KLIK: Kajian Ilmiah Informatika dan Komputer
ISSN 2723-3898 (Media Online)
Vol 4, No 4, February 2024, Hal 2109-2018
DOI 10.30865/klik.v4i4.1710
https://djournals.com/klik

Social Network Analysis and Sentiment Classification of Robotic Restaurant Content using Naïve Bayes Classifier

Yerik Afrianto Singgalen

Faculty of Business Administration and Communication, Tourism Department, Atma Jaya Catholic University of Indonesia, Jakarta, Indonesia
Email: yerik.afrianto@atmajaya.ac.id
Email Corresponding Author: yerik.afrianto@atmajaya.ac.id

Abstract−Sentiment analysis is crucial in understanding public opinion, particularly in emerging technologies such as automation AI and robotic restaurant services. However, achieving accurate sentiment classification in sentiment analysis tasks poses challenges, especially when dealing with imbalanced data. This study employs the Cross-Industry Standard Process for Data Mining (CRISP-DM) through the Naive Bayes Classifier (NBC) algorithm and Synthetic Minority Over-sampling Technique (SMOTE) to address imbalanced data challenges in sentiment analysis. Social network analysis (SNA) collects and analyzes user-generated content related to automation AI and robotic restaurant services, providing insights into public sentiment. Additionally, the occurrence of frequently used words such as "people" (182), "food" (158), "jobs" (135), "robots" (137), "wage" (102), "work" (78), "robot" (79), "minimum" (78), "fast" (70), and "workers" (65) is examined. The performance of the NBC algorithm with and without SMOTE integration is compared. With SMOTE, the algorithm exhibits an accuracy of 70.11% +/- 3.52%, precision of 88.82% +/- 5.06%, recall of 46.06% +/- 6.13%, AUC of 0.967 +/- 0.016, and F-measure of 60.46% +/- 6.02%. Without SMOTE, the algorithm yields an accuracy of 48.90% +/- 4.36%, precision of 72.15% +/- 5.25%, recall of 44.32% +/- 7.15%, AUC of 0.777 +/- 0.051, and F-measure of 54.57% +/- 5.78%. Recommendations to further enhance the algorithm's performance include exploring additional optimization techniques, such as feature engineering and ensemble methods, and continuing data collection and augmentation efforts to improve dataset representativeness. Regular monitoring and evaluation and iterative refinement based on evolving data patterns are also recommended to ensure sustained effectiveness in sentiment analysis tasks.

Keywords: Social; Network; Sentiment; Classification; Restaurant

1. INTRODUCTION

The emergence of robotic restaurants has gained popularity in Artificial Intelligence (AI) automation. These establishments employ advanced robotics and AI technologies to streamline the dining experience, from food preparation to service [1]. Integrating robotics enhances efficiency and ensures precision and consistency in culinary tasks, thus catering to the growing demand for standardized quality in the food industry [2] . Furthermore, adopting robotic systems in restaurants reflects a paradigm shift towards embracing automation to mitigate labor shortages and rising operational costs [3]. In conclusion, the proliferation of robotic restaurants underscores the transformative potential of AI-driven automation in redefining conventional dining practices and optimizing operational workflows.

Adopting AI and robotic technologies in the service industry poses significant challenges. While these technologies promise to enhance efficiency and improve customer experiences, their integration requires substantial investment in infrastructure, training, and maintenance [4]. Concerns regarding job displacement and the need for human oversight persist, raising ethical and societal implications [5]. Despite these challenges, the potential benefits of AI and robotics in revolutionizing service delivery cannot be overlooked [6]. Thus, navigating the complexities of adoption requires careful consideration of both the opportunities and risks involved, alongside proactive measures to address associated challenges and ensure a balanced approach to technological integration in the service sector.

Implementing AI automation technology in the service industry significantly influences consumer perceptions and comfort. By streamlining processes and enhancing efficiency, AI automation often leads to quicker service delivery and reduced waiting times, positively shaping consumer perceptions of service quality [7]. Moreover, personalized recommendations and tailored experiences facilitated by AI algorithms contribute to increased customer satisfaction and loyalty [8]. However, concerns regarding privacy, data security, and the erosion of human interaction may arise, potentially impacting consumer comfort levels [9]. Despite these challenges, the strategic deployment of AI automation has the potential to redefine service standards and create value for both businesses and consumers alike, necessitating a nuanced approach to balance technological advancements with human-centric considerations.

The circulation of content surrounding AI automation in restaurants and robotic establishments has garnered public sentiment. As media coverage and online discourse increase, there is a growing awareness and curiosity regarding the implications of such technological advancements on various societal facets, including employment dynamics, dining experiences, and cultural perceptions [10]. While proponents applaud the benefits of increased efficiency and innovation, skeptics express concerns about job displacement, loss of human touch in service, and ethical considerations surrounding AI decision-making [11], [12]. This perspective divergence underscores the complex interplay between technological progress and societal values [13], [14]. Thus, fostering informed dialogue and addressing public concerns are essential in shaping the trajectory of AI automation in the restaurant industry and fostering societal acceptance and integration.

Considering public sentiment regarding the evolution of AI automation in the service industry, this research aims to analyze social network patterns based on review data concerning AI automation content in restaurants through a combined approach of social network analysis and sentiment analysis. By examining the interconnectedness of individuals and their sentiments expressed in online reviews, this study seeks to elucidate prevailing attitudes, concerns, and trends surrounding the adoption of AI technology in restaurant settings. By integrating social network analysis and sentiment analysis methodologies, this research provides valuable insights into the public discourse on AI automation, informing policymakers, industry stakeholders, and the public about these technological advancements' implications and potential ramifications.

The urgency of this research lies in the imperative to comprehend and address the societal implications of the rapid advancements in AI automation within the service industry, particularly in the context of restaurants [15], [16]. The growing prevalence of AI technologies in restaurant operations has sparked widespread public discourse, accompanied by a spectrum of sentiments ranging from enthusiasm to skepticism [17], [18]. A thorough investigation is warranted as these sentiments significantly impact consumer behavior, business strategies, and societal norms [1], [19], [20]. Integrating social network analysis and sentiment analysis methodologies is crucial in unraveling the intricate patterns of public sentiment and the interconnectedness of opinions within online platforms [21], [22]. By undertaking such an analysis, this research aspires to provide timely and nuanced insights that can inform decision-makers, policymakers, and industry practitioners about the multifaceted nature of public reactions to AI automation in restaurants, facilitating informed discourse and strategic responses to the evolving landscape of technological integration in the service sector.

The practical and theoretical implications of this research are manifold and significant. By employing a combined approach of social network analysis and sentiment analysis to examine public discourse surrounding AI automation in the restaurant industry, this study offers practical insights for policymakers, industry practitioners, and stakeholders. By identifying prevailing sentiments and elucidating underlying social network structures, this research equips decision-makers with valuable knowledge to navigate the complexities of AI integration in service settings effectively. Furthermore, from a theoretical standpoint, this study contributes to the burgeoning literature on the societal impact of AI technologies by advancing our understanding of public perceptions and attitudes toward automation in the context of restaurants. By bridging the gap between theory and practice, this research informs strategic decision-making and enriches scholarly discourse on the broader implications of technological innovation in contemporary society.

The limitation of this research lies in the methods employed, namely social network analysis and sentiment classification of robotic restaurant content using the Naive Bayes Classifier (NBC) model and SMOTE operator. While these methodologies offer valuable insights into public sentiment and social network structures, they also present inherent constraints and biases. For instance, relying on online review data may overlook individuals who do not actively engage in online platforms, potentially skewing the analysis. Additionally, while widely used, the Naive Bayes Classifier model and SMOTE operator may oversimplify the complex nuances of sentiment expression and fail to capture subtle variations in opinions. Despite these limitations, this research contributes valuable insights into the public discourse surrounding AI automation in the restaurant industry. It provides a foundation for further exploration using more nuanced methodologies and datasets.

Incorporating related research as a reference for this study enables a comprehensive examination of the existing literature, facilitating the identification of distinctions between the current research and previous studies. Prior investigations into sentiment analysis, particularly those focusing on the performance of machine learning algorithms in handling imbalanced datasets, offer valuable insights. Notable references include works where Naive Bayes Classifier (NBC) was employed in sentiment classification tasks. While these studies provide foundational knowledge, the present research contributes by extending the analysis to the context of robot hotel reviews and further exploring the impact of the SMOTE on model performance. This juxtaposition with previous research enriches the scholarly discourse and elucidates the novel contributions of the current study within the broader landscape of sentiment analysis research.

The contribution to knowledge stemming from this research is substantial, as it sheds light on the intricate dynamics of public sentiment and social network structures surrounding AI automation in the restaurant industry. By employing social network analysis and sentiment classification techniques, this study enriches our understanding of the multifaceted implications of technological innovation on societal perceptions and behaviors. Moreover, this research opens avenues for further exploration and development in several directions. Future research endeavors may focus on refining sentiment analysis methodologies to capture nuanced opinions more accurately, exploring additional factors influencing public sentiment, such as cultural differences or economic factors, and investigating the long-term effects of AI automation on consumer behavior and industry practices [23]–[25]. Through continued inquiry and interdisciplinary collaboration, scholars can advance our understanding of the complex interplay between technology, society, and human behavior, thus informing strategic decision-making and societal responses to technological advancements.

2. RESEARCH METHODOLOGY

2.1 Gap Analysis of Robotic and AI Automation Restaurant Topics using Vosviewer

The research gap within the realm of robotic restaurants and AI automation in restaurants is notable and warrants scholarly attention. While existing literature provides valuable insights into the technological advancements and potential benefits of AI integration in restaurant operations, there remains a dearth of comprehensive studies addressing the holistic impact of these innovations on various stakeholders, including consumers, employees, and industry practitioners. Additionally, empirical research focusing on the nuanced interplay between technological adoption, societal perceptions, and cultural contexts is limited, highlighting the need for interdisciplinary investigations incorporating social sciences, ethics, and human-centered design principles. Addressing these research gaps is essential for fostering a comprehensive understanding of the implications of AI automation in restaurant settings and informing evidence-based strategies for sustainable integration and societal acceptance.

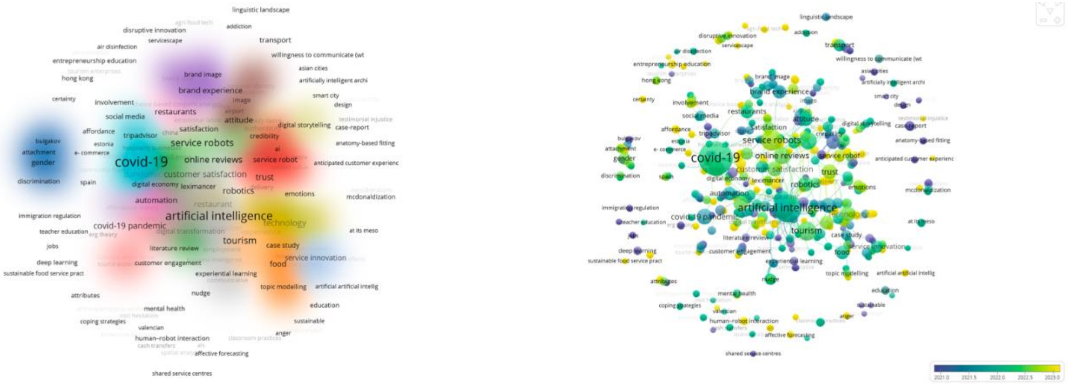

Figure 1. Gap Analysis of Automation AI and Robotic Restaurant Topics

Figure 1 shows the gap analysis of automation AI and robotic restaurant topics. Based on the research gap analysis, it is evident that the topics of artificial intelligence, robotics, and COVID-19 form distinct clusters, with studies focusing on the intersection of robotics and AI in the food and beverage industry emerging as a relatively new and popular area of inquiry. While research on AI and robotics spans various sectors, including healthcare and manufacturing, the specific application of these technologies in the food and beverage sector has garnered increasing attention due to its potential to revolutionize operational efficiency, customer experiences, and safety protocols, particularly in light of the ongoing challenges posed by the COVID-19 pandemic. This observation underscores the significance of further exploring and understanding the implications of AI and robotics in the food and beverage industry, thereby contributing to both academic scholarship and practical innovation in the field.

2.2 Cross-Industry Standard Process for Data Mining (CRISP-DM)

The research methodology employed in this study is the Cross-Industry Standard Process for Data Mining (CRISP-DM). CRISP-DM provides a structured framework for guiding the entire data mining process, encompassing six phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment. By adhering to this rigorous methodology, researchers systematically navigate the complexities of data analysis and modeling, ensuring transparency, reproducibility, and rigor in the research process [26]–[35]. Leveraging CRISP-DM facilitates methodological clarity and enhances the validity and reliability of findings, thereby bolstering the credibility and trustworthiness of research outcomes.

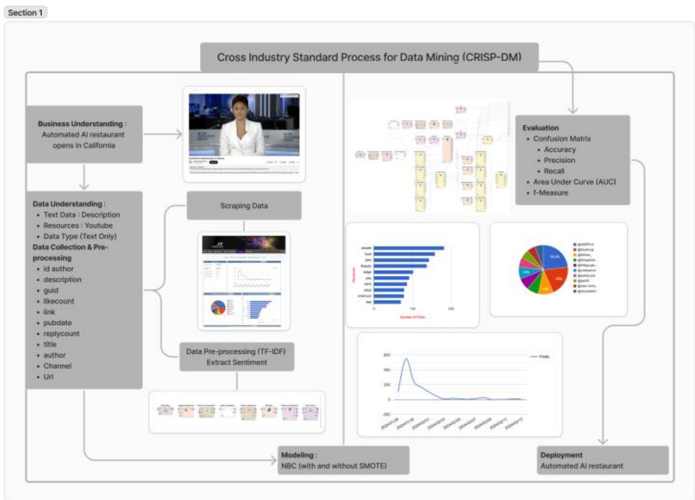

Figure 2. Implementation of CRISP-DM in Social Network Analysis and Sentiment Classification

Figure 2 shows the framework of CRISP-DM in social network analysis and sentiment classification process. In the business understanding phase, it is crucial to comprehend the contextual intricacies of the discussion and analysis of AI automation and robotic restaurants. This phase is the foundation for delineating the research endeavor's objectives, requirements, and constraints within the broader business context. Understanding the nuances of AI automation and robotic restaurant technologies entails examining the technical functionalities, socio-economic implications, ethical considerations, and market dynamics associated with their implementation. By elucidating the underlying business landscape and stakeholders' perspectives, researchers can effectively frame research questions, identify relevant data sources, and delineate the scope of analysis, laying the groundwork for a comprehensive and insightful investigation into the intersection of AI automation and robotic restaurants.

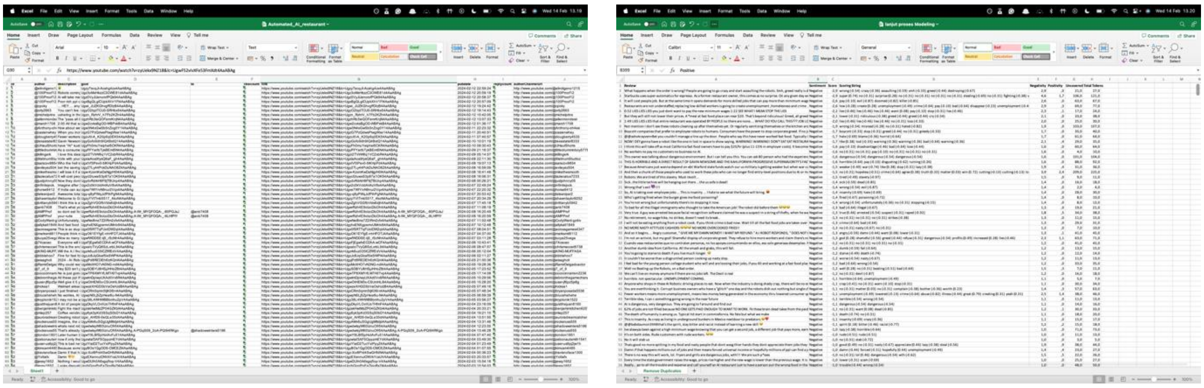

Figure 3. Scraping Data and Social Network Analysis through Netlytic

Figure 3 shows the result of scraping data regarding social network analysis through Netlytic. In the data understanding phase, it is essential to comprehend the sources of data and platforms utilized in implementing social network analysis through Netlytic. Content related to AI automation in restaurants was obtained from the YouTube platform, specifically from the video with the identifier "zyUekx9NZ18," comprising 1655 comments. Understanding the data sources and platforms is paramount for ensuring the collected data's reliability, validity, and relevance for subsequent analysis. By acknowledging and delving into the specifics of data acquisition, researchers can effectively navigate the complexities of social network analysis and derive meaningful insights from the data.

After obtaining data about author ID, description, GUID, like count, link, publication date, reply count, title, author, channel, and URL, the analysis examines social networks using the following Social Network Analysis (SNA) equation. The availability of such comprehensive data sets facilitates a nuanced exploration of the interconnections and dynamics within social networks, allowing researchers to uncover patterns, relationships, and influential nodes that shape information flow and interaction patterns. By applying the principles of SNA to the gathered data, researchers can gain valuable insights into the structure, cohesion, and influence dynamics of the social networks under study, thus enriching our understanding of the underlying phenomena and informing strategic decision-making processes.

$$[C = \frac{L}{N(N-1)}] \tag{1}$$

Where:
(C) is the connectivity coefficient
(L) is the number of links
(N) is the number of nodes

The aim of Social Network Analysis (SNA) in identifying public sentiment on video content regarding AI automation in restaurants is multifaceted. Primarily, SNA unveils the underlying network structures and patterns of interactions among individuals engaging with such content, thereby providing insights into the dissemination of sentiments and forming opinion clusters within online communities [36]. By analyzing the connections between users, their interactions, and the sentiments expressed in comments or interactions, SNA enables researchers to discern influential nodes, identify opinion leaders, and understand the dynamics of sentiment propagation within social networks [37]. Consequently, applying SNA in this context facilitates a comprehensive understanding of public sentiment surrounding AI automation in restaurant settings, contributing to informed decision-making and strategic interventions to address concerns or capitalize on positive perceptions.

Subsequently, in the modeling phase for sentiment analysis, the NBC algorithm with the SMOTE operator is employed to address data imbalance. This phase transforms raw data into actionable insights by applying appropriate analytical techniques. Using the NBC algorithm and the SMOTE operator is advantageous in mitigating the challenges posed by imbalanced datasets commonly encountered in sentiment analysis tasks [38]. By leveraging the strengths of NBC in probabilistic classification and the SMOTE operator in generating synthetic minority samples, this approach enhances sentiment analysis models' robustness and generalization capabilities, yielding more accurate and reliable sentiment classifications [39]. Thus, integrating NBC with the SMOTE operator in the modeling phase represents a methodological refinement that enhances the effectiveness and applicability of sentiment analysis techniques in capturing the nuanced sentiments expressed in the online content [40]. The Naive Bayes Classifier (NBC) algorithm calculates the probability of a class label given a set of features using Bayes' theorem. Mathematically, it can be represented as:

$$[P(C_k|x_1, x_2, \dots, x_n) = \frac{P(x_1, x_2, \dots, x_n|C_k) \cdot P(C_k)}{P(x_1, x_2, \dots, x_n)}] \tag{2}$$

Where:
$(P(C_k|x_1, x_2, \dots, x_n))$ is the probability of class (C_k) given features $(x_1, x_2, \dots, x_n)$,

$(P(x_1, x_2, \ldots, x_n \mid C_k))$the conditional probability of observing features $(x_1, x_2, \ldots, x_n)$given class (C_k),
$(P(C_k))$ is the prior probability of class (C_k),
$(P(x_1, x_2, \ldots, x_n))$ the probability of observing features $(x_1, x_2, \ldots, x_n)$.

The "naive" assumption in NBC is that the features are conditionally independent given the class, simplifying the conditional probability calculation. The Naive Bayes Classifier (NBC) algorithm possesses several advantages in sentiment classification tasks. Primarily, NBC is computationally efficient and requires relatively low computational resources, making it suitable for handling large datasets commonly encountered in sentiment analysis applications [41]. Additionally, NBC performs well even with limited training data and is robust to noise, making it particularly suitable for scenarios where data quality may vary [42]. Moreover, the simplicity of NBC allows for straightforward implementation and interpretation, facilitating its adoption and usage across diverse domains [43]. Furthermore, NBC's probabilistic framework quantifies uncertainty in classification decisions, providing valuable insights into the confidence levels of sentiment predictions [44]. The NBC algorithm offers a compelling solution for sentiment classification tasks, combining computational efficiency, robustness, simplicity, and interpretability to analyze and classify sentiment in various textual data sources effectively.

3. RESULT AND DISCUSSION

The discourse surrounding AI automation and robotic restaurants elicits both favorable and unfavorable sentiments, as evidenced by public sentiment analysis of video content. Proponents of AI automation and robotic restaurants often highlight the potential benefits, such as improved efficiency, enhanced customer experiences, and innovative dining concepts [45]. They argue that these technologies can streamline operations, reduce labor costs, and offer unique dining experiences that cater to modern consumer preferences [46]. However, critics express concerns regarding job displacement, loss of human touch in service, and ethical implications associated with AI decision-making [47]. They argue that excessive reliance on automation may undermine the importance of human interaction and lead to social and economic disparities [48]. Despite these contrasting viewpoints, analyzing public sentiment on video content provides valuable insights into the multifaceted nature of opinions surrounding AI automation and robotic restaurants, underscoring the need for informed discourse and balanced approaches to technological integration in the food service industry.

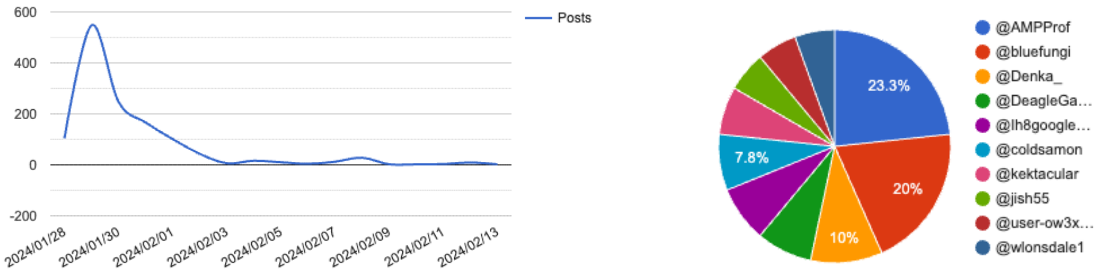

Figure 4. Post over Time and Top Ten Posters

Figure 4 sohws the post over time and top ten posters of the content. Based on the analysis of user data providing reviews on video content, it is discernible that comments related to AI automation and robotic restaurants have increased in 2024 based on posts over time. This trend underscores individuals' growing interest and engagement in discussions surrounding integrating AI technology and robotics in restaurant operations. The surge in comments suggests heightened awareness and evolving attitudes towards these technologies within the public sphere. This observation highlights the importance of continuously monitoring and analyzing user-generated content to glean insights into shifting perceptions and trends, informing strategic decisions and interventions in response to emerging issues or opportunities in AI automation and robotic restaurants.

Figure 5. Chain Network "who replies to whom" Layout

Figure 5 shows the chain network layout with "who replies to whom" categories. Based on the results of Social Network Analysis (SNA), it is apparent that the network exhibits a Diameter of 3, indicating the maximum distance between any pair of nodes. The Density of 0.001068 suggests a low level of connectivity within the network, implying that only a tiny fraction of possible connections are realized. The absence of Reciprocity (0.000000) indicates a lack of mutual connections among nodes. At the same time, the Centralization value of 0.007923 suggests that the network is relatively decentralized, with no single node exerting significant control over the network's flow of information. Furthermore, the high Modularity value of 0.977900 indicates the presence of distinct communities or clusters within the network, indicating the segmentation of nodes based on common attributes or interactions. Overall, these metrics provide valuable insights into the structural characteristics and organization of the network, facilitating a deeper understanding of the patterns of interaction and information flow among nodes.

After conducting Social Network Analysis (SNA), the sentiment analysis proceeded by extracting sentiment through RapidMiner and displaying frequently used words. This subsequent analysis aims to delve deeper into the sentiments expressed within the network and uncover prevalent themes or topics of discussion. By extracting sentiment from textual data, researchers can gain insights into the prevailing attitudes, emotions, and opinions circulating within the network, thereby enhancing the understanding of user perceptions and interactions. Moreover, identifying frequently used words provides valuable context and allows for the identification of critical topics or concerns driving the sentiment patterns observed in the network. Overall, the integration of sentiment analysis complements the insights gleaned from SNA, offering a comprehensive view of the dynamics and sentiment landscape within the analyzed social network.

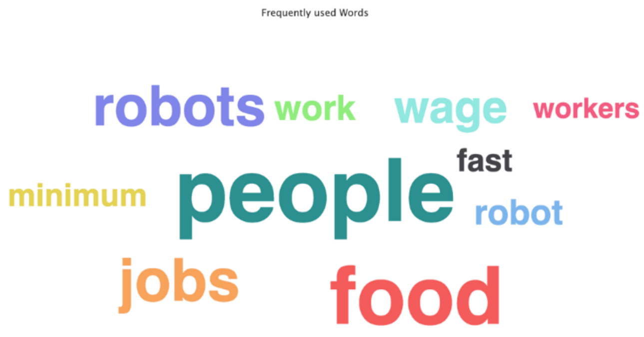

Figure 6. Frequently used Words

Figure 6 shows the frequently used words in the content revies. Based on the results of identifying frequently used words, it is apparent that specific prominent themes emerge within the discourse. The data reveals that words such as "people" (182 occurrences), "food" (158 occurrences), "jobs" (135 occurrences), "robots" (137 occurrences), and "wage" (102 occurrences) indicate a focus on topics related to employment, automation, and labor conditions, suggesting that discussions within the analyzed social network revolve around the impact of AI automation and robotic technologies on workforce dynamics and economic considerations. Additionally, terms like "minimum" (78 occurrences), "fast" (70 occurrences), and "workers" (65 occurrences) further underscore concerns regarding labor rights, wages, and the pace of work in the context of technological innovation. This analysis highlights the significance of these themes within the discourse. It underscores the importance of addressing socio-economic implications and ethical considerations for adopting AI automation and robotic systems in various industries.

The analysis of frequently used words, mainly terms like "robots," "jobs," "wage," "workers," and "food," suggests a significant intersection between the discourse surrounding AI automation, robotic restaurants, and socio-economic concerns. The prominence of terms related to employment and labor conditions indicates a preoccupation with the impact of automation on workforce dynamics and livelihoods within the context of the food service industry. The high frequency of "robots" suggests a keen interest in the role of robotic technologies in restaurant operations, potentially reflecting debates surrounding job displacement, skills retraining, and the quality of work in an increasingly automated environment. Additionally, the recurrence of terms like "wage" and "workers" underscores concerns about fair compensation, working conditions, and the distribution of economic benefits amidst technological advancements. Moreover, "food" as a frequently used word implies a broader discussion encompassing the quality, accessibility, and sustainability of food products and services in the context of automation. Overall, the analysis highlights the multifaceted nature of discussions surrounding AI automation and robotic restaurants, emphasizing the need for nuanced approaches that consider socio-economic implications alongside technological innovation.

Table 1. Classification based on Extract Sentiment Result

Review	String Score	Score	Classification
What happens when the order is wrong? People are going to go crazy and start	wrong (-0.54) crazy (-0.36) assaulting (-0.59) smh (-0.33) greed (-	-2,92307692307692	Negative

assaulting the robots. Smh, greed really is destroying everything.	0.44) destroying (-0.67)		
The key difference between machines used until now and new ones based on Artificial Intelligence (AI) is in their ability to learn, adapt, and respond to various situations. Previously, most robots in factories were programmed for specific tasks with limited ability to handle changes or unforeseen circumstances. AI-based machines, such as those in a restaurant making hamburgers and fries, can learn and adapt to different scenarios. They process data, recognize patterns, and perform more complex tasks efficiently, considering various variables. For example, an AI machine can tailor hamburger preparation to specific customer preferences or handle anomalies during cooking. Additionally, AI-based machines can improve over time through continuous learning, making them more efficient and adaptable to changing needs.	intelligence (0.54) ability (0.33) limited (-0.23) ability (0.33) efficiently (0.44) improve (0.49) efficient (0.46)	2,35897435897436	Positive

Table 1 indicates the classification result based on extract sentiment in Rapidminer. The extract sentiment operator can optimize classifying negative and positive classes by computing scores from words containing positive or negative sentiments. This operator plays a crucial role in sentiment analysis tasks by systematically identifying and evaluating the polarity of words within textual data, enabling the classification of sentiments as either positive or negative. By assigning scores to words based on their sentiment orientation, the operator facilitates quantifying and interpreting sentiment patterns, thereby enhancing the accuracy and effectiveness of sentiment classification algorithms. Overall, the extract sentiment operator is a valuable tool in sentiment analysis workflows, providing researchers with the means to comprehensively discern and analyze sentiments expressed within textual data.

Performance testing of machine classifiers can be conducted by adopting the NBC algorithm to measure accuracy, recall, precision, F-measure, and even AUC. This approach offers a comprehensive evaluation of classifier performance across various metrics, allowing researchers to assess the effectiveness and robustness of the classification model. By measuring accuracy, recall, precision, F-measure, and AUC, researchers can gain insights into the classifier's ability to correctly classify instances, identify true positives, avoid false positives, balance precision and recall, and effectively discriminate between positive and negative instances. Consequently, adopting the NBC algorithm for performance testing enables researchers to make informed decisions regarding the suitability and reliability of the classifier for sentiment analysis tasks, thereby enhancing the quality and rigor of analytical outcomes.

Table 2. Confusion Matrix of NBC with and without SMOTE

Confusion Matrix	NBC with SMOTE	NBC without SMOTE
Accuracy	70.11%	48.90%
Precision	88.82%	72.15%
Recall	46.06%	44.32%
F-measure	60.46%	54.57%
AUC	0.683	0.446

Table 2 shows the confusion matrix of NBC with and without SMOTE. The integration of SMOTE with NBC significantly enhances the classification performance, as evidenced by the notable improvement in accuracy to 70.11% +/- 3.52%, compared to the model without SMOTE. The confusion matrix reveals a substantial reduction in false negatives (37) and false positives (342), while actual positive instances (292) and true negatives (597) increase. Furthermore, the AUC scores indicate improved discrimination ability, with an optimistic AUC at 0.967 +/- 0.016 and a pessimistic AUC at 0.462 +/- 0.059. Precision for the positive class notably increases to 88.82% +/- 5.06%, indicating a higher proportion of correctly classified positive instances. However, recall for the positive class remains relatively low at 46.06% +/- 6.13%, suggesting challenges in identifying all positive instances. Nonetheless, the F-measure for the positive class improves to 60.46% +/- 6.02%, reflecting an overall enhancement in classification performance with SMOTE integration. In contrast, the performance of NBC without SMOTE demonstrates moderate classification performance, with accuracy at 48.90% +/- 4.36%. While the model accurately predicts negative instances (162 true negatives), it struggles with positive instances (281 true positives), resulting in a significant number of false negatives (110) and false positives (353). The AUC scores suggest moderate discrimination ability, with an optimistic AUC at 0.777 +/- 0.051 and a pessimistic AUC at 0.276 +/- 0.052. Precision, recall, and F-measure values for the positive class further elucidate the model's shortcomings in correctly identifying positive instances. These findings underscore the limitations of NBC without SMOTE in handling imbalanced data effectively and accurately classifying positive instances, emphasizing the efficacy of SMOTE in enhancing classification performance for sentiment analysis tasks.

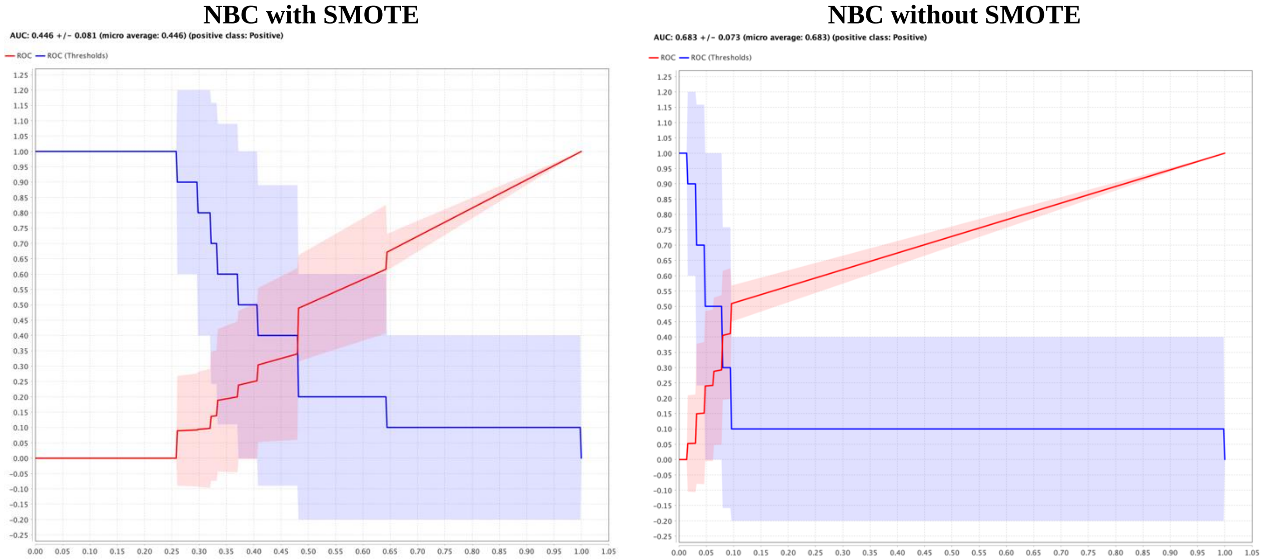

Figure 7. Area Under Curve of NBC With and Without SMOTE

Figure 7 shows the AUC of NBC with and without SMOTE. The comparison and analysis of the Area Under the Curve (AUC) values provide valuable insights into the discrimination ability of the classification models. In this comparison, the AUC values indicate the models' capacity to distinguish between positive and negative instances accurately. The AUC values of the NBC model with SMOTE integration demonstrate a significant improvement compared to the model without SMOTE. Specifically, the optimistic AUC value of 0.967 +/- 0.016 indicates a high level of confidence in the model's discriminatory power, supported by the narrower confidence interval. Conversely, the NBC model without SMOTE exhibits lower AUC values, reflecting a less robust discriminatory performance. This comparison underscores the efficacy of SMOTE in enhancing the classification model's ability to accurately differentiate between positive and negative instances, highlighting its importance in addressing imbalanced data challenges in sentiment analysis tasks.

The confusion matrix analysis reveals crucial aspects of the algorithm's performance based on the provided data. In the case of the model without SMOTE, a relatively high number of false positives (353) and false negatives (110) compared to true positives (281) and true negatives (162) suggests a significant challenge in correctly identifying positive instances. This imbalance in the distribution of classification outcomes indicates potential limitations in the model's ability to classify positive instances, affecting its overall performance accurately. Conversely, with SMOTE integration, the model demonstrates a notable reduction in false positives (342) and false negatives (37), accompanied by an increase in true positives (292) and true negatives (597). This improvement reflects a more balanced distribution of classification outcomes and suggests enhanced performance in correctly classifying positive instances. Overall, the confusion matrix analysis highlights the effectiveness of SMOTE in mitigating imbalanced data challenges and improving the algorithm's ability to accurately classify positive instances, thereby enhancing its overall performance in sentiment analysis tasks.

In providing recommendations, it is essential to consider strategies that address the identified challenges and enhance the overall effectiveness of the algorithm. Firstly, implementing further optimization techniques, such as feature engineering or algorithm tuning, may help improve the model's classification performance. Additionally, exploring ensemble methods, such as combining multiple classifiers or utilizing ensemble learning frameworks, could enhance the algorithm's predictive capabilities and robustness. Furthermore, continued efforts in data collection and augmentation, mainly focusing on increasing the diversity and representativeness of the dataset, may help mitigate imbalanced data challenges and improve the model's performance, especially in classifying positive instances. Lastly, regular monitoring and evaluation of the algorithm's performance and iterative refinement and adaptation based on evolving data patterns and insights are crucial for ensuring sustained effectiveness in sentiment analysis tasks. In conclusion, by incorporating these recommendations into algorithm development and implementation processes, researchers can strive towards achieving more accurate and reliable sentiment analysis outcomes in practical applications.

4. CONCLUSION

In summary, this research has significantly advanced sentiment analysis within the context of automation AI and robotic restaurant services. Through the implementation of sophisticated algorithms such as the Naive Bayes Classifier (NBC) and Synthetic Minority Over-sampling Technique (SMOTE), in conjunction with Social Network Analysis (SNA), the study has effectively addressed challenges associated with imbalanced data and classification accuracy. The findings reveal substantial improvements in various performance metrics. Specifically, with SMOTE integration, there is a noteworthy increase in accuracy to 70.11%, precision to 88.82%, recall to 46.06%, and F-measure to 60.46%. Additionally, the Area Under the Curve (AUC) significantly enhances to 0.967. The analysis of frequently used words further enriches our understanding, with "people" occurring 182 times, "food" 158 times, "jobs" 135 times, "robots" 137 times, "wage" 102 times, "work" 78 times, "robot" 79 times, "minimum" 78 times, "fast" 70 times, and "workers" 65 times. These outcomes underscore the effectiveness of SMOTE in augmenting classification performance and highlight the importance of gauging public sentiment in evolving technological landscapes. Future research endeavors should explore additional optimization strategies and continual data refinement to sustain and enhance sentiment analysis outcomes. Regular monitoring, evaluation, and iterative refinement based on evolving data patterns are imperative to ensure continued efficacy of sentiment analysis methodologies. Overall, this study contributes invaluable insights and methodologies applicable across diverse industries and decision-making contexts, laying the groundwork for future advancements in sentiment analysis research.

REFERENCES

[1] Z. Li, F. Yuan, and Z. Zhao, “Robot restaurant experience and recommendation behaviour: based on text-mining and sentiment analysis from online reviews,” *Curr. Issues Tour.*, pp. 1–15, 2024, doi: 10.1080/13683500.2024.2309140.

[2] V. Sharma, M. Kapse, J. Poulose, and Y. Mahajan, “Robotic dining delight unravelling the key factors driving customer satisfaction in service robot restaurants using PLS-SEM and ML,” *Cogent Bus. Manag.*, vol. 10, no. 3, 2023, doi: 10.1080/23311975.2023.2281053.

[3] D. H. Zhu and W. Q. Zhang, “Do we have to know the hen that lays the eggs? How the disclosure timing of robot chef identity affects consumer perceived ethicality,” *J. Sustain. Tour.*, vol. 0, no. 0, pp. 1–18, 2023, doi: 10.1080/09669582.2023.2266781.

[4] J. Na Nongkhai, C. Suttikun, and P. Mahasuweerachai, “Why Young Consumers Engage with Mobile Self-Ordering Applications (MFOAs) at Restaurants During COVID-19: The Role of Functional and Psychological Factors,” *J. Int. Food Agribus. Mark.*, vol. 0, no. 0, pp. 1–25, 2023, doi: 10.1080/08974438.2023.2255573.

[5] P. Plumwongrot and P. Pholphirul, “Are Robots stealing jobs? Empirical evidence from 10 developing countries,” *Econ. Innov. New Technol.*, vol. 32, no. 6, pp. 873–889, 2023, doi: 10.1080/10438599.2022.2051020.

[6] G. El Hajal and B. Rowson, “The future of hospitality jobs,” *Res. Hosp. Manag.*, vol. 10, no. 1, pp. 55–61, 2020, doi: 10.1080/22243534.2020.1790210.

[7] O. Chehab, M. Bouzari, and S. Ilkhanizadeh, “Job Motivation Mediating the Effects of Standardization on Restaurant Frontline Employees’ Job Outcomes: Evidence from Lebanon,” *J. Qual. Assur. Hosp. Tour.*, vol. 24, no. 6, pp. 836–858, 2023, doi: 10.1080/1528008X.2022.2089944.

[8] S. Alotaibi, K. Shafieizadeh, and K. Alsumait, “Customers’ perceptions toward service robots at restaurants: The moderating role of consumer innovativeness and the mediating role of perceived hedonic benefits,” *J. Qual. Assur. Hosp. Tour.*, vol. 00, no. 00, pp. 1–23, 2022, doi: 10.1080/1528008X.2022.2161693.

[9] L. Xie, X. Liu, and D. Li, “The mechanism of value cocreation in robotic services: customer inspiration from robotic service novelty,” *J. Hosp. Mark. Manag.*, vol. 31, no. 8, pp. 962–983, 2022, doi: 10.1080/19368623.2022.2112354.

[10] M. Nuyken, D. Zilbershtein, and A. Rauf, “Generation Z’s perspective on restaurant service quality,” *Res. Hosp. Manag.*, vol. 12, no. 3, pp. 309–318, 2022, doi: 10.1080/22243534.2023.2202495.

[11] D. M. Zemke, C. Raab, and W. Han, “Factors influencing QSR guests’ acceptance of robotic technology,” *J. Foodserv. Bus. Res.*, vol. 00, no. 00, pp. 1–21, 2023, doi: 10.1080/15378020.2023.2265795.

[12] F. Seyitoğlu, O. Atsız, S. Taş, and F. Kaya, “Double-edged perspectives on service robots: working with robots and robots’ future career impacts,” *J. Teach. Travel Tour.*, vol. 23, no. 1, pp. 1–19, 2023, doi: 10.1080/15313220.2022.2076768.

[13] H. Yang, H. Song, L. Xia, and A. Yang, “Understanding the consequences of robotic interaction quality and outcome quality: A three-phased affordance theory-based approach,” *J. Hosp. Mark. Manag.*, vol. 00, no. 00, pp. 1–21, 2024, doi: 10.1080/19368623.2024.2308531.

[14] F. Seyitoğlu, O. Atsız, and A. Acar, “The future of restaurant labour: evidence from the U.S. restaurants,” *Curr. Issues Tour.*, vol. 0, no. 0, pp. 1–18, 2023, doi: 10.1080/13683500.2023.2280150.

[15] A. Ndiuini and T. Baum, “Underemployment and lived experiences of migrant workers in the hotel industry: policy and industry implications,” *J. Policy Res. Tour. Leis. Events*, vol. 13, no. 1, pp. 36–58, 2021, doi: 10.1080/19407963.2020.1743487.

[16] R. Katz, F. Callorda, and J. Jung, “The impact of automation on employment and its social implications: evidence from Chile,” *Econ. Innov. New Technol.*, vol. 32, no. 5, pp. 646–662, 2023, doi: 10.1080/10438599.2021.1991798.

[17] N. H. Md Saad, F. Mei Nie, and Z. Yaacob, “Exploring sentiment analysis of online food delivery services post COVID-19 pandemic: grabfood and foodpanda,” *J. Foodserv. Bus. Res.*, vol. 00, no. 00, pp. 1–25, 2023, doi: 10.1080/15378020.2023.2294406.

[18] J. Im, T. Song, Y. Lee, and J. Kim, “Confirmatory aspect-level opinion mining processes for tourism and hospitality research: a proposal of DiSSBUS,” *Curr. Issues Tour.*, vol. 25, no. 12, pp. 1876–1894, 2022, doi: 10.1080/13683500.2021.1974357.

[19] H. T. Phan, N. T. Nguyen, D. Hwang, and Y. S. Seo, “M2SA: a novel dataset for multi-level and multi-domain sentiment analysis,” *J. Inf. Telecommun.*, vol. 7, no. 4, pp. 494–512, 2023, doi: 10.1080/24751839.2023.2229700.

[20] A. Hashimoto, D. J. Telfer, and K. Telfer, “Eurocentric cultural theme parks in Japan: domestic tourists’ perspectives on place branding,” *J. Tour. Cult. Chang.*, vol. 21, no. 3, pp. 344–363, 2023, doi: 10.1080/14766825.2023.2178314.

[21] X. Li, J. Francois, and J. Kwon, “Investigating Consumers’ Online Restaurant Selection Behaviors Using Eye-tracking Technology and Retrospective Think-aloud Interviews,” *Int. J. Hosp. Tour. Adm.*, vol. 24, no. 5, pp. 720–752, 2023, doi: 10.1080/15256480.2022.2055690.

[22] N. Sangkaew and H. Zhu, “Understanding Tourists’ Experiences at Local Markets in Phuket: An Analysis of TripAdvisor Reviews,” *J. Qual. Assur. Hosp. Tour.*, vol. 23, no. 1, pp. 89–114, 2022, doi: 10.1080/1528008X.2020.1848747.

[23] G. M. Majid, I. Tussyadiah, Y. R. Kim, and A. Pal, "Intelligent automation for sustainable tourism: a systematic review," *J. Sustain. Tour.*, vol. 31, no. 11, pp. 2421–2440, 2023, doi: 10.1080/09669582.2023.2246681.

[24] A. Rydzik and C. S. Kissoon, "Decent work and tourism workers in the age of intelligent automation and digital surveillance," *J. Sustain. Tour.*, vol. 30, no. 12, pp. 2860–2877, 2022, doi: 10.1080/09669582.2021.1928680.

[25] X. Zhang and H. Jin, "How does smart technology, artificial intelligence, automation, robotics, and algorithms (STAARA) awareness affect hotel employees' career perceptions? A disruptive innovation theory perspective," *J. Hosp. Mark. Manag.*, vol. 32, no. 2, pp. 264–283, 2023, doi: 10.1080/19368623.2023.2166186.

[26] 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.

[27] 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.

[28] Y. A. Singgalen, "Analisis Perilaku Wisatawan Berdasarkan Data Ulasan di Website Tripadvisor Menggunakan CRISP-DM : Wisata Minat Khusus Pendakian Gunung Rinjani dan Gunung Bromo," *J. Comput. Syst. Informatics*, vol. 4, no. 2, pp. 326–338, 2023, doi: 10.47065/josyc.v4i2.3042.

[29] 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.

[30] 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.

[31] Y. A. Singgalen, "Analisis Sentimen Wisatawan terhadap Kualitas Layanan Hotel dan Resort di Lombok Menggunakan SERVQUAL dan CRISP-DM," *Build. Informatics, Technol. Sci.*, vol. 4, no. 4, pp. 1870–1882, 2023, doi: 10.47065/bits.v4i4.3199.

[32] 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.

[33] 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.

[34] 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.

[35] 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.

[36] M. Valeri and R. Baggio, "Italian tourism intermediaries: a social network analysis exploration," *Curr. Issues Tour.*, vol. 24, no. 9, pp. 1270–1283, 2021, doi: 10.1080/13683500.2020.1777950.

[37] L. J. Hsu, W. Lin, and H. P. Yueh, "Connecting Twitter With Scholarly Networks: Exploring HCI Scholars' Interactions From an SNA Approach," *IEEE Trans. Prof. Commun.*, vol. 65, no. 4, pp. 502–515, 2022, doi: 10.1109/TPC.2022.3205511.

[38] 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.

[39] C. Cahyaningtyas, Y. Nataliani, and I. R. Widiasari, "Analisis Sentimen Pada Rating Aplikasi Shopee Menggunakan Metode Decision Tree Berbasis SMOTE," *Aiti*, vol. 18, no. 2, pp. 173–184, 2021, doi: 10.24246/aiti.v18i2.173-184.

[40] A. Andreyestha and Q. N. Azizah, "Analisa Sentimen Kicauan Twitter Tokopedia Dengan Optimalisasi Data Tidak Seimbang Menggunakan Algoritma SMOTE," *Infotek J. Inform. dan Teknol.*, vol. 5, no. 1, pp. 108–116, 2022, doi: 10.29408/jit.v5i1.4581.

[41] V. A. Permadi, "Analisis Sentimen Menggunakan Algoritma Naive Bayes Terhadap Review Restoran di Singapura," *J. Buana Inform.*, vol. 11, no. 2, pp. 141–151, 2020, doi: 10.24002/jbi.v11i2.3769.

[42] F. Ariani, Amir, N. Alam, and K. Rizal, "Klasifikasi Penetapan Status Karyawan Dengan Menggunakan Metode Naive Bayes," *J. Paradig.*, vol. 20, no. 2, pp. 33–38, 2018, doi: 10.31294/p.v.

[43] H. A. R. Harpizon, R. Kurniawan, and ..., "Analisis Sentimen Komentar Di YouTube Tentang Ceramah Ustadz Abdul Somad Menggunakan Algoritma Naïve Bayes," *Anal. Sentimen …*, 2022, [Online]. Available: http://repository.uin-suska.ac.id/59746/

[44] F. Fitriana, E. Utami, and H. Al Fatta, "Analisis sentimen opini terhadap vaksin Covid-19 pada media sosial twitter menggunakan Support Vector Machine dan Naive Bayes," *J. Komtika (Komputasi …*, 2021, [Online]. Available: https://journal.unimma.ac.id/index.php/komtika/article/view/5185

[45] S. H. W. Chuah, E. C. X. Aw, and C. F. Cheng, "A silver lining in the COVID-19 cloud: examining customers' value perceptions, willingness to use and pay more for robotic restaurants," *J. Hosp. Mark. Manag.*, vol. 31, no. 1, pp. 49–76, 2022, doi: 10.1080/19368623.2021.1926038.

[46] F. Angelini, M. Castellani, and L. Vici, "Restaurant sector efficiency frontiers: a meta-analysis," *J. Foodserv. Bus. Res.*, vol. 27, no. 2, pp. 138–156, 2022, doi: 10.1080/15378020.2022.2077090.

[47] J. Y. Choe, J. J. Kim, and J. Hwang, "Innovative robotic restaurants in Korea: merging a technology acceptance model and theory of planned behaviour," *Asian J. Technol. Innov.*, vol. 30, no. 2, pp. 466–489, 2022, doi: 10.1080/19761597.2021.2005466.

[48] J. Hwang, H. Kim, J. J. Kim, and I. Kim, "Investigation of perceived risks and their outcome variables in the context of robotic restaurants," *J. Travel Tour. Mark.*, vol. 38, no. 3, pp. 263–281, 2021, doi: 10.1080/10548408.2021.1906826.