s reception of it as student satisfaction is an important indicator of the quality of learning experiences, thus the environment of the online course is critical to its success [19][31][25]. In addition, the new generation of students, who have become digital natives, are engaged in the use of multiple types of technology and social media [25]. Some of the tools used in online courses include the following: (a) e-mail, (b) chat rooms, (c) threaded discussions, (d) bulletin boards, (e) file transfer protocol, and (f) digital audio and video [5]. This growth is a result of several factors which include program marketability, student adaptability, and convenience, and the rapid growth of multimedia technology and internet access.

Studies show that instructor variables are the most important factor when it comes to student satisfaction in the online environment. Student satisfaction has a strong positive correlation with the performance of the instructor, particularly with his or her availability and response time [11][15]. Instructors should pay attention to students and provide feedback to students in a timely fashion or encourage students to ask questions through different mechanisms [32]. Furthermore, Rothman, Romeo, Brennan, Mitchell [25] pointed out that students appear to require more immediate feed-back from the professor than in a traditional, on campus course such as frequent, written communication regarding threaded discussion and content. The instructors of online classes must adapt, polish their technical skills, and foster new students’ technical knowledge in this environment. This is essential as instructor quality is a significant determinant of student satisfaction during online classes amid a pandemic [13]. The second factor that explains student satisfaction is technology [7]. Students need to have access to reliable equipment both personally and on the part of the institution. Online access is one of the most important factors influencing student satisfaction [9]. Students must have access to reliable equipment and must be familiar with the technology used in the course to be successful [5]. Lack of confidence in using information and communication technology (ICT) may decrease students’ satisfaction during online instruction and in turn lower their performance [32].

The third factor is interactivity [7]. Collaborative learning tools can improve student satisfaction in the online learning environment [8][14]. Online learners must be given plenty of opportunities to participate in discussions in order to feel involved and stay engaged in online courses. Yu-Chun Kuo, Belland and Schroder [32] said that instructors and course designers should pay attention to content design and organization given that learner-content interaction substantially contributes to student satisfaction. Moreover, instructors are encouraged to design more collaborative activities in undergraduate courses to enhance learner-learner interaction. Flexibility is a principal benefit for online courses [17]. Students are no longer tied to a specific place and time. Hence, studies done consider online learning as a valuable learning experience. This may be due to the novelty effects of the online course [30]. Because it is a new or different way to learn, its “novelty” creates a perception of increased value.

This may be because face to face interaction is not a must. Students can work at their own pace and take time to understand and appreciate the material [17]. E-learning makes the exchange easier for both student and teacher even during pandemics as it enables them to effectively fit learning around their own lifestyle [3].
B. Sentiment Analysis

The objective of sentiment analysis is to accurately extract people’s opinions from many unstructured review texts and classify them into sentiment classes, i.e., positive, negative, or neutral. It is also called opinion mining, which refers to the use of natural language processing and text mining to identify the emotional information from text minerals [6]. Word of mouth (WOM) is the process of conveying information from person to person and plays a major role in customer buying decisions [20]. People relay their feelings or opinions regarding products, services, and business to their social network, influencing consumer satisfaction. With the growing availability of resources on the internet using online review sites, blogs, and social networking sites, it made the “decision-making process” easier. Sentiment analysis then has various purposes such as finding if a recently released movie is good or bad, knowing consumers’ opinions to get feedback on one’s product, assisting consumers themselves by giving insight whether a product is worth purchasing, and many more. Sentiment analysis approaches aim to extract positive and negative sentiment bearing words from a text and classify the text as positive, negative, or else objective if it cannot find any sentiment bearing words [20]. When taken into perspective, it can be thought of as a text categorization task given with 3 broad categories. Although sentiment analysis has 3 categories, it is shown to be much more complicated than text classification. The study then provided various instances where challenges would occur when categorizing sentiments. An example is handling implicit sentiment and sarcasm. A sentence may possibly have an implicit sentiment even without providing any sentiment bearing words.

Hussein [16] discussed the importance and effects of sentiment analysis challenges in sentiment evaluation based on two comparisons among forty-seven papers. The study concluded in its first comparison that the topic nature and the review structure determine the suitable challenges for the evaluation of sentiment reviews. Students express their feelings and opinions regarding an issue through social media. Social network sites and microblogging sites are considered very good sources of information because people share and discuss their opinions about a certain topic freely [18]. Ullah, Marium, Begum, and Dipa [28] presented that the age of getting meaningful insights from social media data has now arrived with the advance in technology. Furthermore, they elaborate that over the years, people considered emoticons as a medium of communication that is used in texts or solely to dedicate one’s sentiments in an efficient manner, emoticons are symbolic expressions of happiness, frustration, anger, sadness, etc. which are then classified in positive, negative, and neutral. Hence, they are considered as “data sources in the sentiment analysis process”.

C. Naïve Bayes

Naïve Bayes is a popular and simple classification algorithm based on the Bayes Theorem, used in a variety of classification tasks. The classifier computes the posterior probability of a class and considers the way words are distributed in the document [18]. The classifier was compared with the K-Nearest Neighbor (K-NN) to movie review datasets and found out that the NB approach produced better results than the K-NN as it provided 80% accuracy [12]. It has then been used in various studies that involve analysis of sentiment.

Naïve Bayes has mostly been used to analyze data from social media such as Twitter, Facebook, Instagram, and others. Twitter has been one of the main sources of gathering data because through this medium people directly share their opinions with the public. Pratama, Tampubolon, Sianturi, Manalu and Pangaribuan [22] have implemented such techniques to know the people’s responses to election debates to some public figures. The researchers have used specific keywords to be calculated for each category with the NB classifier to get the value of the sentiment of the people in Jakarta to the three candidates for the Jakarta governor.

D. Support Vector Machine

A support vector machine (SVM) is a powerful and versatile supervised machine learning model that is used for classification and regression analysis. It is a type of linear classifier that works by finding the optimal hyperplane that separates the data points into different classes. Some advantages of using SVMs include effectiveness in high-dimensional data, versatility to use kernel functions for a specified function, and efficiency in conserving memory. Due to its advantages, it has been one of the popular choices in analyzing sentiments.

Ahmad, Aftab, and Ali [2] presented SVM being used in analyzing the sentiments of tweets. They tested the performance of the SVM model with two pre-labeled datasets that contained tweets about self-driving cars and apple products. The researchers concluded that the SVM model’s performance is dependent on the input dataset based on the results of the analysis. They further proposed that it should be investigated as to what machine algorithm would perform better on what specific type of dataset as this would lead to better and improved versions of the algorithms for sentiment analysis.
It has been proven that the performance of the machine learning algorithm would depend as to what type of dataset it is being implemented on [2]. Rahat, Kahir, Ma-summ [23] compared the SVM model to another popular algorithm in machine learning, Naive Bayes classification. Both algorithms have been trained and tested for sentiment analysis of airline reviews gathered from Twitter. The SVM model has the highest accuracy and precision with 82.48 and 90.33, respectively while Naive Bayes has 76.56 and 89.00. In other metrics such as recall and F-measure, Naive Bayes provided better results having 83.75 and 86.37 respectively while SVM had 81.79 and 85.85.

E. Lexicon-Based Classification

Lexicon-based classification is a type of text classification that uses a predefined set of opinion words to classify text into positive, negative, or neutral categories. In this approach, the text is assigned labels by comparing the number of words that appear from two opposed lexicons, such as positive and negative sentiment. A semantic orientation refers to a numeric value used to express the sentiment and intensity of words or expressions.

Bari, Sharoff, and Thomas [4] had pointed out in their research that the commonly used Bag-of-Words approach in sentiment analysis is not ideal because it ignores how the meaning of each word changes depending on the context of the sentiment and the relationship between words. They have proposed a way to automatically generate a Twitter-specific opinion lexicon through an annotation schema called “SentiML”. The schema has been developed to facilitate identifying and analyzing the targets (object, person, or concept) of the opinions, and their modifiers (what modifies the target), along with their linkage (appraisal group). The appraisal group has four possible values, positive, negative, neutral, or ambiguous.

F. Code Switching

Code switching is the practice of shifting between two languages or between two dialects or registers, depending on the social context or conversational setting. It is studied by linguists to examine when people do it, such as under what circumstances do bilingual speakers switch from one to another, and it is studied by sociologists to determine why people do it, such as how it relates to their belonging to a group or the surrounding context of the conversation [21]. In natural language processing (NLP), it poses challenges in analyzing sentiments because each sentiment would have different languages based on where the data were collected. Cetinoglu, Schulz, and Vu [10] mentioned that context-sensitive methods would suffer due to increased combinatorial possibilities crossing syntactic and lexical systems of different languages. They further emphasized that code switching genres are often close to spoken text which leads to dealing with problems that colloquial text poses from non-canonicity to incomplete syntactic structures to out-of-vocabulary words.

The Philippines is a diverse linguistic environment with more than 8 major languages spoken and a complicated language policy affected by its colonization history [1]. Because of the country’s diversity in conversing with one another, it provides us reason when gathering sentiments to use code switching techniques because it denotes the shift of language from the data being collected. This makes collecting data much easier for researchers rather than going through the contents of every sentiment to understand the user’s input.

III. METHODOLOGY

This section elaborates and defines the methods used in training the models used for this study. It also presents the metrics used in testing the data. The conceptual framework presented below show the gathering of the Filipino’s sentiments on online classes with the use of Naïve Bayes and Support Vector Machine with Code Switching.

![Tweet Data Analysis Conceptual Framework](image-url)
A. Tweet Scraping

Data were gathered from Twitter using snscrape implemented in Python. A total of 5,839 Tweets were scraped. Removal of duplicate Tweets was performed which reduced the dataset to 5,296 Tweets.

B. Code Switching and Conversion

The dataset is then code switched and preprocessed by removing unnecessary words, special characters, hyperlinks, and stop words. After preprocessing, the data are converted to lower case.

C. Tokenization and Lemmatization

Lemmatization is the process of grouping together similar words so that they can be analyzed as a single item and then converted to their base form. Tokenization is then performed afterwards for processing. It is the means of converting a sentence or a string into a sequence of lexical tokens. Both processes have been imported from the NLTK library.

D. Opinion Lexicon

SentiWordNet from the NLTK library was used as the opinion lexicon for feature extraction. It breaks down the data using POS (part of speech) tagging to determine the word class of each term. It also determines the polarity of the text, whether it is positive, neutral, or negative. If the sentiment is higher or equal to 0.05, it is determined that the Tweet is positive, if it is lower or equal to -0.05, it is negative, and if it is 0, then it is neutral.

E. Machine Learning Algorithm

Both Multinomial Naïve Bayes and SVM models have similar processes as they only differ on training and testing the model.

The TF-IDF Vectorizer was used as it utilizes both Count Vectorizer and TF-IDF Transformer all at once. Term Frequency – Inverse Document Frequency or TF-IDF, determines how significant the word is in the dataset and provides its respective weight. Data were split into two for training and testing the data. 70% of the data were allotted for training the model and 30% for testing it.

F. Training the Data

The models have been implemented using SciKit-Learn with different parameters in the Vectorizer to provide the best results. The kernel used for SVM was linear as it provided the best outcome.

G. Testing the Data

The models were tested based on their performance by measuring the accuracy, precision, recall, and F-measure.

IV. RESULTS

The data present the results and discussions on the sentiment analysis of Filipinos towards online classes using machine learning algorithms specifically Naïve Bayes and Support Vector Machine.

\[
TF-IDF = TF \times IDF
\]

\[
TF = \frac{\text{number of times the term appears in the document}}{\text{total number of terms in the document}}
\]

\[
IDF = \log \left( \frac{\text{number of documents in the corpus}}{\text{number of documents in the corpus that contain the term}} \right)
\]

**Term Frequency - Inverse Document Frequency** measures the accuracy of text vectorization which is the process of transforming text into numerical data. Term frequency refers to the frequency of a word in a document as compared to the total number of words in a text and inverse document frequency shows the proportion of documents in the overall corpus which have the specific term. Technical jargon and similar words which are less likely to be found in multiple documents and are therefore unique only to a few are given a higher ranking as compared to common words shared throughout all documents.

**Precision and Recall** are both necessary to measure the effectiveness of a machine learning algorithm. Precision evaluates the portion of correctly identified positives thus determining the significance of the results.
While Recall measures what portion of actual positives was identified correctly and thus determines how many significant results are found.

\[
\text{Accuracy} = \frac{\text{total correct data}}{\text{total data}}
\]  
\[(4)\]  
\[
\text{Precision} = \frac{\text{TP}}{\text{TP+FN}}
\]  
\[(5)\]  
\[
\text{Recall} = \frac{\text{TP}}{\text{TP+FN}}
\]  
\[(6)\]

**F-measure** measures the accuracy of the model in identifying positive instances. It differs from accuracy as it only considers positive instances whereas accuracy takes all factors into account (TP, TN, FP, FN). F-measure is then far more accurate when attempting to identify positive instances, especially with a potentially imbalanced data set.

<table>
<thead>
<tr>
<th>Table I. Multinomial Naïve Bayes Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>Negative</td>
</tr>
<tr>
<td>Neutral</td>
</tr>
<tr>
<td>Positive</td>
</tr>
<tr>
<td>Accuracy</td>
</tr>
<tr>
<td>Macro Avg</td>
</tr>
<tr>
<td>Weighted Avg</td>
</tr>
<tr>
<td>Accuracy</td>
</tr>
<tr>
<td>Precision Avg</td>
</tr>
<tr>
<td>Recall Avg</td>
</tr>
<tr>
<td>F-Measure Avg</td>
</tr>
</tbody>
</table>

Table 1 shows that the model’s accuracy of 0.6570 in correctly classifying the data along with a 0.6727 precision of determining the significance of each data, a 0.6570 accurate recall of significant data, with the F-Measure of 0.6494 which is the harmonic mean of precision and recall. These results came to be from the data that fell under the respective classification as shown by the “count”.

<table>
<thead>
<tr>
<th>Table II. Support Vector Machine Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>Negative</td>
</tr>
<tr>
<td>Neutral</td>
</tr>
<tr>
<td>Positive</td>
</tr>
<tr>
<td>Accuracy</td>
</tr>
<tr>
<td>Macro Avg</td>
</tr>
<tr>
<td>Weighted Avg</td>
</tr>
<tr>
<td>Accuracy</td>
</tr>
<tr>
<td>Precision Avg</td>
</tr>
<tr>
<td>Recall Avg</td>
</tr>
<tr>
<td>F-Measure Avg</td>
</tr>
</tbody>
</table>

Table 2 presents significantly better results than those presented in Table 1. It shows the model’s accuracy of 0.7539 in correctly classifying the data along with a 0.7572 precision of determining the significance of each data, a 0.7539 accurate recall of significant data, with the F-Measure of 0.7532 which is the harmonic mean of precision
and recall. These results came to be from the data that fell under the respective classification as shown by the “count” which also shows the similar split in data used for both models as it matches the previous table.

Table III. Multinomial Naïve Bayes Confusion Matrix

<table>
<thead>
<tr>
<th></th>
<th>Negative</th>
<th>Neutral</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>168</td>
<td>111</td>
<td>116</td>
</tr>
<tr>
<td>Neutral</td>
<td>27</td>
<td>420</td>
<td>143</td>
</tr>
<tr>
<td>Positive</td>
<td>21</td>
<td>127</td>
<td>456</td>
</tr>
</tbody>
</table>

Table IV. Support Vector Machine Confusion Matrix

<table>
<thead>
<tr>
<th></th>
<th>Negative</th>
<th>Neutral</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>266</td>
<td>76</td>
<td>53</td>
</tr>
<tr>
<td>Neutral</td>
<td>44</td>
<td>484</td>
<td>62</td>
</tr>
<tr>
<td>Positive</td>
<td>40</td>
<td>116</td>
<td>448</td>
</tr>
</tbody>
</table>

Table 3 and 4 present the confusion matrix used for the models used. The rows of the table represent the actual number of words tagged with that classification while the columns of the table represent the number of words predictively classified by the algorithm.

The results generated from tables above validated that the Filipinos Sentiment towards online classes are predominantly positive despite prevalent negative media during the pandemic.

V. CONCLUSION AND RECOMMENDATION

Online classes in the Philippines are a novelty not a norm. The pandemic brought forth the necessity of its abrupt implementation as a norm which then generated a confused reception. However, the general sentiment of Filipinos to online class is positive and clamor for its continuity as an optional mode of delivering instruction in education.

The following recommendations are provided based on the findings of the study:
1. That online classes delivery be improved and enhanced in response to the Filipinos positive sentiments.
2. For further study on the matter consider acquiring a larger data set to improve accuracy.
3. That a study be done comparing both Naïve Bayes and SVM with a different data set and more data points to further differentiate the two algorithms.

REFERENCES


