Analyzing Customer Satisfaction using Support Vector Machine and Naive Bayes Utilizing Filipino Text

JOSEPH B. CAMPIT
College of Arts, Sciences, and Technology,
Pangasinan State University - Bayambang Campus,
Zone VI, Bayambang, Pangasinan,
PHILIPPINES

Abstract: - The study aimed to compare the classification performance of Support Vector Machine (SVM) and Naive Bayes (NB) machine learning models for estimating customer satisfaction utilizing Filipino text. Specifically, it analyzed the characteristics of the customer satisfaction data. It also examined the impact of different model configurations, including n-gram, stop words, and stemming, on the classification performance of the two models. The research employed qualitative and quantitative methods, utilizing text analytics and sentiment analysis to extract and analyze valuable information from unstructured responses from a satisfaction survey of the University President's leadership performance conducted among PSU personnel and students. The dataset comprised 56,000 Filipino and English-word responses, manually annotated and randomly split into training and testing datasets. The study followed a general framework encompassing data pre-processing, modeling, and model comparison. To validate the classifiers' classification performance, a 10-fold cross-validation approach was employed. The findings revealed that most personnel and students expressed positive sentiment toward the University President's leadership performance. SVM outperformed the NB model across all different model configurations. With both stop word removal and stemming, the SVM trigram model achieved the highest classification performance for estimating customer satisfaction, using 75% of the data for training and 25% for testing. The proposed model holds the potential for estimating customer satisfaction using other unstructured customer satisfaction data utilizing Filipino text.

Key-Words: - Machine Learning, Text Analytics, Sentiment Analysis, Support Vector Machine, Naïve Bayes, Customer Satisfaction, Classification Performance

Received: December 22, 2022. Revised: April 9, 2023. Accepted: May 10, 2023. Published: June 6, 2023.

1 Introduction

Text analytics, powered by machine learning techniques, has gained significant attention in various fields such as economics, social sciences, bioinformatics, business, engineering, education, marketing, and logistics, [1], [2], [3], [4], [5], [6], [7], [8], [9]. It enables the extraction of valuable insights from vast amounts of unstructured textual data, including written survey responses, corporate documents, emails, customer messages, news articles, social media posts, and blogs, [10], [11], [12], [13], [14], [15], [16], [17]. With the exponential growth of unstructured data, automatic retrieval of meaningful knowledge from such data has become crucial for evidence-based decision-making, [18], [19].

One of the critical tasks in text analytics is sentiment or text classification using machine learning models, which automatically assign text documents to predefined categories based on their content and linguistic features, [20]. Machine learning models build classifiers that can effectively categorize new documents by learning the characteristics of pre-classified records from a training dataset, [21].

This study compares the performance of two popular machine learning models for text classification: Support Vector Machines (SVM) and Naïve Bayes (NB). Based on computational learning theory, SVM is a discriminative classification method that minimizes structural risk. It excels in pattern recognition, classification, and regression analysis, [22]. In contrast, NB is a simple probabilistic classifier that applies Bayes' Theorem with an intense independence assumption. It assumes that the presence or absence of a feature is unrelated to the presence or absence of other components, [23].

The main objective of this study is to compare the classification performance of SVM and NB in estimating customer satisfaction utilizing Filipino text. Specifically, it aims to analyze the characteristics of customer satisfaction data and evaluate the classification performance of SVM and NB by varying parameters such as n-gram, stop words, and stemming. Lastly, it will determine which machine learning model yields the best results in estimating customer satisfaction.

Various studies have compared the performance of NB and SVM models in different classification tasks. For instance, [24], compared NB and SVM models in text classification and found that SVM outperformed NB with large feature sets. Similarly, [25], observed that SVM achieved higher accuracy than NB in sentiment analysis. Likewise, [26], compared NB and SVM models in image classification and discovered that SVM outperformed NB in handling high-dimensional image data. Additionally, [27], compared NB and SVM models in spam email detection, and SVM outperformed NB regarding precision and recall.

Despite the performance differences between NB and SVM models, both have been extensively utilized in various machine learning applications. NB has found applications in text classification, [28], spam email detection, [29], and image recognition, [30], while SVM has been applied to face recognition, [31], speech recognition, [32], and gene expression analysis, [33].

Understanding customer sentiments and satisfaction levels is crucial for businesses and organizations, enabling them to tailor their offerings and strategies to meet customer needs and expectations. However, research gaps must be addressed, particularly in estimating customer satisfaction utilizing Filipino text. Although sentiment analysis has been widely studied in different languages and domains, research specifically focused on sentiment classification and customer satisfaction estimation using Filipino documents is limited.

This study contributes to the existing sentiment analysis and text classification knowledge by addressing this research gap. It provides practical implications for businesses and organizations aiming to enhance customer satisfaction and improve their understanding of customer sentiments.

2 Related Works

The study conducted by [34], focused on exploring the impact of stemming and n-gram techniques on sentiment classification for Arabic text. Additionally, it investigated the influence of feature selection on the performance of SVM, K-nearest Neighbor (KNN), and NB classifiers. The experimental findings demonstrated the highest performance when employing a hybrid representation incorporating tokens with character 3-grams. Furthermore, the results indicated that the application of feature selection techniques significantly enhanced the accuracy of all three classifiers in the task of opinion classification. Specifically, SVM exhibited superior performance compared to the other classifiers when utilizing all the features. However, when employing the SVM feature selection technique to select the most relevant features, both SVM and NB classifiers yielded the best outcomes.

In a related study by [35], the NB machine learning classifier was employed to classify Gujarati documents. The study focused on six predefined categories: sports, health, entertainment, business, astrology, and spiritual. A corpus of 280 records for each type was utilized for training and testing the categorizer. K-fold cross-validation was conducted with varying values of k (2, 4, 6, 8, and 10). The study's results indicated that the lowest error rate was achieved using 10-fold cross-validation, while the highest error rate was observed with 2-fold cross-validation. The classifier demonstrated a maximum accuracy of 88.96% when incorporating feature selection techniques, whereas the accuracy without feature selection was slightly lower at 75.74%.

In a study conducted by [36], the focus was on evaluating the performance of machine learning classifiers on Spanish Twitter data for opinion mining. The study aimed to identify the best configuration of parameters and features that would yield high precision in classifying opinions. Various factors were explored, including the size of n-grams, corpus size, number of sentiment classes, balanced versus unbalanced corpus, and the potential influence of different domains. The experimental tools used in the study were SVM, Decision Tree, and NB, which served as language classifiers. The findings indicated that using unigrams as features yielded the best results.

Employing fewer sentiment classes, specifically positive and negative, proved more effective in classifying opinions. Furthermore, the study revealed that a training set comprising at least 3,000 tweets was sufficient, as increasing the size did not significantly enhance precision. Balancing the corpus based on the proportional representation of all classes resulted in slightly worse performance. Finally, among the classifiers tested, SVM demonstrated the highest precision.

A study by [17], focused on estimating Filipino Internet Service Providers' (ISP) customer satisfaction. The study utilized web scraping techniques to extract customer comments and relevant information from popular blog sites featuring the services of major ISPs in the country. It resulted in a dataset comprising 14,000 sentences derived from 5,280 blog comments, which were stored in a database automatically. The researchers compared the classification performance of NB and SVM under different configurations involving stemming, stop word elimination, and n-gram tokenization. The researchers employed SVM and NB as machine learning classifiers and compared their precision, recall, F-measure, and accuracy performance. The classification experiments used 10-fold cross-validation to ensure robustness and reliability. The study's findings revealed that the SVM classifier outperformed the NB classifier in classification performance. The best results were obtained using SVM with trigram, Porter stemming, and stop word elimination. This configuration achieved a classification accuracy of 87%, indicating its effectiveness in accurately classifying customer sentiments related to ISP services.

3 Methodology

3.1 Research Design

The mixed-methods research design was used in this study. It is a type of research that combines quantitative and qualitative research methods to provide a complete understanding of research questions or problems. Text analytics and sentiment analysis were also employed to extract and analyze useful information from the responses of PSU personnel and students in a satisfaction survey of the leadership performance of the University President.

The mixed-methods research design in this study provided a more thorough understanding of the customer satisfaction data by combining qualitative and quantitative methods. Text analytics and sentiment analysis further enhanced the study by systematically analyzing and processing the large volume of unstructured data.

3.2 Respondents

The study targeted the personnel and students of Pangasinan State University, encompassing all nine (9) campuses, during the 2nd semester of the school year 2017-2018. A sample size of 8,000 respondents was randomly chosen from the university population to participate in the survey.

3.3 Dataset

The study utilized a dataset derived from a leadership performance survey comprising seven open-ended questions. The dataset contained 56,000 responses that were manually annotated and coded. The responses exhibited a wide range in length, varying from single-word answers to lengthy paragraphs, and were composed in either Filipino or English text.

The dataset was split into two subsets: the training and testing datasets. The training dataset was utilized for constructing the proposed models, allowing them to learn from the annotated responses. Conversely, the testing dataset evaluated the models' effectiveness and performance, providing an independent measure of their accuracy and predictive capabilities.

The training and testing dataset was selected through random sampling using the random function (Bernoulli) in MS Excel. An initial seed of 122,714 was used to generate the samples. For the division of the dataset, a split of 75% for training and 25% for testing was employed.

3.4 General Framework of the Study

Figure 1 illustrates the general framework used in the study, visually representing the flow of stages from data pre-processing to model comparison.

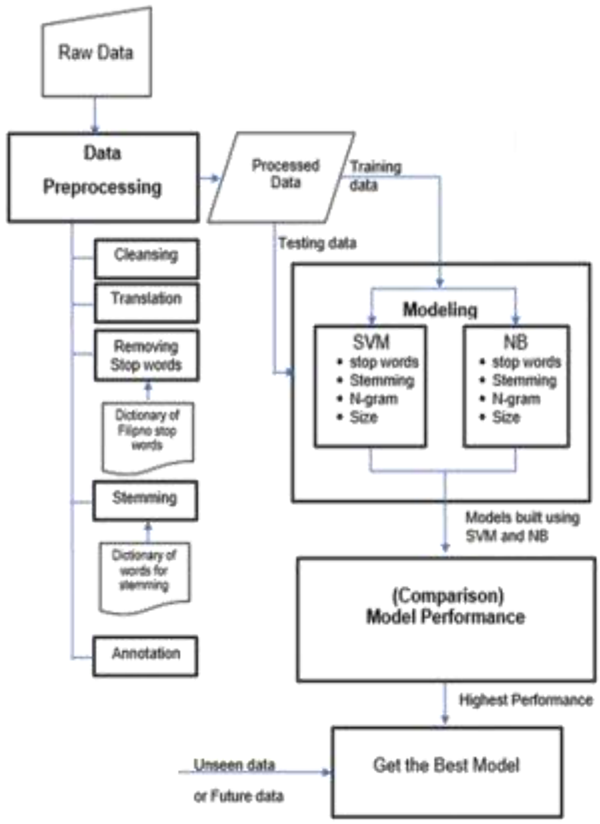

Fig. 1: General Framework Used in the Study

The study followed a general framework comprising three stages: data pre-processing, modeling, and model comparison. This framework provided a structured approach to systematically analyze and compare different machine learning models, enabling the identification of the most effective model for estimating customer satisfaction parameters.

3.3.1 Data Preprocessing

Data pre-processing is crucial in enhancing the relationship between words and document categories. Its main objective is to improve the quality of documents and reduce computational complexity, [37]. In this study, the written responses of the respondents were manually encoded in MS Excel.

During the data pre-processing stage, several steps were performed. Firstly, data cleansing was conducted to remove irrelevant sentence words and characters. Additionally, English words were translated into Filipino since not all responses were initially written in Filipino. The Google Translate tool was employed to convert English words into Filipino automatically.

Stop words, such as "ang," "mga," "sa," "ay," "at," etc., were also eliminated during the pre-processing of data. These familiar words in Filipino hold little value in identifying the category of the documents. A dictionary of Filipino stop words was created and used to remove these stop words from the record.

Stemming, another pre-processing technique, was applied to convert different word forms into their canonical form or root. This process ensures that words with the same canonical form are treated as one, such as "magaling," "pinakamagaling," and "napakagaling," which all share the canonical form "galing." The stemmer removes affixes (prefixes, infixes, and suffixes) and reduplicated parts, retrieving only the root word. Filipino affixes like "um," "ma," and "in" are considered during the stemming process. For example, the words "b(um)ilis," "(ma)ayos," and "s(in)abi" are stemmed from "bilis," "ayos," and "sabi," respectively. Similarly, in the words "aangat" and "tataas," the morphemes "a-" and "ta-" are reduplicated. After stemming, the affixes "a-" and "ta-" are removed, and "angat" and "taas" are retrieved. A dictionary of Filipino words or affixes for stemming was created to perform stemming on the words in the document.

The identification of sentiment polarity was manually annotated in the study. Seven groups, each consisting of three members, were assigned to annotate the polarity of the sentences. The raters underwent orientation and training to ensure consistency in the annotation process. The Fleiss' Kappa inter-rater reliability test was conducted to evaluate the consistency among the raters. The kappa values within each group showed acceptable inter-rater reliability (k>0.75) for the applied test. The mean inter-rater reliability for the sentence polarity raters was calculated to be k = 0.79. Based on the definition of the Fleiss' Kappa statistic, the inter-rater reliability accuracy is considered to have "Substantial agreement", [38].

3.3.2 Modeling

In this stage, the text within the document is transformed into a format suitable for training the algorithm. This training phase involved the experimentation and comparison of two machine learning algorithms: SVM and NB.

Both classifiers were trained and tested using the designated training and testing data. To explore the effectiveness of different model configurations, various techniques, such as n-grams, removal of stop words, and stemming were applied during the training and testing of the classifiers. The models' performances were then evaluated to develop a proposed machine learning model with superior classification capabilities.

To validate the classification performance of the two classifiers, the 10-fold cross-validation technique was employed. This approach involves dividing the dataset into ten equal-sized subsets, each serving as a testing set while the remaining nine subsets are used for training. The process is repeated ten times with a different subset as the testing set. It enables a comprehensive assessment of the classifiers' performance and helps determine their effectiveness in handling various data samples.

3.3.3 Comparison

The performance of the proposed models constructed using SVM and NB was compared to identify the model with the highest performance. Through rigorous evaluation and analysis of the results, the model with the highest performance was determined as the best or recommended machine learning model. This selection was made based on the model's ability to effectively classify and accurately predict the sentiment or category of the data.

This recommended model holds significant potential for estimating the sentiment polarity of forthcoming customer satisfaction data. By leveraging this model, valuable insights can be gained, enabling organizations to effectively analyze and gauge customer sentiment, making informed decisions to enhance overall customer satisfaction.

3.4 Statistical Treatment of Data

To ensure the reliability of the results, the free RapidMiner 8.2 Basic Edition was utilized, restricted to 1 logical processor and 10,000 data rows. This edition encompasses all stages of the text mining process, including data pre-processing, result visualization, and validation.

Various features were analyzed to gain insights into the characteristics of the customer satisfaction data. It involved examining the count of positive and negative sentences and identifying the occurrence of dominant words within the dataset.

The classification performance of the Support SVM and NB machine learning models was evaluated by parameterizing them with different combinations of n-gram, stop words, and stemming techniques. The computed values in the confusion matrix were utilized to determine key performance metrics, including accuracy, precision, recall, and F-measure. These metrics comprehensively assessed the models' effectiveness in accurately classifying sentiment.

A comprehensive performance comparison was conducted to identify the best model for estimating customer satisfaction.

4 Results and Discussion

4.1 Characteristics of the Data

Table 1 presents the distribution of the manually annotated positive, neutral, and negative responses.

Table 1. Distributions of Manually Annotated Positive, Neutral, and Negative Responses

Responses	**Total**
Positive	22,380
Negative	5,786
Neutral	27,834
Total	56,000

The dataset analysis revealed that out of the 56,000 sentences, 22,380 were labeled as positive and 5,786 as negative. Additionally, 27,834 neutral responses were excluded from the analysis. This distribution indicates a skew toward positive sentiments regarding the leadership performance of the University President. From these findings, there is a higher percentage of positive than negative sentiment expressed by both personnel and students toward the University President's leadership performance. These results suggest that the University President's leadership performance is generally well-received, with a majority expressing positive sentiment in their feedback.

Figure 2 presents the dominant words that describe the institutional leadership and performance of the University President.

Fig. 2: Dominant Words that Describe the Institutional Leadership and Performance

Figure 2 reveals that the President (“pangulo”) provides good and effective (“maayos,” “mahusay”, “maganda”) leadership in establishing and maintaining (“pagpapanatili”) excellent (“maganda, “maayos”) student services (“serbisyo”), upgrading (“marami,” “pagbabago,” “gusali”) physical plants and facilities, setting and institutionalizing responsive policies and procedures (“Mabuti,” “polisiya”) for the total improvement (“pagbabago”) of the University (“unibersidad”).

The respondents perceive that these organizational changes ("pagbabago," "unibersidad") in the University brought by the remarkable administration ("pamamalakad," "pamumuno") of the University President resulted in the enhancement to the effective delivery of the University's services ("maganda," "serbisyo") to various clienteles, especially to the students ("estudyante").

Figure 3 presents the respondents' perception concerning the external relations of the University President, particularly in sharing and contacting issues and concerns of the community and other public and private agencies.

Fig. 3: Dominant Words that Describe the External Relations

Figure 3 reveals that the university president is excellent ("mahusay") and orderly ("maganda," "maayos") in taking part ("pakikibahagi") in the activities and functions of the community ("komunidad") and other public and private agencies ("ahensiya"). He encourages involvement ("pakikibahagi") and linkage ("pakikipag-ugnayan") of various stakeholders in attending to issues and concerns ("suliranin") as well as the needs ("pangangailangan") of the University.

Figure 4 presents the dominant words that describe the budgetary fiscal management of the University President.

Fig. 4: Dominant Words that Describe the Budgetary Fiscal Management

It can be seen from Figure 4 that the respondents perceived that the University President has sound ("maayos") financial management ("pangangasiwa," pananalapi"), which also includes the efficient and effective use of resources ("nagagamit"). This can be explained through the effective and transparent leadership of the University President ("pamamalakad"). The President is also perceived as someone who endeavors to make sure that resources ("pera," "salapi") are maximized by prioritizing projects that addresses the needs of the university ("proyekto," "gusali," "paaralan," "building").

Figure 5 presents the dominant words that describe the personal qualities in dealing with issues of the University President.

Fig. 5: Dominant Words that Describe the Personal Qualities in Dealing with Issues

Figure 5 reveals that the University President ("Presidente") is brave ("matapang," "matatag") to face ("hinaharap") the pressing issues ("isyu") of the University and continue his purpose and pursue his goals ("layunin") of uplifting the current status of the University. The respondents recognize that the University, through the steering leadership and responsive decision-making ("desisyon") of the University President, can resolve ("maayos") its issues and problems.

Figure 6 presents the dominant words that describe the knowledge of the accreditation of the University President.

Fig. 6: Dominant Words that Describe the Knowledge in Accreditation

Figure 6 shows that the respondents perceive that the University President is well-versed ("malawak," "kaalaman," "magaling," "mahusay") in the accreditation requirements and ventures to these endeavors for the improvement of the University's processes, standards, academic offerings, and service delivery. Further, the University President's move for the University's accreditation from various agencies is observed to have a positive effect ("maganda," "maayos," "tumaas," "umangat," "lalo") on the institution. This results from the effective planning and self-assessment processes initiated by the University President and the involvement of all prevailing parties.

Figure 7 presents the dominant words that describe the University President's dealing with his fellow employees.

Fig. 7: Dominant Words that Describe the University President's Dealing with his Fellow Employees

Figure 7 reveals that the University President has a great rapport with his fellow University employees, as hinted by the dominant result words: "maayos," "marunong," "makisama," and "magaling." This might be the result of his excellent connection and dealings ("pakikisama," "pakikitungo") with other people, where his interpersonal skills are highlighted. As an administrator, the University President manages the University's most valuable resource - its people. This is achieved through the implementation of policies that are responsive and promote the well-being of University personnel.

Figure 8 presents the dominant words that describe what the President should focus on in his next term.

Fig. 8: Dominant Words that Describe the President Should Focus on in His Next Term

As can be gleaned from Figure 8, the respondents suggested that the University President give more attention ("pansin") to the University's pressing needs that affect the general welfare of the students and their studies. For example, most of the respondents desire the improvement of the University's physical facilities, such as creating new comfort rooms for their continuous upkeep.

4.2 Classification Performance of SVM and NB

Figure 9 displays the classification performance of SVM using different combinations of n-gram, stop words, and stemming.

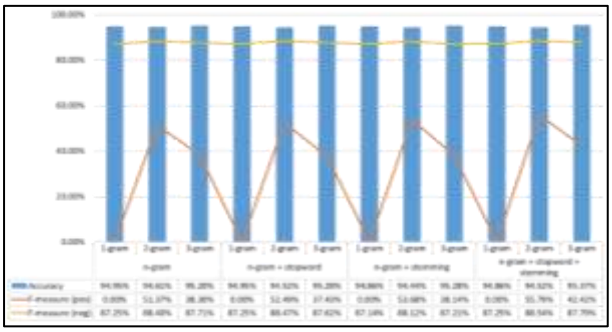

Fig. 9: Classification Performance of SVM

In terms of n-gram, the SVM trigram achieved the highest classification accuracy of 95.20%. Similarly, the SVM trigram obtained the highest classification accuracy of 95.28% when both stop words were removed and when stemming was applied. Furthermore, when combining n-gram, stop word removal, and stemming, the SVM trigram achieved the highest classification accuracy of 95.37%.

It is worth noting that the SVM trigram classification performance was consistently higher than the unigram and bigram in all combinations of model features. There was a slight increase in the classification performance of the SVM trigram from 95.20% to 95.37% when applying stop word removal (95.28%), stemming (95.28%), and the combination of stop word removal and stemming (95.37%). This suggests that the application of stop word removal, stemming, and their combination positively improves the performance of SVM trigram, but not on unigram and bigram.

The results presented in Figure 9 indicate that the highest classification performance of SVM was achieved using trigram with stop word removal and stemming, resulting in a classification accuracy of 95.37%. This model also achieved an F-measure of 96.31% for the positive and 66.67% for the negative classes. The results further demonstrate that the classification performance of SVM remained consistently high across all combinations of model configuration features, ranging from 94.44% to 95.37%. This observation is in line with the previous study by [34], which found that using n-gram is effective for classifying documents, specifically trigram with stop word removal and stemming.

Figure 10 illustrates the classification performance of NB.

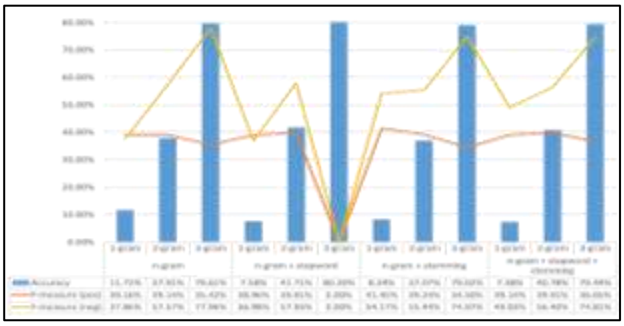

Fig. 10: Classification Performance of NB

The results demonstrate that the NB trigram achieved the highest classification accuracy of 79.44% when considering only the n-gram feature. When stop words were removed with n-gram, the NB trigram obtained the highest accuracy of 80.20%. Furthermore, when stemming was applied to the words and n-gram, the NB trigram attained a classification accuracy of 70.02%. Additionally, when combining n-gram, stop word removal, and stemming, NB trigram achieved the highest classification accuracy of 79.44%. This indicates that the NB trigram consistently outperformed other models regarding classification accuracy across all combinations.

It is worth noting that the classification performance of NB improved as the n-gram representation increased from unigram to trigram within each combination (n-gram, stop word removal, stemming, and the combination of stop word removal and stemming). The performance improvement ranged from 11.72% to 79.61% for n-gram alone, 7.58% to 80.20% for n-gram with stop word removal, 8.34% to 79.02% for n-gram with stemming, and 7.38% to 97.44% for the combination of stop word removal and stemming. These results suggest that utilizing bigram and trigram words as n-gram representations improves the performance of NB compared to unigram.

The analysis reveals that the best classification performance for NB was achieved using NB trigram with stop word removal, resulting in a classification accuracy of 80.20%. Moreover, the figure demonstrates that the classification performance of the NB trigram remained consistently high across all combinations of model configuration features, ranging from 79.02% to 80.20%. This finding coincides with the study of [17], which emphasized the usefulness of n-gram in detecting sentiment and specifically highlighted the effectiveness of trigrams.

4.3 Comparison of the Classification Performance of SVM and NB

A comparison of the classification performance of SVM and NB was made to determine the best model. Table 2 compares the classification performance of SVM and NB applying the different model configurations of n-gram, removing stop words, and stemming.

According to the results presented in Table 2, SVM consistently outperformed NB in terms of classification performance across all categories, including unigram, bigram, and trigram. The superiority of SVM over NB varied from 15.08% to 87.48%. The table further shows that trigram representation had the highest classification accuracy for SVM and NB in all configuration parameters, such as n-gram, stop words, and stemming.

Table 2. Comparison of the Classification Performance of SVM and NB

	SVM	NB	Difference
	Accuracy	**Accuracy**	
n-Gram			
Unigram	94.95%	11.72%	83.23%
Bigram	94.61%	37.91%	56.70%
Trigram	95.20%	79.61%	15.59%
n-Gram + stop words			
Unigram	94.95%	7.58%	87.37%
Bigram	94.52%	41.71%	52.81%
Trigram	95.28%	80.20%	15.08%
n-Gram + Stemming			
Unigram	94.86%	8.34%	86.52%
Bigram	94.44%	37.07%	57.37%
Trigram	95.28%	79.02%	16.26%
n-Gram + stop words + stemming			
Unigram	94.86%	7.38%	87.48%
Bigram	94.52%	40.78%	53.74%
Trigram	95.37%	79.44%	15.93%

The overall results demonstrate that the SVM model with trigram representation, combined with removing stop words and stemming, achieved the highest classification performance with a remarkable accuracy of 95.37%. The high classification accuracy attained by the SVM model with trigram representation indicates that this approach effectively captures the contextual information and dependencies between words within a sentence. The trigram representation considers three consecutive words as a single feature, allowing the model to capture more nuanced patterns and context in the text. This implies that analyzing a text at the trigram level or n-gram of length three is most effective as it provides richer information for sentiment classification and better performance in estimating customer satisfaction. This observation also agrees with the studies of [17], [34], which emphasized that n-gram works well on classifying documents and showed that trigram is effective.

Furthermore, the results suggest that removing stop words and stemming words in the text pre-processing phase improves the classification performance of the SVM model. By removing these stop words, the model focuses on more relevant and informative terms, enhancing its ability to discriminate between different sentiment classes. Similarly, stemming reduces words to their root form, collapsing variations of the same word. This generalization helps the model capture words' essence better and improve classification accuracy. This implies that removing stop words and stemming can significantly enhance the performance of sentiment classification models, particularly the SVM model.

5 Conclusions and Recommendations

5.1 Conclusions

Based on the findings of the study, the following conclusions can be drawn:

1. Most of the participants, including personnel and students, expressed positive sentiment toward the leadership performance of the University President.
2. The application of trigram with stop word removal and stemming techniques proved to be effective in accurately classifying the sentiments of customer satisfaction data for both the SVM and NB classifiers.
3. Among the classifiers evaluated, the SVM model utilizing trigram, stop word removal, and stemming demonstrated the highest performance in classifying the sentiments of customer satisfaction data using Filipino text. This model achieved the most accurate classification results and can be considered the recommended model for sentiment analysis in similar datasets.

5.2 Recommendations

Based on the findings and conclusions of the study, the following recommendations are provided:

1. Since most personnel and students expressed positive sentiment toward the University President's leadership performance, it is recommended to continue supporting and promoting the President's leadership style and initiatives. Regular surveys and feedback mechanisms should be implemented to monitor sentiment toward the President's leadership and identify any areas for improvement. Additionally, similar surveys can be conducted for other leadership roles within the University to gather sentiment data and aid in decision-making and leadership development.
2. The study found that implementing trigram with stop word removal and stemming effectively classified customer satisfaction data for both SVM and NB classifiers. Using these techniques in other sentiment analysis tasks utilizing Filipino text or other languages is recommended. Furthermore, exploring and experimenting with additional combinations of strategies can help improve sentiment analysis accuracy and identify the most effective methods for specific sentiment analysis tasks.
3. The SVM model utilizing trigram, stop word removal, and stemming demonstrated the best performance in classifying sentiments of customer satisfaction data using Filipino text. It is recommended to utilize this model in future sentiment analysis tasks involving Filipino customer satisfaction data. Additionally, considering the application of similar techniques in sentiment analysis tasks for other languages can help identify effective approaches.

References:

[1] Gandomi, A., & Haider, M., (2015). Beyond the Hype: Big Data Concepts, Methods, and Analytics, International Journal of Information Management, Vol. 35, No.2, 2015, pp. 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

[2] Zubin Jelveh, Bruce Kogut, and Suresh Naidu, (2014). Detecting Latent Ideology in Expert Text: Evidence From Academic Papers in Economics, In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014, pp. 1804–1809.

[3] Vapnik, (2000). The Nature of Statistical Learning Theory. Springer, New York, 2000.
[4] Yoshinobu Kano, William A. Baumgartner, Jr, Luke McCrohon, Sophia Ananiadou, K. Bretonnel Cohen, Lawrence Hunter, Jun'ichi Tsujii, U-compare: Share and compare text mining tools with UIMA. Bioinformatics, Vol. 25, No. 15, 2009, pp. 1997-1998. doi:10.1093/bioinformatics/btp289.
[5] Consoli, D., (2009). Analyzing customer opinions with text mining algorithms. AIP Conference Proceedings, Vol. 1148, 2009, pp. 857-860.
[6] Kostoff, R. N., Karpouzian, G., & Malpohl, G., (2005). Text mining the global abruptwing-stall literature. Journal of Aircraft, Vol. 42, 2005, pp. 661-664.
[7] Lin, Hsieh, & Chuang, (2009). Discovering genres of online discussion threads via text mining. Computers and Education, Vol. 52, 2009, pp. 481-495
[8] Gans, Joshua S. and Goldfarb, Avi and Lederman, Mara, (2017). Exit, Tweets and Loyalty, NBER Working Paper No. w23046, 2017.
[9] Jordan MI, Mitchell TM. (2015). Machine learning: trends, perspectives, and prospects. Science Vol. 349, No. 6245, pp. 255–260.
[10] Guran, Aysun, Selim Akyokuş and Nilgun Guler Bayazit, (2009). Turkish Text Categorization Using N-gram Word. International Symposium on Intelligent Systems and Applications, 2009.
[11] J. -S. Xu, (2009). TCBPLK: A New Method of Text Categorization, International Conference on Machine Learning and Cybernetics, Hong Kong, China, 2007, pp. 3889-3892, doi: 10.1109/ICMLC.2007.4370825
[12] Feng Li, (2011). Textual analysis of corporate disclosures: a survey of the literature. Journal of Accounting Literature Vol. 29, 2011, pp. 143-165.
[13] Sameer B. Srivastava, Amir Goldberg, V. Govind Manian, Christopher Potts, (2017). Enculturation Trajectories: Language, Cultural Adaptation, and Individual Outcomes in Organizations. Management Science, Vol. 64, No. 3, 2017, pp. 1-17.
[14] Struhl S., (2015). In the mood for sentiment. In Practical Text Analytics: Interpreting Text and Unstructured Data for Business Intelligence, Kogan Page Publishers: London, U.K., 2015.
[15] Jorge A. Balazs, Juan D. Velasquez, (2016). Opinion mining and information fusion: a survey. Information Fusion Vol. 27, 2016, pp. 95-110.
[16] Tetlock PC. (2007). Giving content to investor sentiment: the role of media in the stock market. The Journal of Finance, Vol. 62, No. 3, 2007, pp. 1139–1168.
[17] Frederick F, Patacsil, and Proceso L. Fernandez, (2015). Blog comments Sentence Level Sentiment Analysis for Estimating Filipino ISP Customer Satisfaction. International Conference Data Mining, Civil and Mechanical Engineering (ICDMCME ‘2015) February 1-2, 2015, Bali (Indonesia)
[18] Allahyari, Mehdi Seyedamin Pouriyeh, Mehdi Assefi, Saied Safaei, Elizabeth D.Trippe, Juan B.Gutierrez, and Krys Kochut. (2017). A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques. In Proceedings of KDD Bigdas, Halifax, Canada, 2017, pp. 1-13.
[19] Raghavan, P., Amer-Yahia, S., & Gravano, L., ((2004). Structure in Text: Extraction and Exploitation. Proceedings of the 7th International Workshop on the Web and Databases (WebDB), ACM SIGMOD/PODS, ACM Press, Vol. 67, 2004.
[20] Manning CD, Raghavan P, Schütze H., (2008). Introduction to Information Retrieval, Cambridge University Press: Cambridge, U.K., 2008.
[21] Sebastiani F. (2002). Machine learning in automated text categorization. ACM Computing Surveys (CSUR) Vol. 34, No. 1, pp. 1–47.
[22] Walaa Medhat, Ahmed Hassan, Hoda Korashy, (2014). Sentiment Analysis Algorithms and Applications: a Survey. Ain Shams University. Ain Shams Engineering Journal, Vol.5, No. 4, 2014, pp. 1093-1113
[23] Vafeiads, Thanasis. “A Comparison of Machine Learning Techniques for Customer Churn Prediction.” Simulation Modelling Practice and Theory Vol. 55, 201, pp. 1–9.
[24] Yang, Y., & Liu, X. (1999). A re-examination of text categorization methods. In Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval (pp. 42-49). ACM.
[25] Zhang, L., Zhu, C., & Li, P. (2014). A comparative study of naive Bayes and support vector machines for sentiment analysis. In 2014 International Conference on Mechatronics, Electronic, Industrial and

Control Engineering (MEIC 2014) (pp. 1705-1708). IEEE.

[26] Li, Y., Zhao, T., & Zhang, J. (2015). Image classification using naive Bayes and support vector machines. In 2015 International Conference on Advanced Cloud and Big Data (pp. 287-291). IEEE.

[27] Li, J., & Wang, X. (2017). Spam email detection based on naive Bayes and support vector machine. In 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (pp. 1410-1414). IEEE.

[28] Wang, S., & Manning, C. D. (2012). Baselines and bigrams: Simple, good sentiment and topic classification. In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers) (pp. 90-94). Association for Computational Linguistics.

[29] Androutsopoulos, I., Koutsias, J., Chandrinos, K. V., Paliouras, G., & Spyropoulos, C. D. (2000). An experimental comparison of naive Bayesian and keyword-based anti-spam filtering with personal email messages. In Proceedings of the 23rd annual international ACM SIGIR conference on Research and Development in Information Retrieval (pp. 160-167). ACM.

[30] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105). Curran Associates.

[31] Phillips, P. J., Moon, H., Rizvi, S. A., & Rauss, P. J. (1998). The FERET evaluation methodology for face-recognition algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(10), 1090-1104.

[32] Li, Y., Wang, J., & Li, X. (2008). Support vector machine-based Chinese speech emotion recognition. In 2008 International Conference on Wavelet Analysis and Pattern Recognition (pp. 589-592). IEEE.

[33] Brown, M. P., Grundy, W. N., Lin, D., Cristianini, N., Sugnet, C. W., Furey, T. S., ... & Haussler, D. (2000). Knowledge-based analysis of microarray gene expression data by using support vector machines. Proceedings of the National Academy of Sciences, 97(1), 262-267.

[34] Brahimi, Belgacem & Touahria, Mohamed & Tari, Abdelkamel, (2016). Data and Text mining Techniques for Classifying Arabic Tweet Polarity, Journal of Digital Information Management, Vol. 14, 2016, pp. 15-25

[35] Rakholia, (2017). Classification of Gujarati Documents using Naïve Bayes Classifier. Indian Journal of Science and Technology. Vol. 10(5), February 2017.

[36] Sidorov, Grigori, Sabino Miranda-Jiménez, Francisco Viveros-Jiménez, Alexander Gelbukh, Noé Castro-Sánchez, Francisco Velásquez, Ismael Díaz-Rangel, Sergio Suárez-Guerra, Alejandro Treviño, and Juan Gordon. (201). "Empirical Study of Machine Learning Based Approach for Opinion Mining in Tweets." Lecture Notes in Computer Science, 2013, pp. 1–14. doi:10.1007/978-3-642-37807-2_1.

[37] Kolchyna, Olga & Souza, Thársis & Treleaven, Philip & Aste, Tomaso, (2016). Twitter Sentiment Analysis: Lexicon Method, Machine Learning Method and Their Combination, Handbook of Sentiment Analysis in Finance. Mitra, G. and Yu, X. (Eds.), 2016.

[38] Gans, Joshua S. and Goldfarb, Avi and Lederman, Mara, (2017). Exit, tweets, and loyalty, NBER Working Paper 23046, National Bureau of Economic Research, Cambridge MA, 2017.

Contribution of Individual Authors to the Creation of a Scientific Article

The sole author of this scientific article independently conducted and prepared the entire work from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The sole author has no conflict of interest to declare.