International Journal of Basic and Applied Science 12 (3) (2023) 100-109

Published by: IOCSCIENCE

International Journal of Basic and Applied Science

Journal homepage: www.ijobas.pelnus.ac.id

Comparative analysis of decision tree and support vector machine algorithm in sentiment classification for birds of paradise content

Yerik Afrianto Singgalen

Tourism Department, Faculty of Business Administration and Communication, Atma Jaya Catholic University of Indonesia, Indonesia

Article Info

Article history:

Received Nov 9, 2023
Revised Nov 30, 2023
Accepted Dec 28, 2023

Keywords:

Birding;
Classification;
Decision Tree;
Sentiment;
Support Vector Machine.

ABSTRACT

This research aims to analyze public sentiments towards National Geographic's content on the bird of paradise from the perspective of nature-based tourism. The method utilized is CRISP-DM, comprising stages of business understanding, data understanding, modeling, evaluation, and deployment. Focusing on sentiments expressed in response to National Geographic's Bird of Paradise content, this study seeks insights into how the public perceives and values nature-oriented tourism experiences. Comparing the results of DT and SVM algorithms with and without the SMOTE reveals noteworthy differences in classification performance. Without SMOTE, both DT and SVM exhibit relatively lower accuracy and AUC values compared to their counterparts with SMOTE. For DT, adding SMOTE substantially improves accuracy (from 92.44% to 95.20%) and AUC (from 0.517 to 0.956), indicating enhanced classification accuracy and model robustness. In addition, SVM demonstrates significant performance gains with SMOTE, achieving notably higher accuracy (from 92.12% to 98.63%) and AUC (from 0.617 to 0.999). The significantly higher values across various performance metrics for SVM underscore its effectiveness in handling imbalanced datasets and accurately classifying sentiment data. Therefore, researchers and practitioners may consider leveraging SVM for sentiment analysis tasks in similar contexts to achieve optimal classification results and enhance decision-making processes.

Corresponding Author:

Yerik Afrianto Singgalen,
Tourism Study Program, Faculty of Business Administration and Communication,
Atma Jaya Catholic University of Indonesia,
Jl. Jenderal Sudirman No. 51, Jakarta, 10220, Indonesia.
Email: yerik.afrianto@atmajaya.ac.id

1. INTRODUCTION

Birding, or birdwatching, has emerged as a prominent trend in tourism, signifying tourists' interest in observing avian species across diverse locations [1]. This activity entails deliberate bird observation, often necessitating visits to specific habitats known for their rich bird populations [2]. Consequently, birding qualifies as a niche tourism interest, attracting enthusiasts who seek to immerse themselves in the natural world through avian observation [3]. This trend underscores travelers' growing appreciation for wildlife and nature-based experiences, contributing to diversifying tourism offerings and fostering environmental awareness [4]. As individuals engage in birding, they contribute to the documentation and appreciation of local avifauna, thereby promoting awareness of the ecological significance of these species [5]. This heightened awareness fosters a sense of responsibility toward conservation efforts to protect habitats crucial for endemic birds [6]. Consequently, birding serves not only as a recreational pursuit but also as a means of advocating for the preservation of biodiversity and the sustainable management of natural resources.

The urgency of this research lies in the perspective utilized to analyze the interest in bird-watching activities based on the nature-based experience perspective [7]. By adopting this perspective, researchers delve into the intricate interplay between individuals' interactions with nature and their motivations for engaging in bird-watching [8]. This approach allows for a nuanced examination of the underlying factors driving the appeal of birding as a form of nature-based recreation [9]. Moreover, it facilitates a deeper understanding of birdwatching's psychological, emotional, and experiential dimensions, offering insights crucial for informing conservation efforts, enhancing visitor experiences, and promoting sustainable tourism practices.

The practical implication of this research underscores the public awareness regarding the sustainability of endemic bird habitats in respective regions, thereby influencing conservation policies or ecosystem protection based on socio-cultural contexts [10]. By shedding light on the significance of preserving habitats for endemic bird species, this study contributes to fostering a deeper understanding of the ecological value of these areas among local communities [11]. This heightened awareness catalyzes community-led conservation initiatives and influences policy decisions to safeguard critical bird habitats [12]. Ultimately, by integrating socio-cultural perspectives into conservation strategies, stakeholders effectively address the complex socio-economic dynamics impacting ecosystem preservation, thus ensuring the long-term viability of endemic bird populations.

The theoretical implication of this research contributes to the nature-based tourism perspective by specifically linking tourism context with bird-watching through sentiment classification analysis and comparing decision tree and support vector machine algorithms in classifying review data on content related to the Bird of Paradise published by National Geographic [13]. By employing advanced machine learning techniques to analyze sentiments expressed in reviews, this study offers valuable insights into the perceptions and preferences of tourists engaging in bird-watching activities [14]. Furthermore, comparing classification algorithms enhances our understanding of the effectiveness and applicability of different computational approaches in handling diverse datasets within the context of nature-based tourism research [15]. Ultimately, this research enriches theoretical frameworks in tourism studies by integrating computational methodologies to explore the complexities of tourist behavior and preferences in nature-oriented activities such as bird-watching.

The limitation of this research lies in the method employed to understand public responses regarding the sustainability of bird-watching tourism through the CRISP-DM method and the comparison of DT and SVM algorithm performance. While the CRISP-DM method offers a systematic framework for data mining, its application in gauging public sentiments may overlook nuanced socio-cultural factors influencing perceptions of sustainability in bird-watching tourism. Additionally, while comparing decision tree (DT) and support vector machine (SVM) algorithms provides valuable insights into classification accuracy, other machine learning techniques could further enhance the robustness of sentiment analysis [16]–[22]. Thus, while this research provides a foundational understanding, future studies should consider integrating data-mining approaches and exploring alternative computational methodologies to comprehensively capture the complexity of public perceptions toward sustainable tourism practices in bird-watching.

The opportunity for further research development in this area lies in expanding the scope of analysis to incorporate a broader range of socio-cultural and environmental variables influencing public perceptions and behaviors towards sustainable bird-watching tourism [23]. Researchers gain deeper insights into the underlying motivations and attitudes shaping tourist engagement with bird-watching activities by integrating qualitative methodologies such as interviews or focus groups [24]. Additionally, exploring alternative machine learning techniques and incorporating real-time data sources could enhance the accuracy and relevance of sentiment analysis in assessing tourist sentiments [25]. Ultimately, by addressing these aspects, future research endeavors have the potential to offer a more comprehensive understanding of sustainable tourism practices in the context of bird-watching, thereby informing effective conservation strategies and policy interventions.

The contribution of this study to knowledge is significant in its integration of advanced computational methodologies with the exploration of public perceptions towards sustainable bird-watching tourism. By applying the CRISP-DM method and comparing the performance of decision tree and support vector machine algorithms in sentiment analysis, this research advances our understanding of the complex interplay between tourism, environmental conservation, and socio-cultural factors [26]. Furthermore, the study's focus on bird-watching as a nature-based tourism activity enriches existing literature by offering insights into the specific motivations and preferences of tourists engaging in avian observation. Consequently, this research contributes to theoretical frameworks in tourism studies and provides practical implications for enhancing sustainable tourism practices and biodiversity conservation efforts.

2. RESEARCH METHOD

The research methodology employed in this study is the CRISP-DM method, encompassing the stages of business understanding, data understanding, modeling, evaluation, and deployment [27]. CRISP-DM, widely recognized in data mining and analytics research, provides a systematic framework for guiding the research process, ensuring methodological rigor and clarity in each analysis stage [28]. By following this structured approach, researchers effectively navigate the complexities of data collection, analysis, and interpretation, thus enhancing the reliability and validity of research findings [29]. Consequently, the utilization of CRISP-DM underscores the methodological robustness of this study and enhances confidence in the accuracy and relevance of its results.

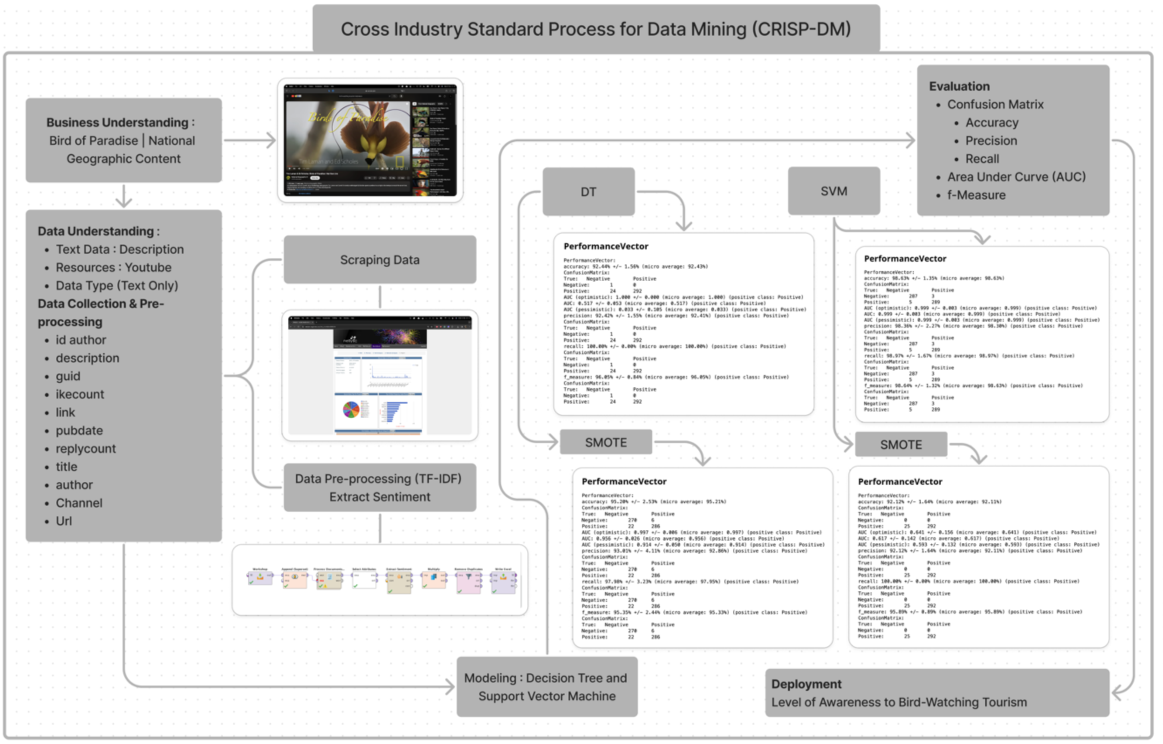

Figure 1. Implementation of CRISP-DM

In implementing CRISP-DM, the data division allocates 70% for testing and 30% for training from the total collected review data. Data pre-processing involves tokenization, transforming cases, and removing stopwords using the English dictionary. This approach ensures a systematic and standardized dataset treatment, allowing reliable and consistent analysis outcomes. By employing these methodologies, researchers effectively prepare the data for further analysis and model development, enhancing the robustness and validity of research findings. In the modeling stage, the decision tree (DT) algorithm is employed for classifying sentiment classes into negative and positive. This choice of algorithm allows for the creation of a hierarchical structure of decision nodes based on features extracted from the dataset, facilitating the classification of sentiments expressed in the reviews. This selection underscores the suitability of DT for handling classification tasks in sentiment analysis, contributing to the comprehensive understanding of public sentiments towards the subject matter.

$$[H(T) = -\sum_{i=1}^{c} p(i|t) \log_2 p\,(i|t)] \quad (1)$$

Where:
$(H(T))$is the total entropy of decision tree (T).
$(p(i|t))$ is the probability that a tuple generated by node (t) belongs to class label (i).
(c) is the number of possible classes.

This is the formula for calculating the entropy of a decision tree, which measures the "uncertainty" or "confusion" present in the dataset, measured in information units (bits) [28]. Lower entropy indicates better performance of the decision tree in classifying data. Following the implementation of the decision tree (DT) algorithm, a performance comparison is conducted with the support vector machine (SVM) algorithm. This comparative analysis aims to evaluate the effectiveness of SVM in sentiment classification tasks, particularly in distinguishing between positive and negative sentiments expressed in the collected reviews. By examining the performance metrics of both algorithms, researchers ascertain each approach's relative strengths and weaknesses, thereby informing the selection of the most suitable algorithm for sentiment analysis in similar contexts. Such comparative assessments are instrumental in advancing the understanding of algorithmic capabilities and optimizing classification accuracy in sentiment analysis research endeavors. The equation of a support vector machine (SVM) is shown as the equation below:

$$[f(x) = \text{sign}(\sum_{i=1}^{N} \alpha_i y_i K(x_i, x) + b)] \quad (2)$$

Where:
$(f(x))$ is the decision function.
(x_i) are the training samples.
(y_i) are the class labels.
$(K(x_i, x))$ is the kernel function that measures the similarity between the input samples.
(α_i) are the coefficients obtained during training.
(b) is the bias term.

This equation represents the decision boundary of the SVM, separating the classes in the feature space. Support Vector Machine (SVM) offers several advantages in classification tasks. One notable advantage is its ability to handle high-dimensional data efficiently, making it suitable for applications with complex feature spaces [30]. Additionally, SVM is robust against overfitting, thanks to its margin maximization objective, which promotes generalization to unseen data. Moreover, SVM effectively handles non-linear data using kernel functions, allowing for flexible decision boundaries [31]. These characteristics render SVM a powerful tool for various machine learning tasks, particularly sentiment analysis, where accurate classification and generalization are paramount. Thus, the advantages of SVM make it a popular choice in both research and practical applications for achieving optimal classification performance.

3. RESULTS AND DISCUSSIONS

Nature-based tourism, particularly bird-watching activities, plays a crucial role in maintaining ecosystem sustainability through institutional approaches; thus, positive public sentiments pressure stakeholders to establish regulations appropriate for bird habitat protection contexts. Bird-watching tourism fosters environmental awareness and appreciation for avian biodiversity, encouraging stakeholders to prioritize conservation efforts and implement effective management practices. Consequently, public support for nature-based tourism activities contributes to formulating and enforcing policies conducive to bird habitat preservation, ultimately ensuring the long-term viability of avian populations and their ecosystems.

Based on the data collection results using data mining techniques, insights into posts over time, top ten users, and most frequently used words were gleaned from respondents engaging with National Geographic's bird of paradise content. Through data mining, patterns emerge regarding the frequency and timing of posts, the most active users contributing to discussions, and the prevalent themes or topics conveyed through the respondents' language. This comprehensive analysis enables a deeper understanding of user engagement dynamics and content preferences within the context of bird of paradise discussions on National Geographic's platform. Thus, data mining is a valuable tool for uncovering meaningful insights that inform content creation strategies and user interaction approaches in online platforms dedicated to nature-related topics.

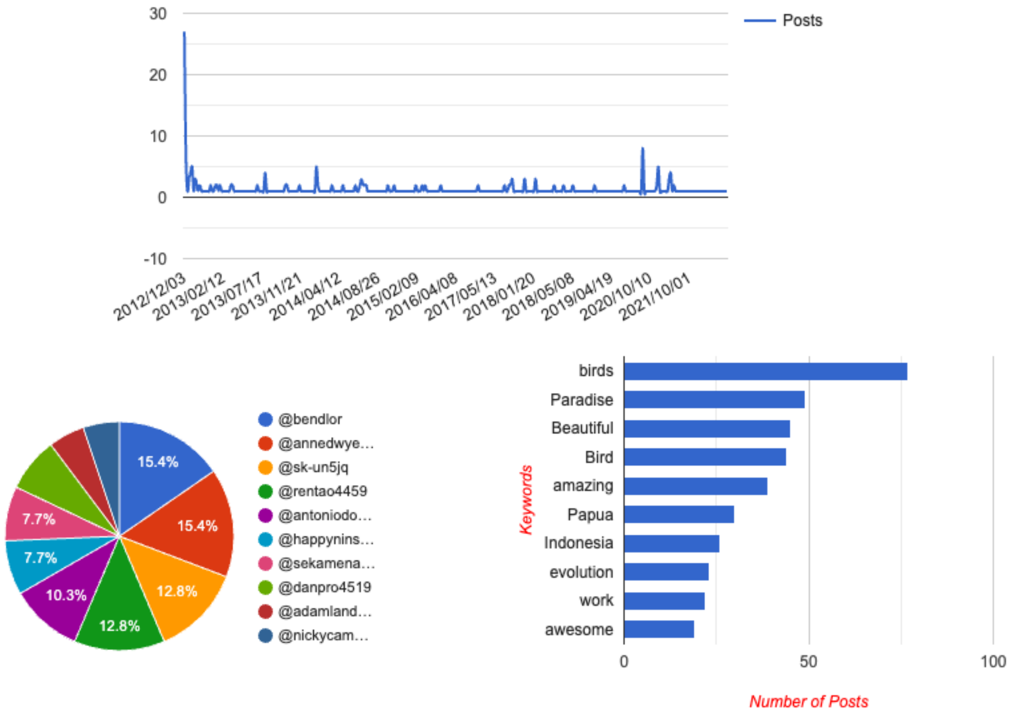

Figure 2. Post Over Time, Top Ten User, and Most Frequently Used Words

Based on the data extracted from Post Over Time, Top Ten Users, and Most Frequently Used Words, it is evident that certain words frequently appear in the reviews, such as "paradise," "beautiful," "bird," "amazing," "Papua," "Indonesia," "evolution," "work," and "awesome." This observation suggests that Indonesia, particularly Papua, emerges as an appealing bird-watching destination. The recurrence of these words underscores the captivating allure of Papua's avian biodiversity and natural landscapes and the positive sentiments expressed by users towards encounters with birds of paradise. Such findings highlight the significance of Papua as a focal point for nature-based tourism, particularly for bird-watching enthusiasts seeking awe-inspiring experiences amidst diverse ecosystems and endemic avifauna.

Based on the sentiment extraction results obtained from the RapidMiner application, it is discernible that certain words recur most frequently in the dataset, including "birds" (76 occurrences), "paradise" (48 occurrences), "bird" (44 occurrences), "beautiful" (44 occurrences), "amazing" (38 occurrences), "Papua" (30 occurrences), "evolution" (23 occurrences), "Indonesia" (26 occurrences), "work" (22 occurrences), and "Guinea" (16 occurrences). This frequency distribution underscores the prominence of words associated with bird-watching experiences and the natural beauty of Papua, Indonesia. The prevalence of positive descriptors such as "beautiful" and "amazing" suggests a favorable sentiment towards bird of paradise encounters and highlights the captivating allure of Papua's avian biodiversity. Additionally, mentioning "evolution" and "Guinea" indicates an appreciation for the region's unique evolutionary history and geographic context. Such insights from sentiment analysis contribute to a deeper understanding of public perceptions and preferences regarding bird-watching tourism in Papua, Indonesia.

Figure 3. Frequently used Words in Rapidminer

The most frequently occurring keywords in the dataset are "birds," "paradise," "bird," "beautiful," and "amazing," suggesting that users commonly associate these terms with their bird-watching experiences. These positive descriptors indicate a favorable sentiment towards encounters with birds of paradise and the natural beauty of Papua, Indonesia. The presence of location-specific terms such as "Papua," "Indonesia," and "Guinea" highlights the geographic context of the discussions. This suggests that users specifically refer to the Papua region of Indonesia, renowned for its rich avian biodiversity and bird-watching opportunities. The appearance of the term "evolution" suggests an appreciation for the evolutionary history of bird species in the region. This indicates that users are interested in observing birds and understanding their evolutionary significance and ecological role within the ecosystem. The predominance of positive descriptors like "beautiful" and "amazing" implies that users have positive experiences and sentiments towards bird-watching activities in Papua, Indonesia. This positive sentiment is crucial for promoting the region as a desirable destination for nature-based tourism, particularly bird-watching. The analysis highlights the positive sentiment and geographic significance of bird-watching experiences in Papua, Indonesia. The insights from sentiment analysis inform tourism stakeholders and policymakers in developing strategies to promote and sustain nature-based tourism in the region. Understanding public perceptions and preferences enhances visitor experiences and conservation efforts in Papua's unique and ecologically significant ecosystems.

In the modeling stage, the decision tree (DT) algorithm is utilized for classifying review data into negative and positive classes. However, there exist differences in the performance of DT when employing the Synthetic Minority Over-sampling Technique (SMOTE) and when not using SMOTE. The application of SMOTE aims to address class imbalance by generating synthetic samples of the minority class, thereby enhancing the model's ability to classify negative and positive sentiments accurately. Consequently, the performance of DT is expected to improve significantly when SMOTE is employed, leading to better classification results and increased model robustness.

Table 1. DT Algorithm Performance

DT without SMOTE	**DT with SMOTE**
PerformanceVector:	PerformanceVector:
Accuracy: 92.44% +/- 1.56% (micro average: 92.43%)	Accuracy: 95.20% +/- 2.53% (micro average: 95.21%)
ConfusionMatrix:	ConfusionMatrix:
True: Negative Positive	True: Negative Positive
Negative: 1 0	Negative: 270 6
Positive: 24 292	Positive: 22 286
AUC (optimistic): 1.000 +/- 0.000 (micro average: 1.000) (positive class: Positive)	AUC (optimistic): 0.997 +/- 0.006 (micro average: 0.997) (positive class: Positive)
AUC: 0.517 +/- 0.053 (micro average: 0.517) (positive class: Positive)	AUC: 0.956 +/- 0.026 (micro average: 0.956) (positive class: Positive)
AUC (pessimistic): 0.033 +/- 0.105 (micro average: 0.033) (positive class: Positive)	AUC (pessimistic): 0.914 +/- 0.050 (micro average: 0.914) (positive class: Positive)
precision: 92.42% +/- 1.55% (micro average: 92.41%) (positive class: Positive)	precision: 93.01% +/- 4.11% (micro average: 92.86%) (positive class: Positive)
ConfusionMatrix:	ConfusionMatrix:
True: Negative Positive	True: Negative Positive
Negative: 1 0	Negative: 270 6
Positive: 24 292	Positive: 22 286
recall: 100.00% +/- 0.00% (micro average: 100.00%) (positive class: Positive)	recall: 97.98% +/- 3.23% (micro average: 97.95%) (positive class: Positive)
ConfusionMatrix:	ConfusionMatrix:
True: Negative Positive	True: Negative Positive
Negative: 1 0	Negative: 270 6
Positive: 24 292	Positive: 22 286
f_measure: 96.05% +/- 0.84% (micro average: 96.05%) (positive class: Positive)	f_measure: 95.35% +/- 2.44% (micro average: 95.33%) (positive class: Positive)
ConfusionMatrix:	ConfusionMatrix:
True: Negative Positive	True: Negative Positive
Negative: 1 0	Negative: 270 6
Positive: 24 292	Positive: 22 286

Notable differences arise in comparing the performance of decision tree (DT) algorithms with and without the Synthetic Minority Over-sampling Technique (SMOTE). Without SMOTE, DT yields an accuracy of 92.44%, precision of 92.42%, recall of 100.00%, f_measure of 96.05%, and an AUC of 0.517. Conversely, when utilizing SMOTE, DT achieves higher performance metrics, with an accuracy of 95.20%, precision of 93.01%, recall of 97.98%, f_measure of 95.35%, and an AUC of 0.956. These results highlight the efficacy of SMOTE in improving the classification performance of DT, particularly in addressing class imbalance issues and enhancing model robustness. Therefore, incorporating SMOTE into the DT algorithm demonstrates promise for achieving more accurate sentiment classification results in similar contexts.

Table 2. SVM Algorithm Performance

SVM without SMOTE	**SVM with SMOTE**
PerformanceVector:	PerformanceVector:
Accuracy: 92.12% +/- 1.64% (micro average: 92.11%)	Accuracy: 98.63% +/- 1.35% (micro average: 98.63%)
ConfusionMatrix:	ConfusionMatrix:
True: Negative Positive	True: Negative Positive

Negative: 0 0	Negative: 287 3
Positive: 25 292	Positive: 5 289
AUC (optimistic): 0.641 +/- 0.156 (micro average: 0.641) (positive class: Positive)	AUC (optimistic): 0.999 +/- 0.003 (micro average: 0.999) (positive class: Positive)
AUC: 0.617 +/- 0.142 (micro average: 0.617) (positive class: Positive)	AUC: 0.999 +/- 0.003 (micro average: 0.999) (positive class: Positive)
AUC (pessimistic): 0.593 +/- 0.132 (micro average: 0.593) (positive class: Positive)	AUC (pessimistic): 0.999 +/- 0.003 (micro average: 0.999) (positive class: Positive)
precision: 92.12% +/- 1.64% (micro average: 92.11%) (positive class: Positive)	precision: 98.36% +/- 2.27% (micro average: 98.30%) (positive class: Positive)
ConfusionMatrix:	ConfusionMatrix:
True: Negative Positive	True: Negative Positive
Negative: 0 0	Negative: 287 3
Positive: 25 292	Positive: 5 289
recall: 100.00% +/- 0.00% (micro average: 100.00%) (positive class: Positive)	recall: 98.97% +/- 1.67% (micro average: 98.97%) (positive class: Positive)
ConfusionMatrix:	ConfusionMatrix:
True: Negative Positive	True: Negative Positive
Negative: 0 0	Negative: 287 3
Positive: 25 292	Positive: 5 289
f_measure: 95.89% +/- 0.89% (micro average: 95.89%) (positive class: Positive)	f_measure: 98.64% +/- 1.32% (micro average: 98.63%) (positive class: Positive)
ConfusionMatrix:	ConfusionMatrix:
True: Negative Positive	True: Negative Positive
Negative: 0 0	Negative: 287 3
Positive: 25 292	Positive: 5 289

In evaluating the performance of support vector machine (SVM) algorithms with and without the Synthetic Minority Over-sampling Technique (SMOTE), discernible distinctions emerge. Without SMOTE, SVM achieves an accuracy of 92.12%, precision of 92.12%, recall of 100.00%, f_measure of 95.89%, and an AUC of 0.617. Conversely, with SMOTE, SVM demonstrates significantly improved performance metrics, attaining an accuracy of 98.63%, precision of 98.36%, recall of 98.97%, f_measure of 98.64%, and an AUC of 0.999. These findings underscore the efficacy of SMOTE in enhancing SVM's classification performance, particularly in mitigating class imbalance issues and improving model robustness. Thus, incorporating SMOTE into the SVM algorithm holds promise for achieving more accurate sentiment classification results in similar analytical contexts.

Comparing the results of DT and SVM algorithms with and without the SMOTE reveals noteworthy differences in classification performance. Without SMOTE, both DT and SVM exhibit relatively lower accuracy and AUC values compared to their counterparts with SMOTE. For DT, adding SMOTE substantially improves accuracy (from 92.44% to 95.20%) and AUC (from 0.517 to 0.956), indicating enhanced classification accuracy and model robustness. Similarly, SVM demonstrates significant performance gains with SMOTE, achieving notably higher accuracy (from 92.12% to 98.63%) and AUC (from 0.617 to 0.999). These findings underscore the effectiveness of SMOTE in addressing class imbalance issues and improving the classification performance of both DT and SVM algorithms. Thus, incorporating SMOTE into the classification process holds promise for optimizing sentiment analysis outcomes in similar analytical contexts.

4. CONCLUSION

In conclusion, this study aims to analyze public sentiments toward National Geographic's Bird of Paradise content within the framework of nature-based tourism. Employing the CRISP-DM methodology, involving stages of business understanding, data understanding, modeling, evaluation, and deployment, this research seeks insights into how the public perceives and values nature-oriented tourism experiences. Adopting CRISP-DM ensures a structured and systematic approach to data analysis, facilitating a comprehensive understanding of public sentiments and their implications for nature-based tourism practices. This research contributes significantly by providing theoretical insights and practical applications, particularly for enhancing visitor experiences and conservation efforts in nature-based destinations. Comparing the results of DT and SVM algorithms with and without the SMOTE reveals noteworthy differences in classification performance. Without SMOTE, both DT and SVM exhibit relatively lower accuracy and AUC values compared to their counterparts with SMOTE. For DT, adding SMOTE substantially improves accuracy (from 92.44% to 95.20%) and AUC (from 0.517 to 0.956), indicating enhanced classification accuracy and model robustness. Similarly, SVM demonstrates significant performance gains with SMOTE, achieving notably higher accuracy (from 92.12% to 98.63%) and AUC (from 0.617 to 0.999).

ACKNOWLEDGEMENTS

I sincerely thank the Tourism Department, the Faculty of Business Administration and Communication at Atma Jaya Catholic University of Indonesia, and the Research and Community Service Institute (LPPM).

REFERENCES

[1] N. Li, C. Yang, X. Qin, J. Sun, and J. Liu, "The impact of climate change on birder destination loyalty: examining changes in bird resources," *Curr. Issues Tour.*, vol. 25, no. 11, pp. 1798–1816, 2022, doi: 10.1080/13683500.2021.1983523.

[2] M. Kruger and A. Viljoen, "Bird(er)s of a feather? A typology of birders to South African national parks based on their behavioural involvement," *Ann. Leis. Res.*, vol. 26, no. 1, pp. 1–26, 2023, doi: 10.1080/11745398.2020.1813041.

[3] P. Tryjanowski, S. Murawiec, and C. Randler, "No such Thing as Bad Birding Weather, but Depends on Personal Experience," *Leis. Sci.*, vol. 0, no. 0, pp. 1–13, 2023, doi: 10.1080/01490400.2023.2167026.

[4] C. Randler, N. Staller, N. Kalb, and P. Tryjanowski, "Charismatic Species and Birdwatching: Advanced Birders Prefer Small, Shy, Dull, and Rare Species," *Anthrozoos*, vol. 36, no. 3, pp. 427–445, 2023, doi: 10.1080/08927936.2023.2182030.

[5] N. Groβmann and C. Randler, "Developing an instrument to assess the satisfaction and frustration of basic psychological needs during the leisure activity of birdwatching (Birding-BPNSF)," *J. Leis. Res.*, vol. 0, no. 0, pp. 1–23, 2023, doi: 10.1080/00222216.2023.2287003.

[6] R. Humphreys, "A 'Game' Bird? On Why Hunting is Not a Game and Thus Not a Sport," *Sport. Ethics Philos.*, vol. 17, no. 4, pp. 432–442, 2023, doi: 10.1080/17511321.2023.2189292.

[7] J. Spring, "Nature-based tourism and guided wildlife tours: designing wildlife tour experiences that optimise sustainable learning opportunities," *J. Ecotourism*, vol. 22, no. 1, pp. 187–207, 2023, doi: 10.1080/14724049.2022.2098963.

[8] E. Conti and I. Farsari, "Disconnection in nature-based tourism experiences: an actor-network theory approach," *Ann. Leis. Res.*, vol. 0, no. 0, pp. 1–18, 2022, doi: 10.1080/11745398.2022.2150665.

[9] D. Sumanapala, I. D. Wolf, and B. Weiler, "Enhancing Tour Guide Training for Delivering Nature-Based Tourism Experiences in a Developing Country," *J. Qual. Assur. Hosp. Tour.*, vol. 00, no. 00, pp. 1–25, 2023, doi: 10.1080/1528008X.2023.2253561.

[10] N. Richardson and A. Insch, "Enabling transformative experiences through nature-based tourism," *Tour. Recreat. Res.*, vol. 48, no. 2, pp. 311–318, 2023, doi: 10.1080/02508281.2021.1952396.

[11] E. Sthapit, P. Björk, and D. N. Coudounaris, "Memorable nature-based tourism experience, place attachment and tourists' environmentally responsible behaviour," *J. Ecotourism*, vol. 22, no. 4, pp. 542–565, 2023, doi: 10.1080/14724049.2022.2091581.

[12] L. Kou, X. Xiao, H. Xu, and J. Cheng, "Understanding tourist experiences of sounds at nature-based destinations: from a relational perspective," *Curr. Issues Tour.*, vol. 27, no. 4, pp. 600–618, 2023, doi: 10.1080/13683500.2023.2168522.

[13] A. Douglas, P. Mostert, and L. Slabbert, "Millennials as consumers of wildlife tourism experiences," *World Leis. J.*, vol. 64, no. 4, pp. 487–507, 2022, doi: 10.1080/16078055.2022.2097736.

[14] I. Falardeau, P. Marcotte, and L. Bourdeau, "How 'natural' is innovation in nature-based tourism?," *Loisir Soc.*, vol. 45, no. 1, pp. 134–149, 2022, doi: 10.1080/07053436.2022.2053326.

[15] C. Clark and G. P. Nyaupane, "Understanding Millennials' nature-based tourism experience through their perceptions of technology use and travel constraints," *J. Ecotourism*, vol. 22, no. 3, pp. 339–353, 2023, doi: 10.1080/14724049.2021.2023555.

[16] Y. A. Singgalen, "Analisis Sentimen Pengunjung Pulau Komodo dan Pulau Rinca di Website Tripadvisor Berbasis CRISP-DM," *J. Inf. Syst. Res.*, vol. 4, no. 2, pp. 614–625, 2023, doi: 10.47065/josh.v4i2.2999.

[17] Y. A. Singgalen, "Penerapan Metode CRISP-DM untuk Optimalisasi Strategi Pemasaran STP (Segmenting

, Targeting , Positioning) Layanan Akomodasi Hotel , Homestay , dan Resort," *J. Media Inform. Budidarma*, vol. 7, no. 4, pp. 1980–1993, 2023, doi: 10.30865/mib.v7i4.6896.

[18] Y. A. Singgalen, "Analisis Sentimen Top 10 Traveler Ranked Hotel di Kota Makassar Menggunakan Algoritma Decision Tree dan Support Vector Machine," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 1, pp. 323–332, 2023, doi: 10.30865/klik.v4i1.1153.

[19] Y. A. Singgalen, "Analisis Sentimen dan Sistem Pendukung Keputusan Menginap di Hotel Menggunakan Metode CRISP-DM dan SAW," *J. Inf. Syst. Res.*, vol. 4, no. 4, pp. 1343–1353, 2023, doi: 10.47065/josh.v4i4.3917.

[20] Y. A. Singgalen, "Penerapan Metode Additive Ratio Assessment (ARAS) dan Ranking of Centroid (ROC) dalam Pemilihan Layanan Akomodasi dan Local Cuisine," *J. Comput. Syst. Informatics*, vol. 5, no. 1, pp. 51–60, 2023, doi: 10.47065/josyc.v5i1.4569.

[21] Y. A. Singgalen, "Implementasi Metode CRISP-DM dalam Analisis Model Pendukung Keputusan Simple Additive Weighting dan Pengembangan Basis Data Riwayat Pembelian Layanan Akomodasi Hotel," *J. Sist. Komput. dan Inform.*, vol. 5, no. 2, pp. 308–317, 2023, doi: 10.30865/json.v5i2.7153.

[22] Y. A. Singgalen, "Penerapan CRISP-DM dalam Klasifikasi Sentimen dan Analisis Perilaku Pembelian Layanan Akomodasi Hotel Berbasis Algoritma Decision Tree (DT)," *J. Sist. Komput. dan Inform.*, vol. 5, no. 2, pp. 237–248, 2023, doi: 10.30865/json.v5i2.7081.

[23] J. Räikkönen, M. Grénman, H. Rouhiainen, A. Honkanen, and I. E. Sääksjärvi, "Conceptualizing nature-based science tourism: a case study of Seili Island, Finland," *J. Sustain. Tour.*, vol. 31, no. 5, pp. 1214–1232, 2023, doi: 10.1080/09669582.2021.1948553.

[24] A. Akhshik, H. Rezapouraghdam, A. Ozturen, and H. Ramkissoon, "Memorable tourism experiences and critical outcomes among nature-based visitors: a fuzzy-set qualitative comparative analysis approach," *Curr. Issues Tour.*, vol. 26, no. 18, pp. 2981–3003, 2023, doi: 10.1080/13683500.2022.2106196.

[25] Q. Jiang, C. S. Chan, S. Eichelberger, H. Ma, and B. Pikkemaat, "Sentiment analysis of online destination image of Hong Kong held by mainland Chinese tourists," *Curr. Issues Tour.*, vol. 24, no. 17, pp. 2501–2522, 2021, doi: 10.1080/13683500.2021.1874312.

[26] V. H. Luong, A. Manthiou, J. Kang, and C. Nguyen, "The building blocks of regenerative tourism and hospitality: a text-mining approach," *Curr. Issues Tour.*, 2023, doi: 10.1080/13683500.2023.2228974.

[27] L. G. R. Putra, Mayadi, and I. N. D. Setiaji, "Klasifikasi Jenis Client Menggunakan Algoritma Decision Tree Cart," *JSI J. Sist. Inf.*, vol. 14, no. 2, pp. 2842–2855, 2022, doi: 10.30812/jsi.v14i2.18826.

[28] J. A. Syahid and D. Mahdiana, "Perbandingan algoritma untuk klasifikasi analisis sentimen terhadap Genose pada media sosial Twitter," *semanTIK*, vol. 7, no. 1, pp. 9–16, 2021, doi: 10.5281/zenodo.5034916.

[29] M. R. Muttaqin, T. I. Hermanto, and M. A. Sunandar, "Penerapan K-Means Clustering Dan Cross-Industry Standard Process for Data Mining (CRISP-DM) Untuk Mengelompokan Penjualan Kue," *Komputasi J. Ilm. Ilmu Komput. dan Mat.*, vol. 19, no. 1, pp. 38–53, 2022, [Online]. Available: https://journal.unpak.ac.id/index.php/komputasi

[30] H. J. Christanto and Y. A. Singgalen, "Sentiment Analysis on Customer Perception towards Products and Services of Restaurant in Labuan Bajo," *J. Inf. Syst. Informatics*, vol. 4, no. 3, pp. 511–523, 2022, doi: 10.51519/journalisi.v4i3.276.

[31] A. Munir, E. P. Atika, and A. D. Indraswari, "Analisis Sentimen pada review hotel menggunakan metode pembobotan dan klasifikasi," *Jnanaloka*, vol. 3, no. 1, pp. 33–38, 2022, doi: 10.36802/jnanaloka.2022.v3-no1-33-38.