International Journal of Basic and Applied Science 12 (3) (2023) 121-130

Published by: IOCSCIENCE

International Journal of Basic and Applied Science

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

Sentiment classification of coral reef 101 content using decision tree algorithm through CRISP-DM

Yerik Afrianto Singgalen

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

Article Info

Article history:

Received Oct 2, 2023
Revised Nov 20, 2023
Accepted Dec 11, 2023

Keywords:

Classification;
Coral;
Decision Tree;
Reef;
Sentiment.

ABSTRACT

This research aims to classify public sentiment regarding the content of "Coral Reef 101," published by National Geographic. The methodology employed is the Cross-Industry Standard Process for Data Mining (CRISP-DM), encompassing stages such as business understanding, data understanding, modeling, evaluation, and deployment. The Decision Tree algorithm is utilized in conjunction with the SMOTE operator. This comprehensive approach enables the systematic analysis of public sentiment towards coral reef content, facilitating a deeper understanding of public perception and attitudes. The results of this study indicate that the DT algorithm with SMOTE demonstrates an accuracy of 87.51% +/- 4.28% (micro average: 87.50%), a precision of 80.35% +/- 5.10% (micro average: 80.00%) (positive class: Positive), recall of 100.00% +/- 0.00% (micro average: 100.00%) (positive class: Positive), f-measure of 89.02% +/- 3.22% (micro average: 88.89%) (positive class: Positive), and an AUC of 0.875 +/- 0.044 (micro average: 0.875) (positive class: Positive). These metrics demonstrate the effectiveness of the DT algorithm with SMOTE in accurately classifying public sentiment towards coral reef-related content, particularly in correctly identifying positive sentiment instances. The high accuracy, precision, recall, f-measure, and AUC values underscore the robustness and reliability of the model in sentiment analysis tasks.

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

Coral reefs play a crucial role in maintaining the sustainability of marine ecosystems [1]. As the "rainforests of the sea," coral reefs harbor immense biodiversity and provide habitat for many marine species [2]. Additionally, they serve as breeding grounds and nurseries for numerous fish and invertebrates, contributing significantly to the productivity and resilience of marine communities [3]. Furthermore, coral reefs act as natural barriers, protecting coastlines from erosion and buffering against the impacts of storms and tsunamis [4]. In conclusion, preserving coral reefs is paramount for the continued health and stability of marine ecosystems worldwide [5].

Protecting marine ecosystems represents a strategic step towards safeguarding marine environments through implementing appropriate policies that benefit the economy, society, and culture [6]. By enacting measures to preserve marine habitats and biodiversity, governments and stakeholders can mitigate the adverse impacts of human activities such as overfishing, pollution, and habitat destruction [7]. Furthermore, conserving marine ecosystems ensures the sustainability of vital ecological processes and fosters socio-economic development through ecotourism, sustainable fisheries, and the preservation of cultural heritage linked to marine environments [8]. In conclusion, protecting marine ecosystems is imperative for marine life's and human societies' long-term well-being, fostering harmony between environmental preservation and socio-economic prosperity [9].

The urgency of this research lies in the public response to the threats facing coral reefs stemming from various issues such as climate change and development [10]. This study aims to identify public sentiment towards the content of "Coral Reef 101," published by National Geographic. By understanding public perceptions and attitudes towards coral reefs, particularly in the context of prevalent threats, policymakers and conservationists can effectively formulate informed strategies to address these challenges [11]. Furthermore, raising awareness among the public about the importance of coral reef conservation is essential for fostering collective action towards mitigating the impacts of climate change and unsustainable development on marine ecosystems [12]. In conclusion, the findings of this research are crucial for informing decision-making processes aimed at preserving coral reefs and ensuring the long-term health and resilience of marine environments through sentiment analysis.

The practical implication of this research emphasizes the importance of policymakers' consideration of public sentiment regarding the sustainability of marine ecosystems through various digital media channels [13]. By incorporating insights from public sentiment analysis of coral reefs, policymakers can develop more targeted and effective strategies for conservation and management [14]. This approach facilitates the alignment of policy interventions with public perceptions and concerns, enhancing the likelihood of successful implementation and more outstanding public support for conservation initiatives [15]. Integrating public sentiment into decision-making processes regarding coral reef management is essential for promoting public engagement and fostering collaborative efforts toward preserving marine ecosystems [16].

The theoretical implication of this research underscores the significance of public awareness concerning the sustainability of coral reefs and broader ecological sustainability issues, mainly marine resources. By delving into public sentiment towards coral reefs through digital media campaigns, this study advances theoretical understanding of how awareness and perceptions of marine conservation issues are disseminated and influenced within the public sphere [17]. Moreover, by examining public awareness within the context of coral reef sustainability, this research extends theoretical discussions on environmental communication and advocacy, offering insights into practical strategies for fostering public engagement and support for marine conservation endeavors [18]. Overall, the theoretical implications of this study underscore the critical importance of public awareness in promoting ecological sustainability, particularly concerning marine resources and coral reef conservation.

The limitation of this research is primarily rooted in its methodology and contextual constraints, as the scope of discussion is confined to sentiment classification using the CRISP-DM method with the DT algorithm and SMOTE operator [19]. While this approach offers valuable insights into public sentiment towards coral reef content, it may overlook other pertinent factors influencing public perceptions and attitudes. Additionally, the reliance on digital media platforms for data collection may introduce biases and limitations inherent to online communication channels [20], [21]. Thus, future research endeavors should incorporate diverse methodologies and broader contextual factors to provide a more comprehensive understanding of public sentiment toward coral reef conservation.

The opportunity for further research development lies in comparing the utilization of algorithms within the CRISP-DM method concerning contextual issues of marine tourism in the Indonesia [22], [23]. By exploring the application of various algorithms in analyzing public sentiment towards marine tourism, researchers can gain deeper insights into the effectiveness of different methodologies in addressing the complex socio-economic and environmental challenges associated with marine tourism development in Indonesia [24], [25]. This comparative approach enhances the understanding of algorithmic performance and facilitates the identification of best practices and tailored solutions for sustainable marine tourism management. Ultimately, such research endeavors promise to inform policy interventions and industry practices to promote responsible and sustainable marine tourism in Indonesia.

2. RESEARCH METHOD

The method employed in this research is the CRISP-DM, comprising stages of business understanding, data understanding, modeling, evaluation, and deployment. This systematic approach provides a structured framework for conducting data mining projects, facilitating the progression from initial business objectives to the practical implementation of predictive models [26]. By adhering to the CRISP-DM methodology, researchers can effectively navigate the complexities of data analysis and model development, ensuring the rigor and reliability of their findings [27]. Moreover, the iterative nature of CRISP-DM allows for continuous refinement and optimization of models, enhancing their accuracy and applicability in real-world scenarios [28]. Overall, using CRISP-DM methodology in this research underscores the commitment to methodological rigor and systematic approach in addressing complex research questions.

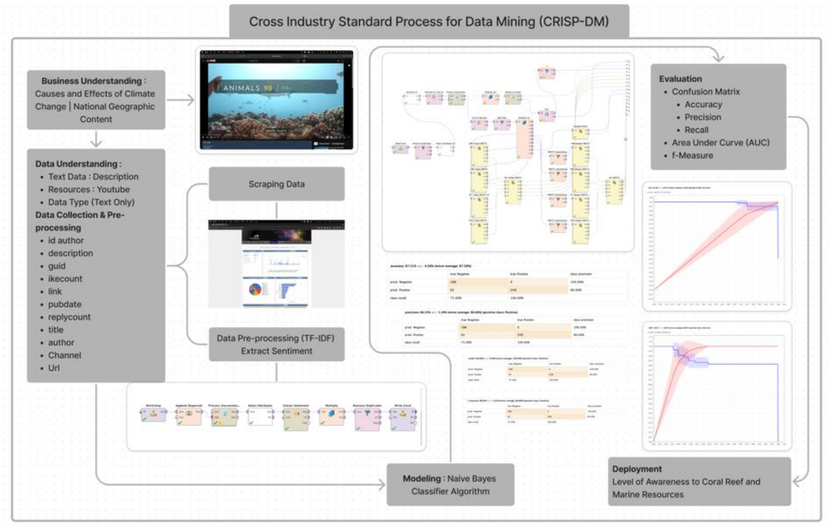

Figure 1. Implementation of CRISP-DM

The algorithm utilized in the modeling stage is the Decision Tree, with the dataset divided into training data (30%) and testing data (70%), further employing the SMOTE operator to address data imbalance. Decision Tree algorithms are well-suited for classification tasks and offer transparency in decision-making processes, making them particularly suitable for analyzing sentiment data [29]. Additionally, utilizing the SMOTE operator enables the generation of synthetic samples to balance class distributions, enhancing the robustness and reliability of the model's predictions [30]. Overall, the combination of the Decision Tree algorithm and SMOTE operator in this research underscores a comprehensive approach to addressing data imbalances and maximizing the predictive performance of the sentiment classification model [31].

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

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) [32]. Lower entropy indicates better performance of the decision tree in classifying data. Decision Trees (DT) offer several advantages in sentiment classification tasks [33]. One of the primary advantages is their interpretability, as DT models generate straightforward decision rules that humans can easily understand and interpret [34]. Additionally, DT models can handle numerical and categorical data without extensive preprocessing, simplifying the data preparation phase [35]. DT models are robust to outliers and missing values, as they partition the feature space based on information gain or Gini impurity [36]. This flexibility enables DT models to effectively capture nonlinear relationships and interactions among features, making them suitable for sentiment classification tasks with complex data patterns [37]. In conclusion, the interpretability, flexibility, and robustness of DT models make them a preferred choice for sentiment classification tasks, providing valuable insights into the factors influencing sentiment polarity.

3. RESULTS AND DISCUSSIONS

Coral Reef 101 content published by the National Geographic Channel has elicited various public perspectives representing negative and positive sentiments. This diversity of views reflects the complex nature of public engagement with environmental issues, particularly those concerning marine conservation. While some individuals may express concerns or criticisms regarding the state of coral reefs and their ecosystems, others may convey admiration and support for preservation efforts. The juxtaposition of these contrasting sentiments underscores the importance of fostering informed discussions and promoting greater awareness of the challenges of coral reefs. Ultimately, this diversity of perspectives enriches the discourse surrounding coral reef conservation and highlights the need for collaborative efforts to safeguard these vital ecosystems.

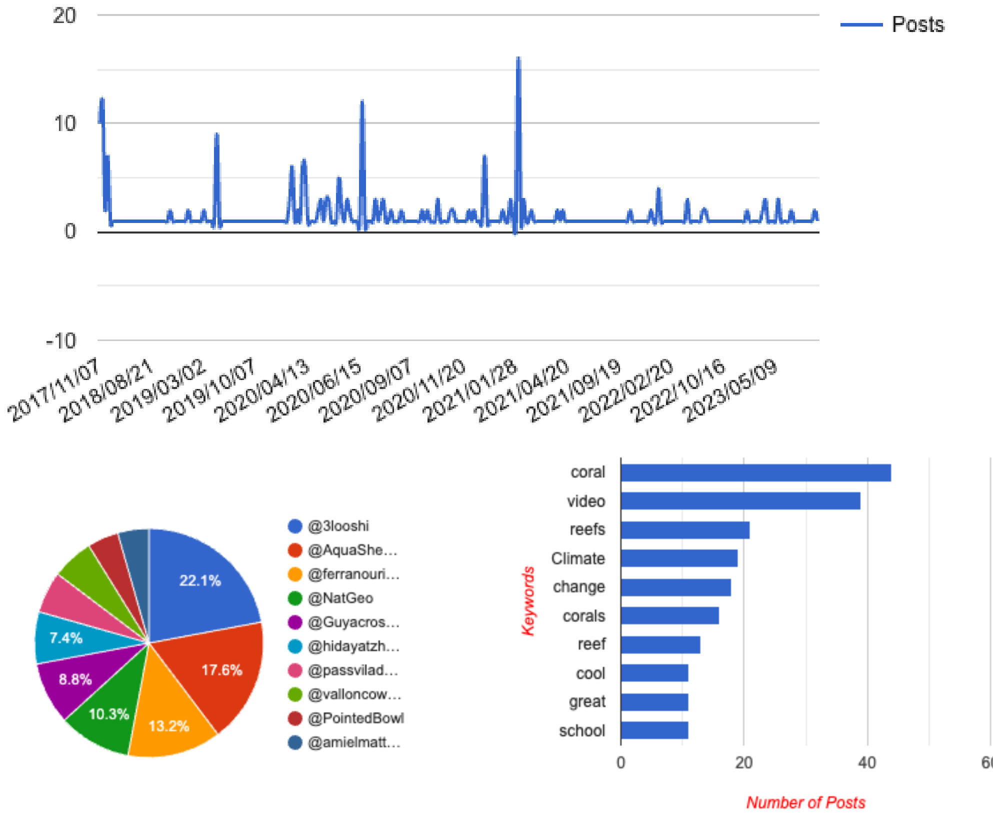

Figure 2. Post Over Time, Top Ten User, and Most Frequently Used Words

Based on the data of the top ten users, top ten frequently used words, and posts over time, insights into initiatives, perceptions, and intentions regarding ecological issues, particularly coral reefs, can be gleaned. This demonstrates that studying public perceptions of coral reef sustainability is an important issue, warranting comprehensive examination. Identifying top users and frequently used words provides a glimpse into the key actors and prevalent topics shaping discussions around coral reef conservation. Additionally, analyzing posts over time allows tracking evolving trends and attitudes toward this ecological issue. The multifaceted nature of these data underscores the significance of understanding public perspectives on coral reef sustainability and the necessity for in-depth investigation to inform effective conservation strategies.

Before performance testing of the classification, public review data that had been collected were extracted using the extract sentiment operator in the RapidMiner application. This initial preprocessing step ensured that the dataset was appropriately prepared for sentiment analysis, allowing for the extraction of sentiment-related information from the raw text data. Using the extract sentiment operator in RapidMiner exemplifies a systematic approach to data preprocessing, underscoring the importance of rigorous data preparation in facilitating accurate and reliable sentiment classification. This initial step lays the foundation for subsequent analysis and evaluation, ultimately contributing to the robustness and validity of the research findings.

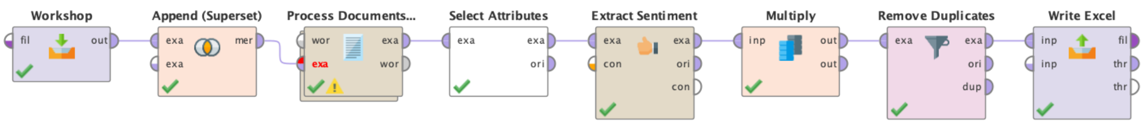

Figure 3. Extract Sentiment Process

This research also utilizes RapidMiner to classify sentiment towards Coral Reef 101. Leveraging RapidMiner enhances the analytical capabilities and efficiency of sentiment classification processes, enabling the extraction of valuable insights from the data. The utilization of RapidMiner underscores the commitment to employing advanced tools and methodologies in addressing complex research questions, particularly in the realm of sentiment analysis and environmental communication. This integration of RapidMiner further enriches the study's methodological framework, facilitating a comprehensive understanding of public perceptions towards coral reef conservation efforts.

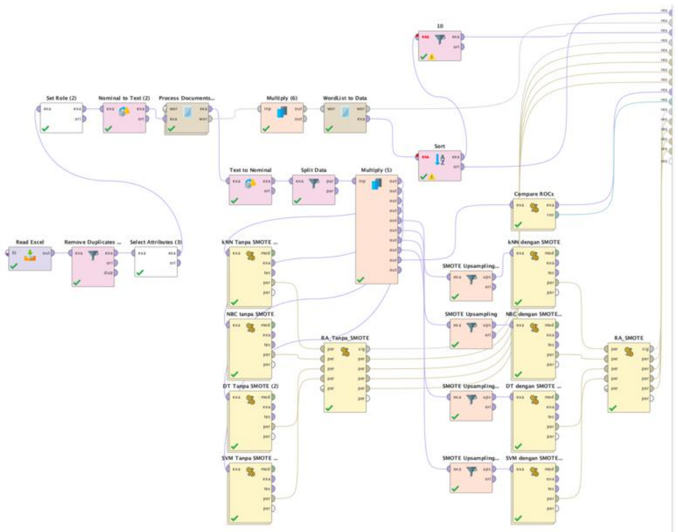

Figure 4. Implementation of Sentiment Classification using Rapidminer Application

The classification results of public sentiment regarding coral reefs show that certain words appear more frequently in the reviews of "Coral Reef 101" content, indicating prevalent themes and topics of interest. Among these, the top 10 most common words identified in the data and their respective frequencies are as follows: "coral" (44), "video" (39), "reefs" (21), "climate" (18), "change" (17), "corals"

(16), "reef" (14), "cool" (13), "plants" (9), and "planet" (9). These findings provide valuable insights into the public's perception and discussion surrounding coral reef-related content. They highlight key terms that resonate with audiences and potentially influence their sentiments toward coral reef conservation efforts. Further analysis of these famous words can inform targeted communication strategies, and content creation approaches to engage and educate the public on coral reef preservation effectively.

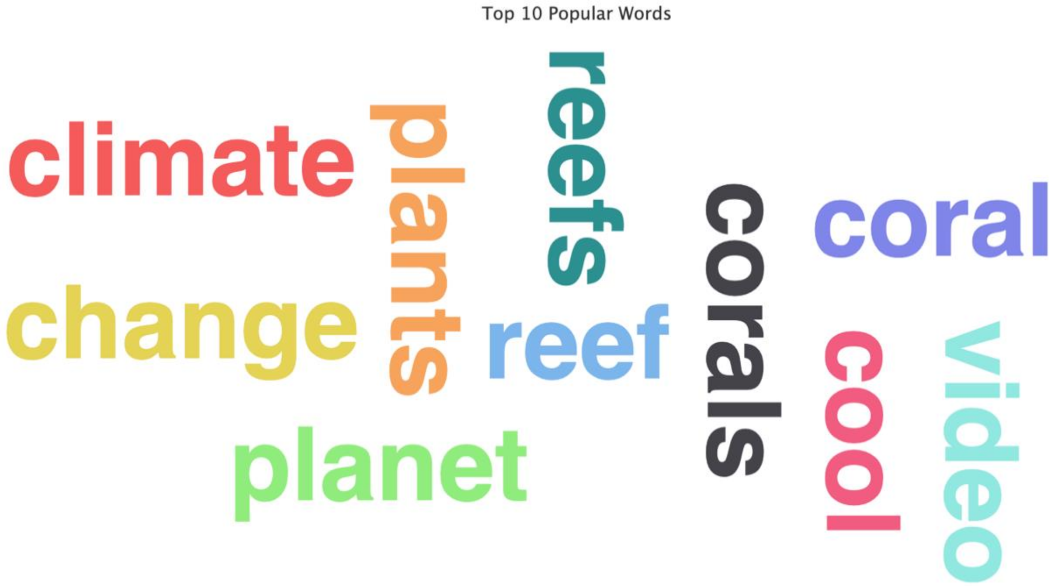

Figure 5. Top 10 Popular Words

Analyzing the most common words in the data set provides insights into the prevailing themes and interests among the public regarding coral reef content. The word "coral" appears most frequently, indicating a strong focus on the coral itself, likely reflecting discussions about its health, diversity, and importance within coral reef ecosystems. The prominence of the word "video" suggests that visual content, such as videos or multimedia presentations, is influential in engaging audiences and disseminating information about coral reefs. Additionally, terms like "reefs," "corals," and "reef" emphasize the significance of the reef ecosystem as a whole, encompassing various aspects such as biodiversity, conservation, and threats. The presence of words like "climate" and "change" highlights the growing concern about the impact of climate change on coral reefs, indicating a recognition of the environmental challenges facing these ecosystems. Moreover, "cool" suggests a positive sentiment or appreciation towards coral reef content, potentially indicating enjoyment or interest in the subject matter. Overall, the analysis underscores the diverse topics and sentiments expressed in public discussions about coral reefs, emphasizing the importance of effective communication and education efforts to promote coral reef conservation and awareness.

After identifying frequently used words, algorithm performance testing was conducted to determine the confusion matrix values. This rigorous evaluation step allows for assessing the classification model's effectiveness in accurately categorizing sentiment. By examining the confusion matrix, researchers can gain insights into the model's performance, including its ability to correctly classify instances into their respective sentiment categories and identify misclassifications. The confusion matrix analysis provides valuable information for refining and optimizing the sentiment classification model, enhancing its overall performance and reliability in capturing public sentiment toward the subject matter.

Table 1. DT performance

DT without SMOTE	DT with SMOTE
PerformanceVector:	PerformanceVector:

Accuracy: 87.07% +/- 3.87% (micro average: 87.11%)	Accuracy: 87.51% +/- 4.28% (micro average: 87.50%)
ConfusionMatrix:	ConfusionMatrix:
True: Negative Positive	True: Negative Positive
Negative: 5 3	Negative: 186 0
Positive: 34 245	Positive: 62 248
AUC (optimistic): 0.989 +/- 0.027 (micro average: 0.989) (positive class: Positive)	AUC (optimistic): 1.000 +/- 0.000 (micro average: 1.000) (positive class: Positive)
AUC: 0.560 +/- 0.075 (micro average: 0.560) (positive class: Positive)	AUC: 0.875 +/- 0.044 (micro average: 0.875) (positive class: Positive)
AUC (pessimistic): 0.132 +/- 0.142 (micro average: 0.132) (positive class: Positive)	AUC (pessimistic): 0.750 +/- 0.087 (micro average: 0.750) (positive class: Positive)
precision: 87.81% +/- 2.52% (micro average: 87.81%) (positive class: Positive)	precision: 80.35% +/- 5.10% (micro average: 80.00%) (positive class: Positive)
ConfusionMatrix:	ConfusionMatrix:
True: Negative Positive	True: Negative Positive
Negative: 5 3	Negative: 186 0
Positive: 34 245	Positive: 62 248
recall: 98.77% +/- 2.79% (micro average: 98.79%) (positive class: Positive)	recall: 100.00% +/- 0.00% (micro average: 100.00%) (positive class: Positive)
ConfusionMatrix:	ConfusionMatrix:
True: Negative Positive	True: Negative Positive
Negative: 5 3	Negative: 186 0
Positive: 34 245	Positive: 62 248
f_measure: 92.94% +/- 2.19% (micro average: 92.98%) (positive class: Positive)	f_measure: 89.02% +/- 3.22% (micro average: 88.89%) (positive class: Positive)
ConfusionMatrix:	ConfusionMatrix:
True: Negative Positive	True: Negative Positive
Negative: 5 3	Negative: 186 0
Positive: 34 245	Positive: 62 248

Based on the evaluation results of the DT algorithm without SMOTE, it is discernible that the model exhibits relatively high-performance metrics across several evaluation criteria. Specifically, the accuracy is reported at 87.07% with a slight standard deviation of +/- 3.87%, indicating consistent performance in correctly classifying sentiment. Moreover, precision, recall, and f-measure metrics demonstrate strong performance in correctly identifying positive sentiments, with precision at 87.81%, recall at 98.77%, and f-measure at 92.94%. However, the AUC score of 0.560 suggests limited discriminatory power in distinguishing between positive and negative sentiments. Despite the high accuracy and precision, the relatively low AUC underscores the need for further optimization or consideration of additional evaluation metrics to enhance the model's performance.

Based on the evaluation results of the DT algorithm with SMOTE, it is evident that the model demonstrates strong performance across various evaluation metrics. The accuracy is reported at 87.51% with a slight standard deviation of +/- 4.28%, indicating consistent performance in correctly classifying sentiment. Furthermore, precision, recall, and f-measure metrics exhibit robust performance in accurately identifying positive sentiments, with precision at 80.35%, recall at 100.00%, and f-measure at 89.02%. The AUC score of 0.875 indicates excellent discriminatory power in distinguishing between positive and negative sentiments. Overall, the evaluation metrics suggest that the DT algorithm with SMOTE effectively addresses data imbalance and yields superior performance in sentiment classification tasks.

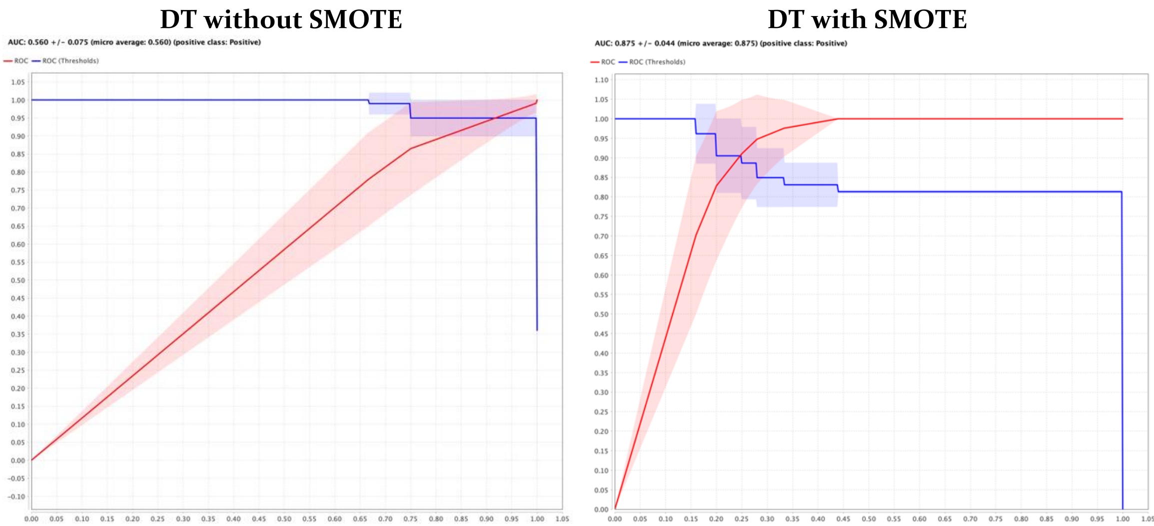

Figure 6. Area Under Curve (AUC) of DT with and without SMOTE

The AUC (Area Under the Curve) metric measures the performance of a classification model across all possible threshold values. A higher AUC value indicates better discrimination between positive and negative classes. We observe a substantial difference when comparing the AUC values between the Decision Tree (DT) algorithm with and without SMOTE. The AUC value is reported as 0.560 in the case without SMOTE, indicating limited discriminatory power in distinguishing between positive and negative sentiments. This suggests that the model without SMOTE may struggle to differentiate between the two classes accurately and may have difficulty correctly classifying instances. With SMOTE, the AUC value significantly improves to 0.875, indicating excellent discriminatory power and a more vital ability to distinguish between positive and negative sentiments. This suggests that the SMOTE technique effectively addressed the data imbalance issue, resulting in a model with enhanced performance and predictive capabilities. Overall, the substantial difference in AUC values highlights the importance of addressing data imbalance through techniques like SMOTE, as it significantly improves the model's ability to classify sentiments and make reliable predictions accurately.

4. CONCLUSION

The results of this study indicate that public sentiment classification towards "Coral Reef 101" content published by the National Geographic Channel is predominantly positive, supporting efforts for marine resource preservation. Furthermore, the utilization of the CRISP-DM methodology alongside the Decision Tree (DT) algorithm and SMOTE technique demonstrates robust performance, as evidenced by high accuracy (87.51% +/- 4.28%), precision (80.35% +/- 5.10%), recall (100.00% +/- 0.00%), f-measure (89.02% +/- 3.22%), and AUC (0.875 +/- 0.044). These findings underscore the effectiveness of the employed approach in accurately classifying sentiment and highlight the importance of leveraging advanced data mining techniques for analyzing public sentiment towards marine conservation efforts. The contribution of this study to knowledge lies in its comprehensive analysis of public sentiment towards coral reef content, mainly focusing on "Coral Reef 101" material from the National Geographic Channel. The research achieves a deep understanding of public perceptions and attitudes toward marine conservation by employing the CRISP-DM methodology in conjunction with the Decision Tree algorithm and SMOTE technique. Additionally, the study's findings provide valuable insights into the effectiveness of sentiment classification models in capturing positive sentiments supportive of marine resource preservation. This research contributes to advancing the field of environmental communication and data mining by offering a systematic approach to analyzing public sentiment and informing targeted conservation strategies.

ACKNOWLEDGEMENTS

I want to express my sincere gratitude to the Tourism Department and the Faculty of Business Administration and Communication at Atma Jaya Catholic University of Indonesia for their invaluable support and encouragement throughout the publication process of this research. Additionally, I would like to extend my heartfelt appreciation to the Research and Community Service Institute (LPPM) for their assistance and resources provided, which have significantly contributed to the realization of this publication.

REFERENCES

[1] T. Kuhn *et al.*, "Literature syntheses to inform marine ecosystem management: lessons learned from stakeholder participation," *Ecosyst. People*, vol. 19, no. 1, pp. 1–16, 2023, doi: 10.1080/26395916.2023.2188970.

[2] F. Cerutti-Pereyra, I. López-Ercilla, G. Sánchez-Rivera, V. Francisco, X. Arvizu-Torres, and T. Adame-Sánchez, "Impact of SCUBA divers on the coral reefs of a national park in the Mexican Caribbean," *J. Ecotourism*, vol. 21, no. 1, pp. 71–86, 2022, doi: 10.1080/14724049.2021.1922422.

[3] Z. Lin, "Chinese legislation on protection of underwater cultural heritage in marine spatial planning and its implementation," *Int. J. Cult. Policy*, vol. 29, no. 4, pp. 500–517, 2023, doi: 10.1080/10286632.2022.2080201.

[4] E. Spooner *et al.*, "Using Integrated Ecosystem Assessments to Build Resilient Ecosystems, Communities, and Economies," *Coast. Manag.*, vol. 49, no. 1, pp. 26–45, 2021, doi: 10.1080/08920753.2021.1846152.

[5] M. He *et al.*, "Coral reef applications of Landsat-8: geomorphic zonation and benthic habitat mapping of Xisha Islands, China," *GIScience Remote Sens.*, vol. 60, no. 1, pp. 1–27, 2023, doi: 10.1080/15481603.2023.2261213.

[6] H. Zhang and S. Chen, "Overview of research on marine resources and economic development," *Mar. Econ. Manag.*, vol. 5, no. 1, pp. 69–83, 2022, doi: 10.1108/maem-11-2021-0012.

[7] H. Tang, M. Lin, J. Yu, and Q. Yue, "New development of marine spatial planning in China: problems and policy suggestions on the implementation of National Plan for Main Functional Zones of Oceans," *Mar. Econ. Manag.*, vol. 5, no. 1, pp. 34–44, 2022, doi: 10.1108/maem-11-2021-0011.

[8] C. Tang, Z. Chen, and J. Peng, "Marine tourism omnichannel coordination," *Mar. Econ. Manag.*, vol. 5, no. 2, pp. 147–172, 2022, doi: 10.1108/maem-01-2022-0003.

[9] V. Reiskarts and K. Savenkovs, "Smarthub for supervising system for resource exploration and pollution control in deep-water and coastal areas based on ICT technologies," *Mar. Econ. Manag.*, vol. 6, no. 1, pp. 23–34, 2023, doi: 10.1108/maem-02-2023-0002.

[10] M. Struwig, A. Van den Berg, and N. Hadi, "Challenges in the ocean economy of South Africa," *Dev. South. Afr.*, vol. 41, no. 1, pp. 1–15, 2023, doi: 10.1080/0376835X.2023.2232396.

[11] A. Mehta, R. Wynberg, C. Ramcharan-Kotze, and A. J. Smit, "Exploring local perceptions around the value of marine biodiversity: the case of kelp in the Western Cape, South Africa," *Ecosyst. People*, vol. 19, no. 1, 2023, doi: 10.1080/26395916.2023.2234499.

[12] J. Davret, B. Trouillet, and H. Toonen, "The digital turn of marine planning: a global analysis of ocean geoportals," *J. Environ. Policy Plan.*, vol. 26, no. 1, pp. 75–90, 2023, doi: 10.1080/1523908X.2023.2283081.

[13] J. M. Ip-Soo-Ching, P. Massingham, and A. Pomering, "Managing diverse knowledge systems of tourism operators in vulnerable marine ecosystems: addressing sustainability challenges through nature-based solutions," *Asia Pacific J. Tour. Res.*, vol. 27, no. 4, pp. 333–356, 2022, doi: 10.1080/10941665.2022.2061364.

[14] E. Macpherson *et al.*, "Designing Law and Policy for the Health and Resilience of Marine and Coastal Ecosystems—Lessons From (and for) Aotearoa New Zealand," *Ocean Dev. Int. Law*, vol. 54, no. 2, pp. 200–252, 2023, doi: 10.1080/00908320.2023.2224116.

[15] D. Eider, S. Partelow, S. Albrecht, L. Adrianto, and L. C. Kluger, "SCUBA tourism and coral reefs: a social-ecological network analysis of governance challenges in Indonesia," *Curr. Issues Tour.*, vol. 26, no. 7, pp. 1031–1050, 2023, doi: 10.1080/13683500.2021.2006612.

[16] A. K. Smith *et al.*, "Innovative local response to cyclone damaged reef leads to rapid tourism recovery," *J. Ecotourism*, vol. 22, no. 3, pp. 354–374, 2023, doi: 10.1080/14724049.2022.2027952.

[17] D. G. Borojo, J. Yushi, Z. Hongyu, L. Xiao, and M. Miao, "A pathway to the green revolution in emerging economies: how does green technological innovation affect green growth and ecological sustainability?," *Econ. Res. Istraz.*, vol. 36, no. 1, pp. 1–24, 2023, doi: 10.1080/1331677X.2023.2167223.

[18] J. J. Lv, N. Wang, H. Ju, and X. F. Cui, "Influence of green technology, tourism, and inclusive financial development on ecological sustainability: exploring the path toward green revolution," *Econ. Res. Istraz.*, vol. 36, no. 1, pp. 1–24, 2023, doi: 10.1080/1331677X.2022.2116349.

[19] Y. A. Singgalen, "Analisis Performa Algoritma NBC , DT , SVM dalam Klasifikasi Data Ulasan Pengunjung Candi Borobudur Berbasis CRISP-DM," *Build. Informatics, Technol. Sci.*, vol. 4, no. 3, pp. 1634–1646, 2022,

doi: 10.47065/bits.v4i3.2766.
[20] Y. A. Singgalen, "Penerapan Metode CRISP-DM dalam Klasifikasi Data Ulasan Pengunjung Destinasi Danau Toba Menggunakan Algoritma Naïve Bayes Classifier (NBC) dan Decision Tree (DT)," *J. Media Inform. Budidarma*, vol. 7, no. 3, pp. 1551–1562, 2023, doi: 10.30865/mib.v7i3.6461.
[21] Y. A. Singgalen, "Analisis Sentimen Wisatawan terhadap Taman Nasional Bunaken dan Top 10 Hotel Rekomendasi Tripadvisor Menggunakan Algoritma SVM dan DT berbasis CRISP-DM," *J. Comput. Syst. Informatics*, vol. 4, no. 2, pp. 367–379, 2023, doi: 10.47065/josyc.v4i2.3092.
[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] 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.
[24] Y. A. Singgalen, "Analisis Sentimen Konsumen terhadap Food , Services , and Value di Restoran dan Rumah Makan Populer Kota Makassar Berdasarkan Rekomendasi Tripadvisor Menggunakan Metode CRISP-DM dan," *Build. Informatics, Technol. Sci.*, vol. 4, no. 4, pp. 1899–1914, 2023, doi: 10.47065/bits.v4i4.3231.
[25] 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.
[26] Y. A. Singgalen, "Analisis Perbandingan Top 10 Best-Value dan Top 10 Traveler- Ranked Hotel Menggunakan MOORA," *J. Comput. Syst. Informatics*, vol. 4, no. 4, pp. 899–911, 2023, doi: 10.47065/josyc.v4i4.4070.
[27] 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.
[28] 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.
[29] H. Christanto *et al.*, "Analisis Perbandingan Decision Tree , Support Vector Machine , dan Xgboost dalam Mengklasifikasi Review Hotel Trip Advisor," *J. Teknol. Inform. dan Komput. MH. Thamrin*, vol. 9, no. 1, pp. 306–319, 2023.
[30] R. Obiedat *et al.*, "Sentiment Analysis of Customers' Reviews Using a Hybrid Evolutionary SVM-Based Approach in an Imbalanced Data Distribution," *IEEE Access*, vol. 10, pp. 22260–22273, 2022, doi: 10.1109/ACCESS.2022.3149482.
[31] D. A. Munawwaroh and A. H. Primandari, "Implementasi Crisp-Dm Model Menggunakan Metode Decision Tree Dengan Algoritma Cart Untuk Prediksi Lila Ibu Hamil Berpotensi Gizi Kurang," *J. Ilm. Pendidik. Mat.*, vol. 10, no. 2, pp. 367–380, 2022, [Online]. Available: http://dx.doi.org/10.31941/delta.v10i2.2172
[32] 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.
[33] A. A. Arifiyanti, M. Fuad, P. Fikri, and B. Utomo, "Analisis Sentimen Ulasan Pengunjung Objek Wisata Gunung Bromo pada Situs Tripadvisor," *Explor. J. Sist. Inf. dan Telemat.*, vol. 13, no. 1, pp. 32–37, 2022.
[34] L. A. Pramesti and N. Pratiwi, "Analisis Sentimen Twitter Terhadap Program MBKM Menggunakan Decision Tree dan Support Vector Machine," *J. Inf. Syst. Res.*, vol. 4, no. 4, pp. 1145–1154, 2023, doi: 10.47065/josh.v4i4.3807.
[35] A. A. Arifiyanti, M. F. Pandji, and B. Utomo, "Analisis Sentimen Ulasan Pengunjung Objek Wisata Gunung Bromo pada Situs Tripadvisor," *Explor. J. Sist. Inf. dan Telemat.*, vol. 13, no. 1, p. 32, 2022, doi: 10.36448/jsit.v13i1.2539.
[36] N. L. W. S. R. Ginantra, C. P. Yanti, G. D. Prasetya, I. B. G. Sarasvananda, and I. K. A. G. Wiguna, "Analisis Sentimen Ulasan Villa di Ubud Menggunakan Metode Naive Bayes, Decision Tree, dan k-NN," *Janapati*, vol. 11, no. 3, pp. 205–216, 2022.
[37] A. T. W. Putra and A. Triayudi, "Analisis Sentimen Pembelajaran Daring menggunakan Metode Naïve Bayes, KNN, dan Decision Tree," *Jurnal JTIK (Jurnal Teknologi* …. pdfs.semanticscholar.org, 2022. [Online]. Available: https://pdfs.semanticscholar.org/177f/4bbf7cd48c9afb2af92cfe8c589bbe81ce6e.pdf