International Journal of Computing Sciences Research (ISSN print: 2546-0552; ISSN online: 2546-115X)
Vol. 8, pp. 3235-3251
doi: 10.25147/ijcsr.2017.001.1.216
https://stepacademic.net

Short Paper*

An Enhancement of Support Vector Machine in Context of Sentiment Analysis Applied in Scraped Data from Tripadvisor Hotel Reviews

Lambert T. Dela Cruz
College of Information Systems and Technology Management, Pamantasan ng Lungsod ng Maynila, Philippines
ltdelacruz2020@plm.edu.ph
(corresponding author)

Marjorie Jasmine C. Racelis
College of Information Systems and Technology Management, Pamantasan ng Lungsod ng Maynila, Philippines
mjcracelis2020@plm.edu.ph
(corresponding author)

Vivien C. Agustin
College of Information Systems and Technology Management, Pamantasan ng Lungsod ng Maynila, Philippines
vaagustin@plm.edu.ph

Jamillah S. Guialil
College of Information Systems and Technology Management, Pamantasan ng Lungsod ng Maynila, Philippines
jsguialil@plm.edu.ph

Gabriel R. Hill
Senior Software Engineer, Accenture, Philippines
gabrhill97@gmail.com

Leisyl M. Mahusay
College of Information Systems and Technology Management, Pamantasan ng Lungsod ng Maynila, Philippines
lmocampo@plm.edu.ph

Jonathan C. Morano
College of Information Systems and Technology Management, Pamantasan ng Lungsod ng Maynila, Philippines
jcmorano@plm.edu.ph

Date received: May 7, 2024
Date received in revised form: July 29, 2024; July 30, 2024
Date accepted: July 30, 2024

Recommended citation:

Dela Cruz, L. T., Racelis, M. J. C., Agustin, V. C., Guialil, J. S., Hill, G. R., Mahusay, L. M., & Morano, J. C. (2024). An enhancement of support vector machine in the context of sentiment analysis applied in scraped data from Tripadvisor hotel reviews. *International Journal of Computing Sciences Research*, *8*, 3235-3251. https://doi.org/10.25147/ijcsr.2017.001.1.216

**Special Issue on International Research Conference on Computer Engineering and Technology Education (IRCCETE). Guest Associate Editors: Dr. Roben A. Juanatas (National University-Manila) and Dr. Nelson C. Rodelas (University of East).*

Abstract

Purpose – The purpose of this study is to improve the efficiency and accuracy of sentiment analysis in the context of hotel reviews, thereby contributing to the advancement of machine learning and natural language processing fields.

Method – The study employs an Enhanced SVM algorithm, incorporating SMO, Random Search, and SMOTE, to address issues of long training time, hyperparameter optimization, and imbalanced data.

Results – The Enhanced SVM outperforms the Traditional SVM, with a 13.48% increase in accuracy, a 100.42% reduction in training time, and improvements of 11.5% and 8.5% in Precision and F1-Score, respectively.

Conclusion – The study successfully enhances the SVM algorithm, providing a more effective tool for sentiment analysis in the context of hotel reviews, with significant improvements in performance metrics.

Recommendations – Future researchers should explore advanced optimization methods for hyperparameter tuning, use additional linguistic features like semantic analysis and context-aware embeddings, and incorporate sarcasm detection. Furthermore, consider deep learning models and ensemble approaches, combining SVM with other algorithms. Lastly, advocating for real-time sentiment analysis is suggested for immediate customer feedback insights.

Research Implications – The study offers valuable insights into the application of machine learning techniques in sentiment analysis, particularly in the tourism industry.

Practical Implications – The Enhanced SVM model can be used by platforms like TripAdvisor to provide more accurate sentiment analysis of hotel reviews, aiding tourists in their decision-making process.

Social Implications – Improved sentiment analysis can enhance the overall travel experience, leading to more satisfying and informed travel decisions.

Keywords – sentiment analysis, support vector machines, sequential minimal optimization, random search, synthetic minority over-sampling technique

---

INTRODUCTION

The abundance of online reviews poses a challenge for tourists in identifying quality attractions. Sentiment Analysis (SA), particularly using Support Vector Machines (SVM), can effectively analyze these reviews, aiding decision-making (Sari et al., 2023). SVM is a popular supervised machine learning algorithm that is used to determine the polarity of text (Abdulhammed & Karim, 2022).

The training process of SVM involves solving a quadratic programming (QP) problem, which requires solving many optimization subproblems. This can be time-consuming, especially when the dataset size is large. Sequential Minimal Optimization (SMO) addresses this issue by significantly reducing training time, ensuring both speed and stability. Their study result indicates that the Support Vector Machine SMO training time is 2.83 seconds, with a low standard deviation of 0.1, implying that the method is not only fast but also consistently stable and reliable, less prone to variations from random factors (Goel et al., 2021).

Hyperparameter tuning is particularly important in sentiment analysis and other natural language processing tasks, as it helps in accurately classifying and categorizing text-based data. The paper compares the performance of Naive Bayes and SVM classifiers. After hyperparameter tuning, SVM outperformed Naive Bayes with an accuracy of 85.65% compared to 68.70%, demonstrating the effectiveness of hyperparameter tuning in improving the performance of both models (Chong & Shah, 2022).

LITERATURE REVIEW

Local Literature

A study introduces a machine learning framework for diabetes classification, utilizing the Support Vector Machine (SVM) algorithm with various kernels. Evaluating SVM's pre and post-pre-processing performance, the polynomial kernel achieves the highest accuracy (83.77%), sensitivity (86.07%), and specificity (81.97%). SVM effectively identifies optimal hyperplanes for class segregation, proving beneficial for high-dimensional, limited training samples, and non-linear datasets. Comparative analysis shows the proposed SVM model's superior diagnostic capability in classifying diabetic disease compared to previous studies (Asaad, 2022).

Foreign Literature

SVM is a popular linear classifier for text, requiring meticulous feature engineering and TFIDF vectorization. Pre-trained Language Models (PLMs) like BERT and ELMO need minimal feature engineering but can be fine-tuned for NLP tasks. A study comparing four PLMs with SVM for text classification suggests that even fine-tuned PLMs don't outperform SVM significantly. Authors suggest using SVM with careful feature engineering for cost-effective and superior performance in text classification (Wahba et al., 2023).

SVM utilizes a kernel to classify data by converting it to a higher-dimensional space. Hyperparameters like C, kernel, and gamma are optimized through random search cross-validation. The SVM model achieved an F1-score of 0.80 without resampling, effectively classifying TikTok review sentiment with a good balance of precision and recall (Isnan et al., 2023).

Local Studies

SVM is used in OCR systems to differentiate Baybayin scripts from the Latin alphabet. SVM models excel in recognizing Baybayin and Latin scripts, characters, and diacritics. The SVM character classifier achieves perfect accuracy, precision, recall, and F1 score, all at 100% (Pino et al., 2022)

Foreign Studies

Support Vector Machines (SVM) is a classification algorithm aiming to find the optimal surface for separating positive and negative training samples. It generates a weight vector from training data to classify new points effectively (Korovkinas et al., 2019).

Support Vector Machine (SVM) is a classification algorithm in sentiment analysis, that outperformed K-NN in handling high-dimensional data by finding the best hyperplane for class separation. Results showed SVM achieved 88% accuracy with the Waterbom dataset (Sari et al., 2023).

SVM is frequently used for classification tasks in text mining, including sentiment analysis in e-commerce reviews. It can be optimized using techniques such as grid search for parameter selection and unigram feature extraction, which have been shown to improve classification accuracy in experiments using customer review datasets (Sulistiana & Muslim, 2020).

Sequential Minimal Optimization (SMO) efficiently trains SVMs by solving the Quadratic Programming problem in smaller subproblems, optimizing two variables iteratively. It reduces computational complexity and training time compared to optimizing the entire problem at once, making it a preferred choice for SVM tuning (Torres-Barrán et al., 2021).

A random search in SVM optimizes hyperparameters by randomly selecting combinations from a defined search space. Table 9 outlines SVM hyperparameters like kernel type, C value, and gamma value. Table 11 shows that preprocessing methods led to the highest F1 score of 83.362%, indicating improved sentiment prediction by the SVM model (Rahmadayana & Sibaroni, 2021).

The Synthetic Minority Over-sampling Technique (SMOTE) addresses imbalanced class data by generating synthetic samples for the minority class. In SVM, it enhances accuracy, precision, and sensitivity by balancing class distribution. Results show notable improvements, such as a 3% accuracy increase, 8% precision enhancement, and 25% sensitivity improvement in SVM performance with SMOTE on imbalanced data (Anggrawan et al., 2023).

METHODOLOGY

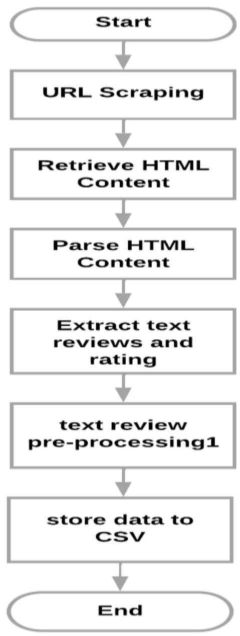
The researchers have first analyzed the existing problems of Support Vector Machines in Sentiment Analysis. The researchers have found three problems including intensive resource demands, hyperparameter optimization, and data imbalance. The researchers set objectives to enhance the SVM's efficiency, optimize the hyperplane for maximal-margin classification, and incorporate SMOTE for better data balance.

The existing problem is Resource-intensive in solving the Quadratic Programming problem making training time much longer and with a Complexity of O(n^3), Accuracy can be further improved if Hyperparameter and Data is Optimized. The researchers explored ways to reduce the time complexity and come up with a framework that aims to solve the existing problems. The new framework includes the use of Synthetic Minority Over-sampling TEchnique (SMOTE), Random Search, and Sequential Minimal Optimization.

Data Gathering and Pre-processing using BeautifulSoup

The data gathering process from TripAdvisor hotel reviews is extracted through BeautifulSoup web scraping. Initially, URLs are extracted to locate web pages containing hotel reviews. Subsequently, HTML content is retrieved via HTTP requests and parsed using BeautifulSoup to extract pertinent information such as text reviews and ratings. The extracted text reviews undergo pre-processing, including tasks such as HTML tag removal, tokenization, lowercasing, and stopword elimination, to enhance data quality.

Finally, the processed data, consisting of reviews and associated ratings, is stored in CSV format for further analysis. This structured approach ensures efficient and accurate acquisition of hotel review data from TripAdvisor, facilitating valuable insights and evaluations relevant to the tourism industry.



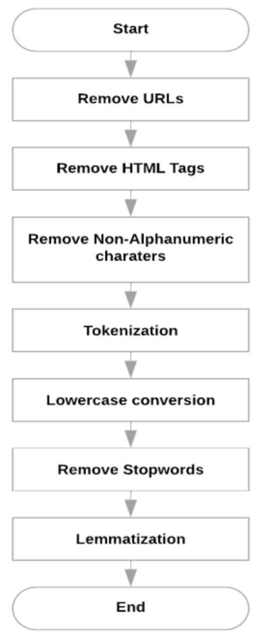


Figure 1. Flow of Data Gathering *Figure 2.* Flow of Data pre-processing

Data pre-processing refines raw text data for analysis in several ways. URLs are removed from the text to remove extraneous information. Only text is extracted once HTML tags are removed. To focus on text, alphanumeric characters like numerals and punctuation marks are filtered out. Tokenization breaks text into words for processing. To standardize representation, all text is transformed to lowercase. Eliminating meaningless words like "the" and "and" simplifies analysis. Finally, lemmatization reduces words to dictionary form for consistency and interpretability. This systematic approach cleans and standardizes text data for analysis and modeling.

Synthetic Minority Over-Sampling Technique (SMOTE)

According to Demidova and Klyueva (2017), the SMOTE algorithm is said to be improving the classification quality of Support Vector Machine classifiers for imbalanced data. Table 1 shows how SMOTE can affect the accuracy of a model when data imbalances are pre-processed. However, the study of Flores et al. (2018), says that SMOTE must be pre-processed well before feeding it to the SVM Classifier for it to yield a good result. In applying SMOTE, 10 Folds validation provides better findings compared to the 70:30 split.

Table 1. A Comparative Analysis between a Traditional SVM vs SVM + SMOTE

Algorithm	Parameter	Dataset A	Dataset B
SVM	Total Instances	2418	1021
	Correctly Classified Instances	1999	858
	Percentage	82.67%	84.04%
SVM + SMOTE	Total Instances	2494	1051
	Correctly Classified Instances	2067	890
	Percentage	82.88%	84.68

Hyperparameter Optimization

According to Chong and Shah (2022), it is proven that choosing the best parameters can highly impact the accuracy of the model. Tailoring these parameters, like the C and kernel, can significantly improve the model's ability to generalize from training data to unseen data, leading to more accurate predictions. This optimization process also helps in avoiding overfitting, where the model performs well on training data but poorly on new, unseen data.

Table 2. A Comparative Analysis between a Traditional SVM vs SVM + SMOTE

Parameters	C	Kernel	F1-Score
Default	1.0	Rbf	83.90%
Optimized	7.0	Linear	88.04

Sequential Minimal Optimization

According to the study of Flores et al. (2018), SMO breaks large QP problems into a series of smallest QP problems which can be solved analytically in such a short period. And can be faster than the chunking algorithms. This plays a pivotal role in reducing training time, particularly when dealing with large datasets.

Table 3. A Comparative Analysis between a Traditional SVM vs SMO+SVM

Algorithm	Samples	Training Time(s)
SVM	20K	124.04s
	30K	216.08s
	50k	796.96
SVM+SMO	20K	81s
	30K	98.87s
	50K	197.02s

The performance is evaluated based on the training time. As the number of samples increases from 20k to 50k, the training time also increases. However, the accuracy decreases slightly when the number of samples is increased to 50k. This suggests that while the SMO + SVM algorithm can handle larger datasets, there might be a trade-off in terms of accuracy.

Testing

To measure the performance of the model, the researchers used various kinds of benchmarking metrics that include Precision, Recall, and F1-Score.

$$\text{Precision} = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Positives\ (FP)} \qquad \textit{Equation 1}$$

Precision quantifies the degree of correctness in the positive predictions generated by the model. The calculation involves determining the proportion of accurate positive forecasts to the overall number of positive predictions, which include both accurate and inaccurate positive predictions.

$$\text{Recall} = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Negative\ (FP)} \qquad \textit{Equation 2}$$

Recall evaluates the model's capacity to accurately detect all true positive instances. The calculation involves dividing the number of true positives by the sum of true positives and false negatives.

$$\text{F1-Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \qquad \textit{Equation 3}$$

The F1-Score is a mathematical average that combines Precision and Recall, achieving a balanced measurement between these two criteria. It is especially beneficial in cases where there is an imbalanced distribution of classes or where both false positives and false negatives hold significant importance.

Conceptual Framework

Figure 3 presents the conceptual framework of the Enhanced Support Vector Machine (SVM) as it is applied in the field of sentiment analysis. This framework illustrates the process of how the Enhanced SVM, an improved version of the traditional SVM, is utilized to analyze and interpret subjective information.

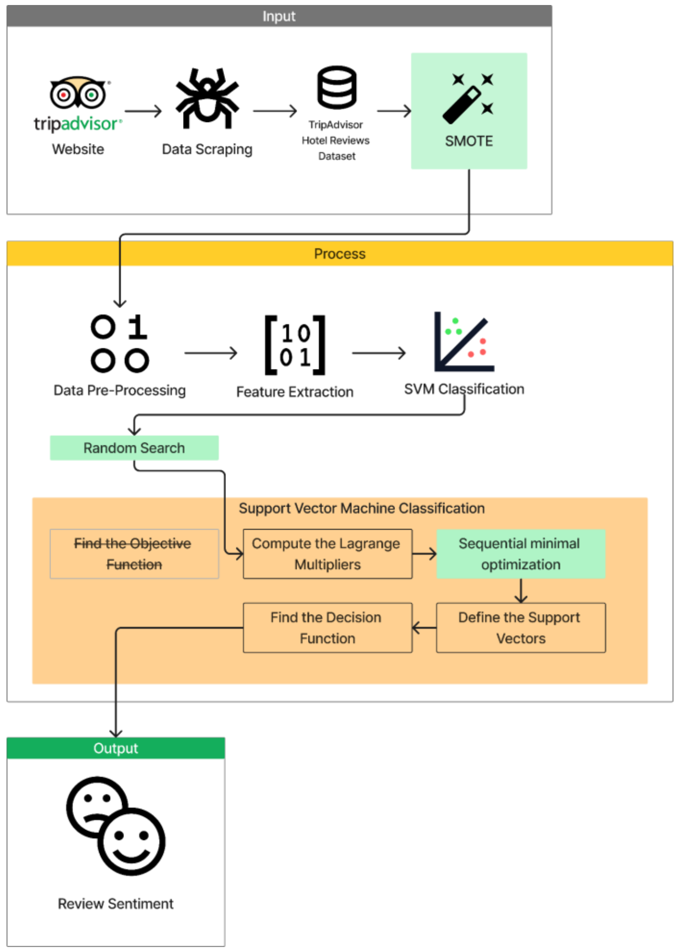


Figure 3. Conceptual Framework of Enhanced SVM in the context of Sentiment Analysis

The process starts with Data Scraping from the TripAdvisor website. To make the dataset balanced for all the classifications, the Synthetic Minority Oversampling Technique (SMOTE) will be used to balance the scraped data. After the data balancing, it will undergo data preprocessing where stop-words, filtering, and cleaning will be done. After data preprocessing is done, Feature extraction will take place where each text will turn into a numerical format that the machine understands. Afterwards, it will undergo SVM Classification and objective functions will be determined using Random Search to find the optimal hyperparameters of the given data. To speed up the training process, Sequential Minimal Optimization will be used.

RESULTS

Table 4 presents a comparative analysis of the traditional Support Vector Machine (SVM) and the enhanced SVM using Sequential Minimal Optimization (SMO) with Random Search and SMOTE.

Table 4. Comparative Analysis of SVM and SMO-RandomSearch-SMOTE

Parameter	SVM	SMO-RandomSearch-SMOTE	Difference
Accuracy	74.47%	87.95%	+13.48(16.6%)
Training Time(s)	379.02s	125.63s	-253.4s(100.42%)
Precision	73%	84.5%	+11.5 (14.60%)
Recall	79%	79%	0
F1-Score	72.5%	81%	+8.5(11.0749%)

The enhanced SVM outperforms the traditional SVM in terms of accuracy, precision, and F1-score, with improvements of 13.48% (16.6%), 11.5% (14.6032%), and 8.5% (11.0749%), respectively. The recall rate remains the same for both models at 79%. Notably, the training time for the enhanced SVM is significantly reduced by 253.4 seconds (100.42%), making it not only more accurate but also more efficient.

The findings of this study demonstrate significant improvements in SVM performance using SMO, Random Search, and SMOTE, aligning with and extending insights from the literature. Traditional SVM models, as highlighted by Wahba et al. (2023) and Isnan et al. (2023), rely heavily on meticulous feature engineering and hyperparameter optimization to achieve high performance. Our study shows that incorporating SMO reduces computational complexity and training time by breaking down the Quadratic Programming problem into smaller subproblems, a benefit highlighted by Torres-Barrán et al. (2021). Additionally, Random Search optimizes hyperparameters by selecting effective combinations from a defined search space, resulting in notable enhancements in accuracy, precision, and F1-score, as supported by Rahmadayana and Sibaroni (2021).

Moreover, the use of SMOTE to address the class imbalance by generating synthetic samples for the minority class significantly improves model performance, particularly in terms of accuracy and precision (Anggrawan et al., 2023). Comparative studies, such as those by Korovkinas et al. (2019) and Sari et al. (2023), illustrate SVM's superior performance in classification tasks, and our study further validates these findings by demonstrating that the enhanced SVM outperforms the traditional SVM in key performance metrics. This underscores the practical benefits of incorporating advanced optimization techniques and resampling methods, highlighting the enhanced SVM model's superior accuracy, precision, F1 score, and reduced training time, which are crucial for efficient and effective classification tasks in various applications.

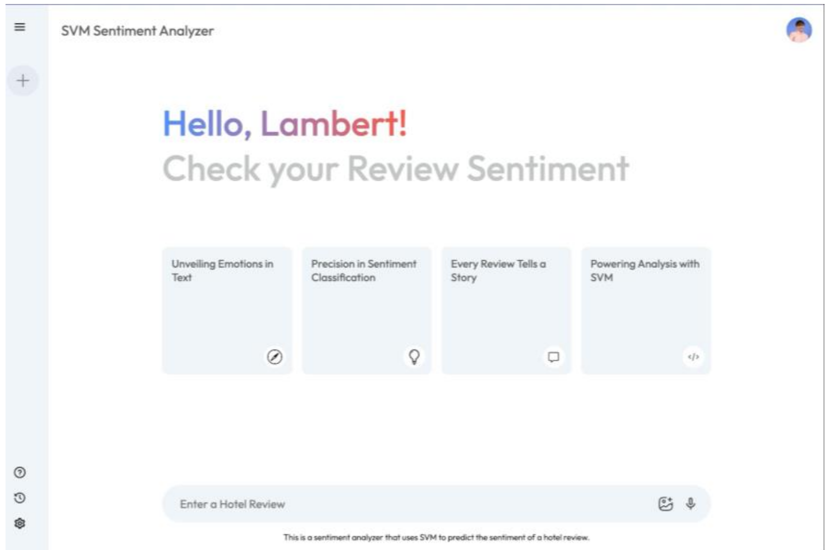


Figure 4. SVM Sentiment Analyzer Homepage

Figure 4 presents the user interface of the website SVM Sentiment Analyzer. The website provides a guide for users through the process of inputting text, analyzing sentiment, and viewing results that enhance decision-making processes. This website is currently using the Enhanced SVM Model.

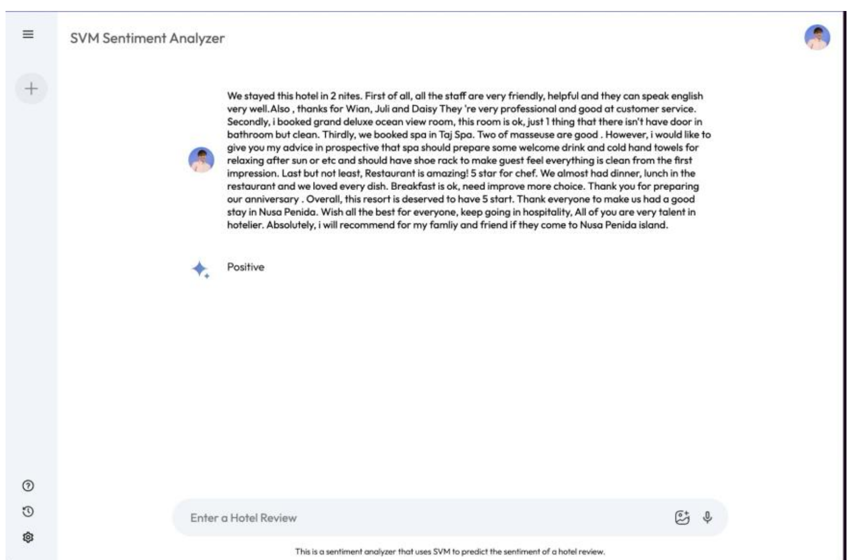


Figure 5. Positive Review

Figure 5 presents positive results, based on the hotel review from Tripadvisor. It reads the input review following the processing of the conceptual framework.

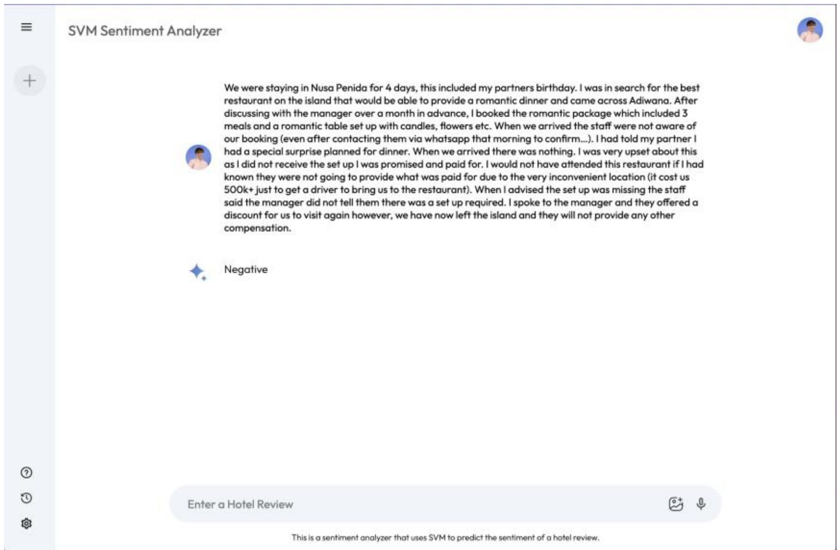


Figure 6. Negative Review

Figure 6 presents negative results, based on the hotel review from TripAdvisor. It resulted in a negative review as the predicted model detected that the choice of words is on the negative side of support vectors.

DISCUSSION

The results of this study demonstrate significant improvements in the performance of the Enhanced Support Vector Machine (SVM) model over the traditional SVM in the context of sentiment analysis of hotel reviews. The Enhanced SVM, which incorporates Sequential Minimal Optimization (SMO), Random Search, and Synthetic Minority Over-sampling Technique (SMOTE), shows marked advancements in accuracy, precision, and F1-score, while also significantly reducing training time.

The Enhanced SVM achieved a 13.48% increase in accuracy, an 11.5% improvement in precision, and an 8.5% enhancement in the F1 score compared to the traditional SVM. These improvements highlight the effectiveness of the combined techniques in addressing the challenges of long training times, hyperparameter optimization, and data imbalance.

The reduction in training time by 100.42% is particularly noteworthy. This efficiency gain is attributed to the use of SMO, which breaks down the quadratic programming problem into smaller, more manageable subproblems, thereby speeding up the training process without compromising stability.

The use of Random Search for hyperparameter tuning proved to be effective in enhancing the model's performance. By exploring a wide range of parameter combinations, the model was able to achieve optimal settings that significantly improved its predictive capabilities.

The application of SMOTE to balance the dataset further contributed to the model's improved performance. By generating synthetic samples for the minority class, SMOTE ensured a more balanced training set, which is crucial for accurate sentiment analysis.

In conclusion, the Enhanced SVM model presents a robust and efficient tool for sentiment analysis, addressing key challenges and offering significant improvements over traditional methods. Future research and practical applications can build on these findings to further advance the field of sentiment analysis in the tourism industry and beyond.

CONCLUSIONS AND RECOMMENDATIONS

The comprehensive analysis shows that Enhanced SVM outperforms Traditional SVM. It exhibits faster training times and consistent accuracy across different sample sizes. Enhanced SVM with Random Search achieves higher accuracy in hyperparameter optimization. Additionally, when combined with SMOTE, it handles imbalanced data better, maintaining high accuracy even with larger datasets. In conclusion, Enhanced SVM is the preferred choice for sentiment analysis on Tripadvisor hotel reviews, addressing issues of training times, hyperparameters, and data imbalances more effectively.

Researchers recommend exploring advanced optimization approaches beyond Grid Search and Random Search for efficient hyperparameter fine-tuning. Future studies can investigate additional linguistic features like semantic analysis and context-aware embeddings to enhance sentiment classification accuracy. They propose considering an ensemble approach, combining SVM with other ML models, to leverage their strengths. Lastly, advocating for the development of real-time sentiment analysis capabilities is suggested to provide immediate insights into customer feedback.

IMPLICATIONS

The study highlights the potential of advanced machine learning methods, such as the Enhanced SVM model, for improving sentiment analysis, notably in the tourism sector. The findings from this research improve computational linguistics and predictive analytics research by allowing for accurate sentiment analysis from massive hotel reviews. The improved accuracy and efficiency of the model can benefit websites such as TripAdvisor by streamlining the review analysis process, allowing tourists to make more informed accommodation choices based on reliable sentiment analysis. Furthermore, these advancements in sentiment analysis technology can enhance the travel experience by enabling tourists to make decisions based on a more nuanced understanding of hotel reviews, leading to greater satisfaction and a more personalized travel experience.

ACKNOWLEDGEMENT

The researcher extends profound gratitude to Prof. Vivien A. Agustin for her guidance and expertise, and to Prof. Jamillah S. Guialil, Prof. Gabriel Hill, Prof. Leisyl Mahusay, and Prof. Jonathan C. Morano for their invaluable support and contributions. Prof. Raymund M. Dioses for fostering an environment conducive to academic growth. Their collective efforts were instrumental in shaping this study, and their contributions are deeply appreciated.

FUNDING

The study did not receive funding from any institution.

DECLARATIONS

Conflict of Interest

The researcher declares no conflict of interest in this study.

Informed Consent

Not applicable. The researchers are utilizing public data available through the internet, hence the reason why it is not applicable.

Ethics Approval

Not applicable. The study involves the use of publicly available data. This study only collects and analyzes hotel reviews that have been publicly posted on the platform.

REFERENCES

Abdulhammed, O., & Karim, P. (2022). Sentiment analysis using SVM-based SSO intelligence algorithm. *Passer Journal of Basic and Applied Sciences, 4*(Special issue), 17-39. https://doi.org/ 10.24271/psr.2022.160801

Anggrawan, A., Hairani, H., & Satria, C. (2023). Improving SVM classification performance on unbalanced student graduation time data using SMOTE. *International Journal of Information and Education Technology, 13*(2), 289-295. https://doi.org/10.18178/ijiet.2023.13.2.1806

Asaad, R. R. (2022). Support vector machine classification learning algorithm for diabetes prediction. *International Research Journal of Science, Technology, Education, and Management, 2*(2), 26-34. https://doi.org/10.5281/zenodo.6975670

Chong, K., & Shah, N. (2022). Comparison of Naive Bayes and SVM classification in grid-search hyperparameter tuned and non-hyperparameter tuned healthcare stock market sentiment analysis. *International Journal of Advanced Computer Science and Applications, 13*(12), 90-94. https://doi.org/10.14569/IJACSA.2022.0131213

Demidova, L., & Klyueva, I. (2017, July). SVM classification: Optimization with the SMOTE algorithm for the class imbalance problem. In *Proceedings of the 2017 6th Mediterranean Conference on Embedded Computing (MECO)* (pp. 1-4). IEEE. https://doi.org/10.1109/MECO.2017.7977136

Flores, A. C., Icoy, R. I., Peña, C. F., & Gorro, K. D. (2018, July). An evaluation of SVM and Naive Bayes with SMOTE on sentiment analysis data set. In *Proceedings of the 2018 International Conference on Engineering, Applied Sciences, and Technology (ICEAST)* (pp. 1-4). IEEE. https://doi.org/10.1109/ICEAST.2018.8434401

Goel, K., Gu, A., Vaswani, A., & Susskind, J. (2021). *Efficiently modeling long sequences with structured state spaces*. arXiv. Retrieved from https://arxiv.org/pdf/2111.00396

Isnan, M., Elwirehardja, G. N., & Pardamean, B. (2023). Sentiment analysis for TikTok review using VADER sentiment and SVM model. *Procedia Computer Science, 227*, 168-175. https://doi.org/10.1016/j.procs.2023.10.514

Korovkinas, K., Danėnas, P., & Garšva, G. (2019). SVM and k-means hybrid method for textual data sentiment analysis. *Baltic Journal of Modern Computing, 7,* 47-60. https://doi.org/10.22364/bjmc.2019.7.1.04

Pino, R., Mendoza, R., & Sambayan, R. (2022). Block-level optical character recognition system for automatic transliterations of Baybayin texts using support vector machine. *Philippine Journal of Science, 151*, 303-315. https://doi.org/10.56899/151.01.23

Rahmadayana, F., & Sibaroni, Y. (2021). Sentiment analysis of work-from-home activity using SVM with randomized search optimization. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, *5*(5), 936-942. https://doi.org/10.29207/resti.v5i5.3457

Sari, A. W., Hermanto, T. I., & Defriani, M. (2023). Sentiment analysis of tourist reviews using k-nearest neighbors algorithm and support vector machine. *Sinkron: Jurnal Dan Penelitian Teknik Informatika, 8*(3), 1366-1378. https://doi.org/10.33395/sinkron.v8i3.12447

Sulistiana, & Much Aziz Muslim. (2020). Support Vector Machine (SVM) optimization using grid search and unigram to improve e-commerce review accuracy. *Journal of Soft Computing Exploration*, *1*(1), 8-15. https://doi.org/10.52465/joscex.v1i1.3

Torres-Barrán, A., Alaíz, C. M., & Dorronsoro, J. R. (2021). Faster SVM training via conjugate SMO. *Pattern Recognition, 111*, 107644. https://doi.org/10.1016/j.patcog.2020.107644

Wahba, Y., Madhavji, N., & Steinbacher, J. (2023). *A comparison of SVM against pre-trained language models (PLMs) for text classification tasks*. In G. Nicosia, V. Ojha, E. La Malfa, P. Pardalos, G. Di Fatta, G. Giuffrida, & R. Umeton (eds.), *Machine learning, optimization, and data science* (pp. 304-313). https://doi.org/10.1007/978-3-031-25891-6_23

Author's Biography

Lambert T. Dela Cruz, a Computer Science student at Pamantasan ng Lungsod ng Maynila, has already distinguished himself as a consistent academic achiever and former PLM Computer Science Society President. His leadership extended to his role as Cloud Development Lead at Google Developers Student Clubs PLM. Professionally, Lambert has made significant strides as a Google Cloud Engineer at Tenet Global Business Center, holding certifications as an AWS Certified Cloud Practitioner and AWS Certified Solutions Architect Associate. He also possesses a Hashicorp Terraform Associate certification, underscoring his expertise in infrastructure automation.

Marjorie Jasmine C. Racelis, a Computer Science student at Pamantasan ng Lungsod ng Maynila, possesses an excellent academic record. Her professional experience includes working as an IT Project Manager at XScribe Solutions, Inc., where she displayed exceptional project management and quality assurance skills. Marjorie has also completed internships at DFI Retail Group, Creative Interlace Technologies, and Dashlabs.ai, where she strengthened her technical and leadership skills. As a core member of the Google Developer Student Club PLM and a former CSS Vice President of the PLM Computer Science Society, she has demonstrated a strong dedication to community and organizational growth.

Vivien C. Agustin serves as a Thesis Adviser at the College of Information Systems and Technology Management, Pamantasan ng Lungsod ng Maynila, Philippines. With a deep understanding of information technology and systems, Vivien guides students through the rigorous process of thesis development, ensuring that their research is both innovative and impactful.

Jamillah S. Guialil is a Thesis Panelist at the College of Information Systems and Technology Management, Pamantasan ng Lungsod ng Maynila, Philippines. As an expert in information systems, Jamillah plays a critical role in evaluating and providing feedback on student theses, ensuring they meet academic and industry standards. Her insights help students refine their research and strengthen their projects.

Gabriel R. Hill serves as a Thesis Panelist at the College of Information Systems and Technology Management, Pamantasan ng Lungsod ng Maynila, Philippines, in addition to his role as a Senior Software Engineer at Accenture Philippines. Gabriel brings his extensive experience in software development and engineering to the evaluation of student theses, offering a practical perspective on the application of their research.

Leisyl M. Mahusay is the Thesis Coordinator at the College of Information Systems and Technology Management, Pamantasan ng Lungsod ng Maynila, Philippines. In her role, Leisyl oversees the thesis process, ensuring that all aspects, from proposal to defense, are conducted smoothly and efficiently.

Jonathan C. Morano is a Thesis Coordinator at the College of Information Systems and Technology Management, Pamantasan ng Lungsod ng Maynila, Philippines. In this role, Jonathan is responsible for managing the thesis process, ensuring that students receive the guidance and resources they need to complete their research successfully.