



Hotel Customer Segmentation for Marketing Strategy Optimization Using CRAF Framework

Yerik Afrianto Singgalen

Faculty of Business Administration and Communication, Tourism Study Program, Atma Jaya Catholic University of Indonesia, Jakarta

Jl. Jend. Sudirman No.51 5, RT.004/RW.4, Karet Semanggi, Kecamatan Setiabudi, Kota Jakarta Selatan, Daerah Khusus Ibukota Jakarta 12930, Indonesia

Email: yerik.afrianto@atmajaya.ac.id

Correspondence Author Email: yerik.afrianto@atmajaya.ac.id

Submitted: **01/05/2024**; Accepted: **29/06/2024**; Published: **30/06/2024**

Abstract—This research explores the implementation of the Customer Reviews and Analysis Framework (CRAF) as a crucial tool for optimizing marketing strategies in the hospitality industry. By conducting thorough data analysis and sentiment evaluation, CRAF provides valuable insights into guest preferences and behaviors, facilitating the creation of highly targeted marketing campaigns. A comparative study of SVM algorithms with and without SMOTE demonstrated the significance of data balancing techniques, with accuracies of 85.19% and 93.86%, respectively. Additionally, integrating Oracle Apex for data visualization and decision support enhances strategic planning and operational efficiency. The findings highlight that the combined use of advanced data analytics and sophisticated digital tools leads to improved customer satisfaction, refined marketing strategies, and sustained competitive advantage, contributing to the overall growth and success of the hospitality sector.

Keywords: Customer Segmentation; Marketing Strategy; Strategic Planning; Operational Efficiency; CRAF

1. INTRODUCTION

Market segmentation is a strategic approach to devising sales strategies that align precisely with consumer preferences. By categorizing the market into distinct segments based on various criteria such as demographics, psychographics, and behavioral traits, companies can tailor their offerings to meet specific needs, enhancing customer satisfaction and loyalty (Han, 2021; Suleri, Meijer, & Tarus, 2021). This targeted approach optimizes resource allocation and increases marketing efforts' efficiency, ensuring higher conversion rates (Deng, Shi, Chen, & Li, 2022; Whalen & Sisson, 2022). Implementing a well-defined segmentation strategy enables businesses to anticipate and respond to market dynamics more effectively, ultimately leading to a competitive advantage and sustained growth. Thus, the systematic application of market segmentation proves indispensable in achieving precise and effective sales strategies.

The hospitality industry leverages digital data for market segment analysis in accommodation services. Hotels can identify distinct customer segments and tailor their services to meet specific preferences and needs by systematically collecting and analyzing data from various digital sources (Ogada & Lindberg, 2022; Wen et al., 2023). This data-driven approach enhances the ability to predict customer behavior, optimize pricing strategies, and deliver personalized experiences, fostering customer loyalty and maximizing revenue (Tata, Sharrock, & Westerlanden, 2023; Whalen & Sisson, 2022). The strategic utilization of digital data in market segmentation empowers hotels to stay competitive and responsive in a dynamic market environment, underscoring its critical role in modern hospitality management.

Market segmentation within the hospitality industry can be achieved through various methodologies, including data mining. Utilizing data mining techniques allows hotels to analyze vast amounts of data to uncover patterns and insights about customer preferences and behaviors (Nwokorie, Igbojekwe, & Ukabuilu, 2024; Promnil, Srepirote, & Poopruksachat, 2024). This approach facilitates the identification of distinct market segments, enabling more precise targeting and personalized service offerings (Gazi et al., 2024; Yang, Zhu, Liao, & Wu, 2023). Employing data mining not only enhances the efficiency of marketing strategies but also drives competitive advantage by providing deeper customer insights (Huang & Tsaih, 2022; Moise, Gil-Saura, & Ruiz Molina, 2021). Therefore, integrating data mining in market segmentation is indispensable for optimizing hospitality services and achieving sustained business growth.

This study aims to analyze hotel market segments to produce recommendations for optimizing the marketing of products and services aligned with guest preferences. The research seeks to identify distinct customer groups and understand their unique needs and expectations by examining various demographic, psychographic, and behavioral data. The insights gained will enable hotels to tailor their marketing strategies and service offerings, enhancing customer satisfaction and loyalty. The findings underscore the importance of targeted marketing efforts in the hospitality industry, ultimately leading to improved business performance and competitive advantage. Therefore, this study is essential for developing effective marketing strategies that resonate with diverse guest segments.

The urgency of this research lies in its potential to address critical gaps in the current understanding of market segmentation within the hospitality industry. The study aims to provide actionable insights for enhancing service quality and competitive positioning by focusing on hotel guests' evolving preferences and behaviors. In an increasingly competitive market, the ability to tailor marketing strategies and service offerings to specific customer segments is not just advantageous but necessary for sustained success (Issaka, Bansah, & Kuuder, 2022; Ledi, Dumeda, Bandoma, & Ameza-Xemalordzo, 2024). Consequently, the timely execution of this research is pivotal in informing strategic decisions that drive growth and customer satisfaction in the hospitality sector.

Based on the identified research gap, designing a framework relevant to the data context and processing objectives is necessary. This study proposes the Customer Reviews and Analysis Framework (CRAF) as an approach tailored to the context of data and the goals of processing review data for market segmentation in the hospitality industry. The framework utilizes advanced text mining and sentiment analysis techniques to extract valuable insights from customer reviews, allowing for more accurate identification of market segments. This innovative approach addresses the limitations of traditional segmentation methods and enhances the ability to respond to dynamic customer preferences. Consequently, implementing CRAF is essential for improving market segmentation strategies and achieving better alignment with customer needs in the hospitality sector.

2.2 Customer Reviews and Analysis Framework (CRAF)

The Customer Reviews and Analysis Framework (CRAF) is tailored to the context of data and the goals of processing review data for market segmentation in the hospitality industry. Leveraging advanced text mining and sentiment analysis techniques, CRAF extracts and interprets customer sentiments and preferences from vast review data. This enables the identification of distinct market segments with greater accuracy and granularity. Such an approach enhances the precision of market segmentation and facilitates the development of more personalized and effective marketing strategies. Consequently, CRAF represents a significant advancement in the field, addressing the limitations of traditional methods and aligning segmentation practices with contemporary data-driven needs.

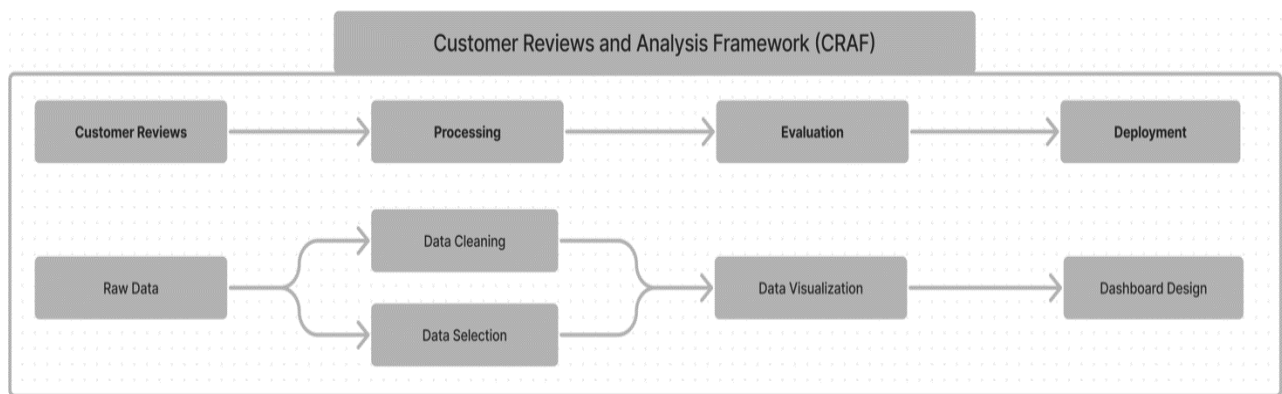


Figure 2. Customer Reviews and Analysis Framework (CRAF)

Figure 2 shows the implementation of CRAF. The Customer Reviews and Analysis Framework (CRAF) comprises the stages of customer reviews, processing, evaluation, and deployment. Initially, customer reviews are collected from various online platforms to form the raw data set. Advanced text mining and sentiment analysis techniques are employed during processing to extract meaningful insights and identify patterns. The evaluation stage involves assessing the quality and relevance of the processed data to ensure accurate market segmentation. Finally, the deployment stage applies the insights gained to develop targeted marketing strategies and improve service offerings. Consequently, CRAF provides a comprehensive and systematic approach to leveraging customer reviews for effective market segmentation in the hospitality industry.

Considering the flexibility of the Customer Reviews and Analysis Framework (CRAF) in customer segmentation and analysis, this research can generate contextual information and recommendations for optimizing hotel marketing strategies. The adaptable nature of CRAF allows it to process diverse and dynamic review data, ensuring that the insights are relevant and current. By analyzing customer sentiments and preferences effectively, the framework provides detailed segmentation crucial for developing targeted and efficient marketing plans. Therefore, the implementation of CRAF in this study promises to deliver actionable intelligence that enhances marketing efficacy and aligns closely with customer expectations.

2.2.1 Customer Reviews

At the customer review stage, it is essential to identify the brands providing the products and services being reviewed. This initial step involves gathering reviews from multiple online platforms, ensuring a comprehensive collection of customer feedback specific to each brand. Identifying the brands allows for a more targeted analysis, as the context of customer sentiments can vary significantly between different providers. This specificity enhances the accuracy and relevance of the insights derived from the reviews. Consequently, clarifying the brand context is crucial for producing detailed and actionable recommendations tailored to optimize the marketing strategies for each hotel brand.

In the context of this research, the data source for reviews is derived from the Agoda application, explicitly focusing on Gallery Prawirotaman Hotel. The total number of reviews collected is 3,076, but only 1,615 verified guest comments are displayed. Utilizing verified comments ensures the authenticity and reliability of the data, providing a more accurate representation of customer sentiments. This refined dataset is critical for thorough analysis, leading to more precise and actionable insights. Consequently, the emphasis on verified guest comments underscores the importance of data integrity in achieving reliable research outcomes.



Figure 4. Rating on Agoda and Booking.com

Figure 4 shows the rating of Gallery Prawirotaman Hotel on Agoda and Booking.com. Based on the rating calculations on the Agoda platform, several key performance indicators for Gallery Prawirotaman Hotel have been identified: Cleanliness received a score of 9.0, Service also scored 9.0, Value for money achieved 9.0, while Facilities, Location, and Room comfort and quality each scored 8.7. These ratings highlight the hotel's strengths in cleanliness, service quality, and value proposition, suggesting a high level of guest satisfaction in these areas. However, slightly lower scores in facilities, location, and room comfort indicate areas with room for improvement. Consequently, these insights can guide targeted enhancements to elevate the overall guest experience further and maintain competitive advantage.

Based on the rating calculations on the Booking.com platform, several performance metrics for Gallery Prawirotaman Hotel have been identified: Cleanliness scored 8.7, Facilities received 8.6, Location was rated at 8.7, Room comfort and quality also scored 8.7, Service achieved the highest rating at 9.1, and Value for money was rated at 8.7. These ratings illustrate a generally high level of satisfaction among guests, particularly in service, which stands out with the highest score. The slightly lower score in facilities suggests a potential area for improvement to enhance the guest experience further. Consequently, these insights provide valuable guidance for targeted enhancements in maintaining and elevating the hotel's overall quality.

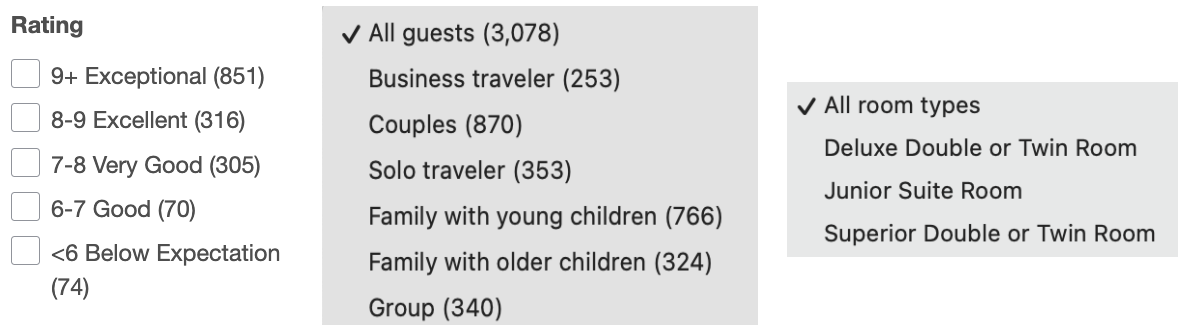


Figure 5. Rating Description, Guest, and Room Type

Figure 5 shows the platform's rating description, guest, and room type. The platform's rating description, guest, and room type data require thorough scraping to ensure they are cleaned, selected, and visualized effectively, thereby facilitating comprehension by decision-makers for optimizing marketing strategies. Scraping this data enables the extraction of detailed and relevant information, which can then be processed and organized to highlight key insights. This systematic approach allows for creating clear and informative visualizations that simplify complex data, making it accessible and actionable. Ultimately, such refined data presentation empowers policymakers to make informed decisions, enhancing the precision and effectiveness of marketing strategies.

Thus, the hotel's verified reviews can be further processed for marketing. This entails utilizing advanced data analysis techniques to extract valuable insights from the reviews, which can inform targeted marketing strategies. By systematically analyzing verified feedback, hotels can identify critical areas of satisfaction that need improvement, enabling a more personalized and practical approach to customer engagement. Consequently, leveraging verified review data enhances the precision of marketing efforts, ultimately leading to improved customer satisfaction and increased loyalty.

2.2.2 Processing

The processing stage consists of two main components: data cleansing and selection. Data cleansing involves removing inaccuracies, inconsistencies, and irrelevant information to ensure the dataset's integrity and reliability. Following this, data selection identifies and extracts the most relevant data points that align with the research objectives and marketing goals. These steps are critical for refining the dataset, enabling more accurate and insightful analysis. Consequently, a meticulous approach to data processing enhances the overall quality and applicability of the findings, supporting more effective decision-making in marketing strategies.

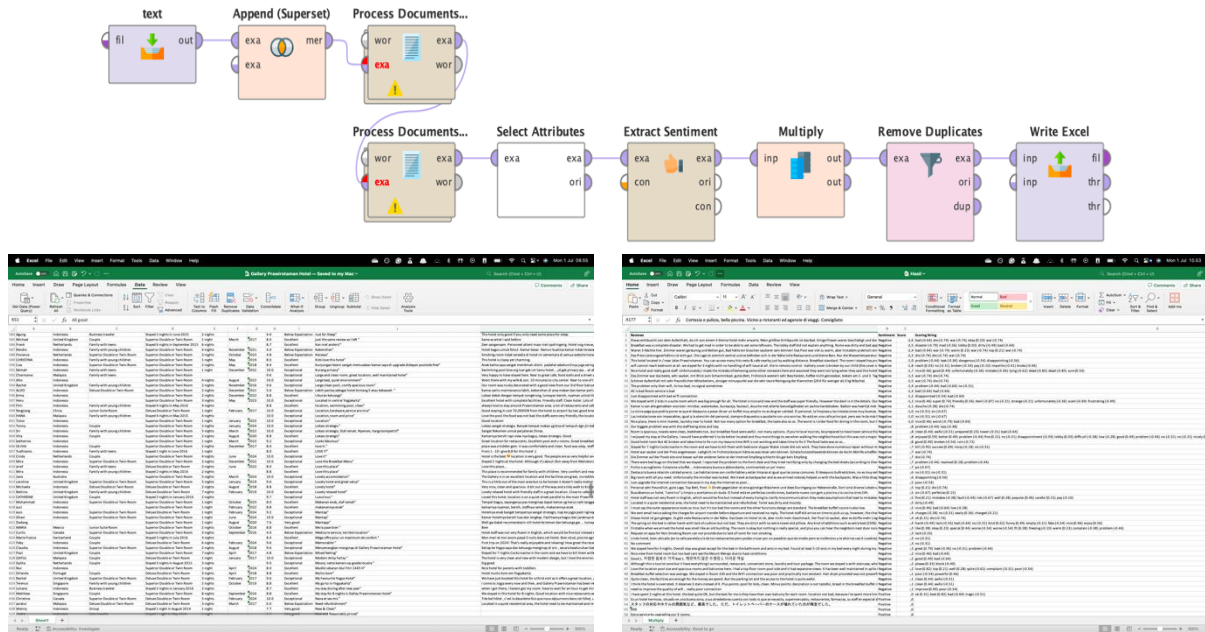


Figure 6. Data Cleaning and Extraction

Figure 6 shows the data-cleaning process. The data cleansing process employs several operators in Rapidminer, including tokenize (non-letters), transform cases (regular expression), filter tokens, filter stopwords, and stemming. Tokenization breaks down text into individual elements, excluding non-letter characters, while case transformation standardizes text using regular expressions. Filtering tokens and stopwords removes irrelevant words, ensuring that only meaningful data is retained. Stemming then reduces words to their base forms, facilitating more accurate analysis. This systematic approach to data cleansing enhances the quality and reliability of the dataset, laying a solid foundation for subsequent analytical processes and ensuring the validity of the research outcomes.

At the data selection stage, the text in the review column is filtered based on language to classify sentiments as positive or negative. This involves first identifying and isolating reviews written in the target language to ensure consistency in analysis. Subsequently, sentiment analysis algorithms are applied to categorize the reviews, distinguishing between positive and negative feedback. This classification provides a clear understanding of customer sentiments, which is crucial for developing targeted marketing strategies. Effective data selection based on language and sentiment ultimately enhances the relevance and accuracy of the analytical insights.

2.2.3 Evaluation

At the evaluation stage, the algorithms used in the review data classification process are assessed based on their accuracy, precision, recall, F-measure, and AUC values. Accuracy measures the overall correctness of the classification, while precision quantifies the proportion of actual positive results among all optimistic predictions. Recall, or sensitivity, evaluates the algorithm's ability to identify all relevant instances, and the F-measure provides a balance between precision and recall. The AUC, or Area Under the Curve, represents the algorithm's capability to distinguish between positive and negative classes. This comprehensive evaluation ensures that the selected algorithm performs optimally, enhancing the reliability and effectiveness of the sentiment analysis.

The significance of accuracy, precision, recall, F-measure, and AUC values in hotel guest sentiment classification is paramount for evaluating the effectiveness of sentiment analysis models. Accuracy reflects the overall correctness of the model by measuring the proportion of correctly classified sentiments. Precision quantifies the proportion of true positive sentiments among all optimistic predictions and highlights the model's reliability in identifying positive feedback. Recall, or sensitivity, evaluates the model's ability to capture all relevant positive sentiments, ensuring comprehensive sentiment detection. The F-measure balances precision and recall, providing a metric that reflects the model's overall performance. The AUC, or Area Under the Curve, assesses the model's capacity to distinguish between positive and

negative sentiments. Collectively, these metrics provide a robust evaluation framework, ensuring the sentiment analysis model is both accurate and reliable in classifying hotel guest feedback.

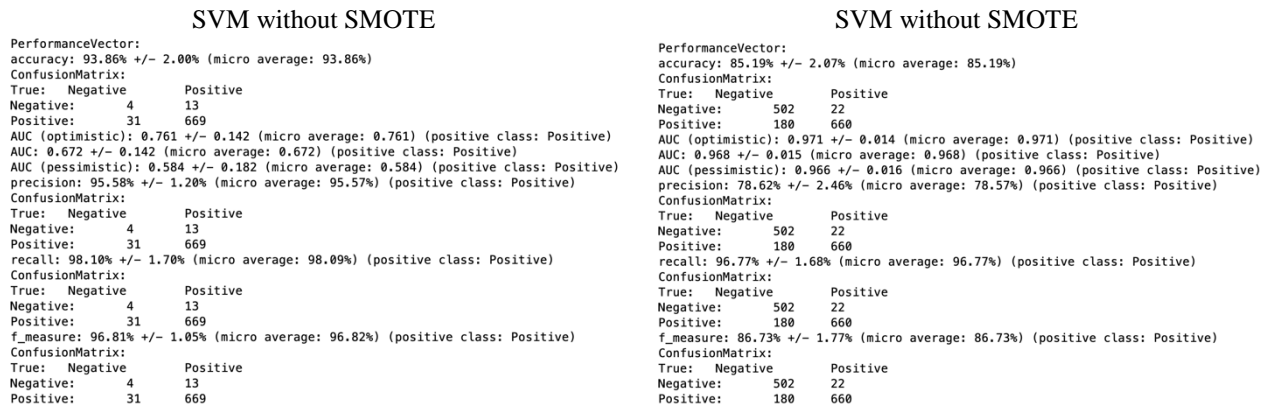


Figure 7. Evaluation of SVM Performance in Sentiment Classification

Figure 7 shows the evaluation of SVM in sentiment classification. The SMOTE operator is employed to address data imbalance issues. By generating synthetic examples, SMOTE enhances the representation of minority classes in the dataset, thereby mitigating the skewed distribution problem. This technique ensures that the classification algorithm is not biased towards the majority class, leading to more balanced and accurate predictive performance. Consequently, applying SMOTE is crucial for improving the reliability and validity of data analysis, particularly in scenarios where class imbalance can significantly affect model outcomes.

Therefore, the performance of algorithms with and without using SMOTE can be analyzed for comparative purposes. This comparative analysis involves evaluating key metrics such as accuracy, precision, recall, F-measure, and AUC to determine the impact of SMOTE on classification outcomes. By assessing these metrics, it becomes evident whether using SMOTE enhances the model's ability to classify minority class instances correctly. Such analysis is essential for validating the effectiveness of SMOTE in addressing data imbalance, ensuring that the selected algorithm provides the most reliable and accurate predictions. Consequently, this comparison aids in selecting the optimal approach for handling imbalanced datasets in sentiment classification tasks.

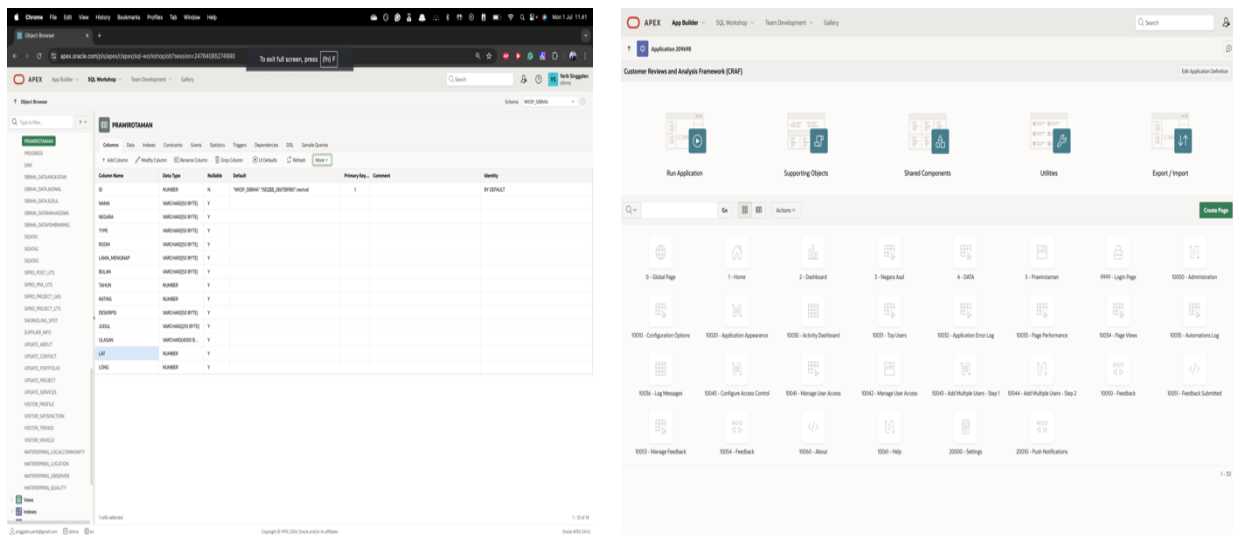


Figure 8. Data Visualization in Oracle Apex

Figure 8 shows the data visualization in Oracle Apex. The selected data is visualized using Oracle Apex through dashboards, maps, and reports. Specifically, the visualizations encompass data on the country of origin, guest type, room type, and length of stay. These visual representations enable a comprehensive analysis of guest demographics and behaviors, facilitating the identification of trends and patterns. Oracle Apex ensures that the data is presented in an interactive and user-friendly manner, enhancing the accessibility of critical insights for decision-makers. Consequently, these visual tools are instrumental in optimizing marketing strategies and improving service delivery in the hospitality industry.

Subsequently, the data visualizations can be utilized to analyze market segments within the hospitality industry. By examining the visualized data, insights can be derived into guest preferences and behaviors across different segments, such as demographic trends and booking patterns. These analyses enable hoteliers to identify and target specific market segments more effectively, tailoring their marketing strategies to meet the unique needs of each segment. Ultimately,

leveraging data visualizations for market analysis facilitates more informed decision-making, thereby enhancing the competitiveness and profitability of hotel operations.

2.2.4 Deployment

Data clustering and sentiment analysis results can be applied at the deployment stage to optimize the marketing strategies for hotel products and services. By leveraging these analytical outcomes, hotels can tailor their marketing efforts to better align with the identified customer segments and sentiment trends. This approach ensures that promotional campaigns and service enhancements are precisely targeted, maximizing their effectiveness and customer appeal. Consequently, integrating these analytical insights into marketing strategies significantly enhances customer satisfaction and loyalty, ultimately driving business growth and competitive advantage in the hospitality industry.

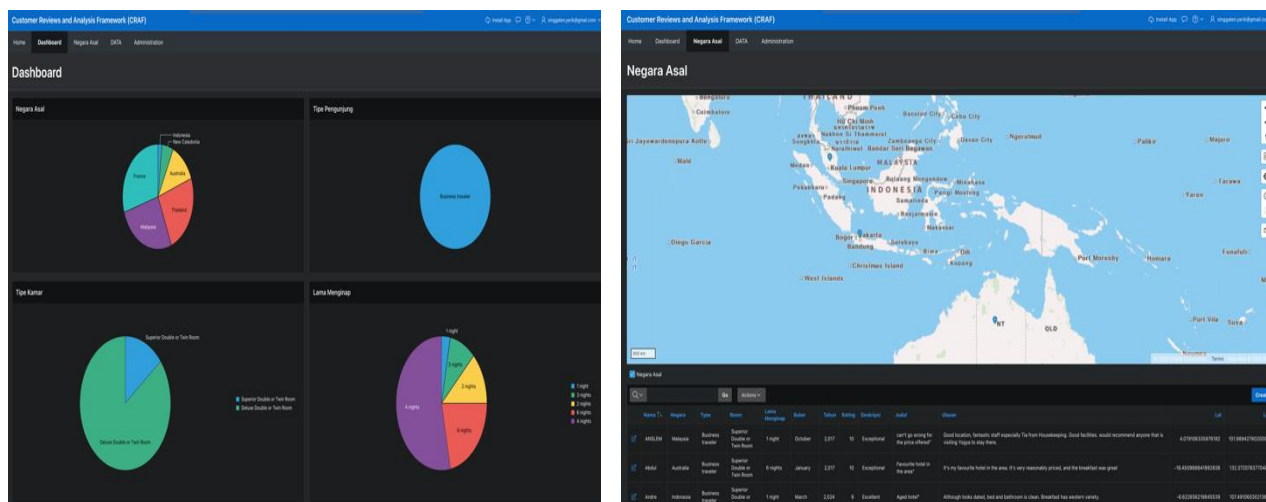


Figure 9. Data Integration in Oracle Apex

Figure 9 shows the data integration in Oracle Apex. Data visualization using Oracle Apex facilitates users in adding, modifying, updating, and deleting data from the system. This platform provides an intuitive interface that simplifies data management, ensuring users can efficiently handle data operations without extensive technical expertise. The ability to seamlessly update and maintain data enhances the accuracy and relevance of the visualizations, thereby supporting more informed decision-making. Consequently, Oracle Apex's robust data management capabilities significantly streamline data handling processes, contributing to the overall effectiveness of data-driven strategies in various applications.

Thus, this research extends beyond merely evaluating the results of market segmentation analysis in the hospitality industry by enabling visualization through an interactive system. This approach allows stakeholders to engage with the data more dynamically, facilitating a deeper understanding of the market segments and their characteristics. Using an interactive system enhances the accessibility and usability of the analytical outcomes, making it easier for decision-makers to interpret and act on the insights. Consequently, integrating interactive visualizations enriches the analysis and empowers more effective and informed strategic planning.

3. RESULTS AND DISCUSSION

The discussion in this research is divided into two sections: hotel customer segmentation based on the context of Gallery Prawirotaman Hotel, followed by the optimization of accommodation service marketing strategies utilizing the Customer Reviews and Analysis Framework (CRAF). The first section delves into customer segmentation, analyzing specific demographic and behavioral data from Gallery Prawirotaman Hotel to identify distinct market segments. The second section leverages insights from CRAF to enhance marketing strategies, ensuring they are precisely tailored to the identified segments and their preferences. This comprehensive approach provides a detailed understanding of market dynamics and actionable strategies for improving marketing effectiveness. Consequently, the research contributes value to market analysis and strategic marketing optimization in the hospitality industry.

3.1 Hotel Customer Segmentation Based on Type of Guest and Room

Hotel customer segmentation based on the type of guest and room reveals significant insights into guest preferences and behaviors. Analysis indicates that Superior Double or Twin Rooms are most popular among couples and families with young children, with 80 couples and 64 families recorded, respectively. Business travelers and solo travelers also show a notable preference for this room type, albeit to a lesser extent. Understanding these segments allows for developing tailored marketing strategies that cater to dominant guest types, enhancing customer satisfaction and optimizing room

occupancy. Therefore, effective segmentation by guest type and room preference is essential for strategic decision-making in the hospitality industry.

Based on the identification and analysis of market segments by guest type, the following distribution for the Deluxe Double or Twin Room has been observed: 8 business travelers, 47 couples, 13 families with teens, 22 families with young children, 12 groups, and 20 solo travelers, with 154 entries unclassified. These figures indicate a predominance of couples among the guests, suggesting a strong preference for this room type within this segment. Additionally, the significant number of unclassified entries highlights the need for improved data collection methods. Consequently, understanding these segment distributions enables more tailored and effective marketing strategies to enhance guest satisfaction and optimize occupancy rates.

Table 1. Type of Guest and Room

Type of Guest	Deluxe Double or Twin Room	Junior Suite Room	Room Assigned on Arrival	Superior Double or Twin Room	(blank)	Grand Total
Business traveler	8			18	25	51
Couple	47	2	1	80	84	214
Family with teens	13			13	21	47
Family with young children	22	3		64	53	142
Group	12	1		21	30	64
Solo traveler	20			31	30	81
(blank)	154	1		315	26	496
Grand Total	276	7	1	542	269	1095

Table 1 shows the type of guest and room based on the data mining result of the Agoda platform. Based on the identification and analysis of market segments by guest type, the following distribution for the Superior Double or Twin Room has been observed: 18 business travelers, 80 couples, 13 families with teens, 64 families with young children, 21 groups, 31 solo travelers, and 315 entries unclassified. These figures indicate that couples and families with young children are the predominant segments for this room type, highlighting their strong preference for these accommodations. Additionally, the substantial number of unclassified entries points to a need for enhanced data categorization processes. Consequently, these insights are crucial for developing targeted marketing strategies to cater to the specific needs of these segments, thereby optimizing room occupancy and guest satisfaction.

Table 2. Type of Guest and Length of Stay

Type of Guest	Blank	Length of Stay										Grand Total			
		1	15	2	20	3	4	5	6	7	8		9		
Business traveler		22		16		7	2	2	2						51
Couple	3	47		73	1	50	27	4	4	4	1				214
Family with teens	1	14		13		9	6	2	2						47
Family with young children		39		58		33	9	2		1					142
Group		18	1	13		24	4	2	1	1					64
Solo traveler	2	19		33		14	8	3	1	1					81
(blank)		135		164		115	51	22	6	2			1		496
Grand Total	6	294	1	370	1	252	107	37	16	9	1	1			1095

Table 2 shows the type of guest and length of stay. Based on the type of guest, the length of stay is as follows: Business travelers have 51 stays, with most staying for one day (22 stays) and three days (16 stays). Couples recorded 214 stays, with most staying for three days (73 stays) and five (50 stays). Families with teens have 47 stays, primarily for three days (14 stays) and five days (13 stays). Families with young children recorded 142 stays, mainly for three days (58 stays) and five (33 stays). Groups have 64 stays, with most staying for three days (18 stays) and five (24 stays). Solo travelers recorded 81 stays, predominantly for one day (19 stays) and three days (33 stays). There are also 496 unclassified stays, with the highest numbers staying for three days (164 stays) and five days (115 stays). This detailed segmentation by guest type and length of stay provides valuable insights for optimizing hotel marketing and service strategies.

Based on the identification of visitor types by country of origin, it is evident that a significant proportion of guests originate from Indonesia, accounting for 57% of the total. Malaysia follows with 12%, while the remaining top countries include Japan and the United States, each contributing 4%. Other notable countries such as Australia, France, Germany, the Netherlands, Singapore, and the United Kingdom each represent 3-6% of the visitor demographics. This distribution highlights Indonesia as the predominant source of guests, emphasizing the importance of focusing marketing efforts

domestically. Consequently, these insights facilitate more targeted and effective marketing strategies to cater to the diverse international visitor base while prioritizing the significant domestic market.

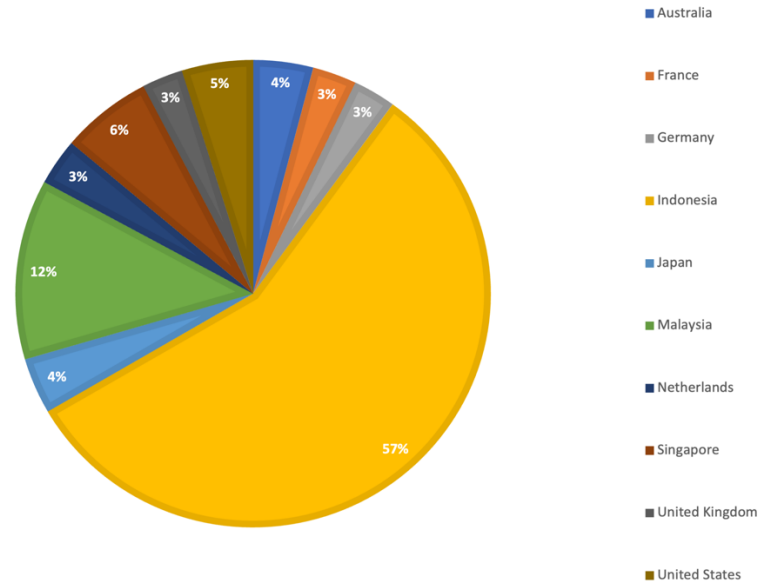


Figure 10. Top 10 Visitors based on Country of Origin

Figure 10 shows the top 10 visitors based on country of origin. Based on the data, an effective sales strategy can be formulated by aligning guest preferences with their country of origin, type of guest, room type, and length of stay. The high proportion of domestic visitors suggests that marketing efforts should prioritize local promotional campaigns. Additionally, understanding that couples and families with young children prefer Superior Double or Twin Rooms allows for targeted promotions and package deals tailored to these segments. The hotel can offer customized packages that encourage longer stays or repeat visits by analyzing the typical length of stay. Consequently, such a data-driven approach ensures that marketing strategies are relevant and appealing, ultimately enhancing guest satisfaction and maximizing revenue.

Thus, marketing strategies can be tailored based on the stay history of previous guests. By analyzing data on past stays, including duration and preferences, hotels can identify trends and patterns that inform targeted promotional efforts. This approach allows for the creation of personalized offers that cater to different guest segments' specific needs and expectations. Implementing such data-driven strategies enhances guest satisfaction by providing more relevant and appealing options, maximizing the potential for repeat business and long-term loyalty. Consequently, leveraging historical guest data is crucial for optimizing marketing effectiveness and achieving sustained growth in the hospitality industry.

3.2 Optimization of Marketing Strategy

Optimization of the marketing strategy for Gallery Prawirotaman Hotel can be achieved by implementing the Customer Reviews and Analysis Framework (CRAF). This framework enables the extraction of valuable insights from customer reviews, identifying key areas of satisfaction and improvement. By leveraging these insights, the hotel can tailor its marketing campaigns to meet guest preferences and expectations better. Applying CRAF ensures that marketing efforts are data-driven and highly targeted, thereby enhancing customer engagement and loyalty. Consequently, utilizing CRAF is essential for refining marketing strategies and achieving a competitive edge in the hospitality industry.

Based on the sentiment analysis of review data, it is evident that guests provide positive feedback on the products and services. Optimized using SMOTE, the SVM algorithm demonstrated the best performance in sentiment classification with the following performance vector: accuracy of 85.19% with a margin of error of +/- 2.07%. The confusion matrix reveals 502 true negatives and 22 false negatives, alongside 180 false positives and 660 true positives. The AUC values are notably high, with an optimistic AUC of 0.971, a regular AUC of 0.968, and a pessimistic AUC of 0.966. Precision stands at 78.62% with a margin of error of +/- 2.46%, while recall is 96.77% with a margin of error of +/- 1.68%, and the F-measure is 86.73% with a margin of error of +/- 1.77%. These results indicate that the algorithm effectively classifies sentiments, providing a reliable basis for understanding guest opinions and enhancing service quality.

Similarly, the performance of the SVM algorithm without SMOTE also demonstrates strong results: an accuracy of 93.86% with a margin of error of +/- 2.00%. The confusion matrix shows four true negatives, 13 false negatives, 31 false positives, and 669 true positives. The AUC values are optimistic AUC at 0.761, regular AUC at 0.672, and pessimistic AUC at 0.584. Precision is recorded at 95.58% with a margin of error of +/- 1.20%, recall at 98.10% with a margin of error of +/- 1.70%, and the F-measure at 96.81% with a margin of error of +/- 1.05%. These metrics indicate that the SVM algorithm, even without using SMOTE, performs effectively in classifying sentiments, providing reliable insights for understanding guest feedback and enhancing service quality.

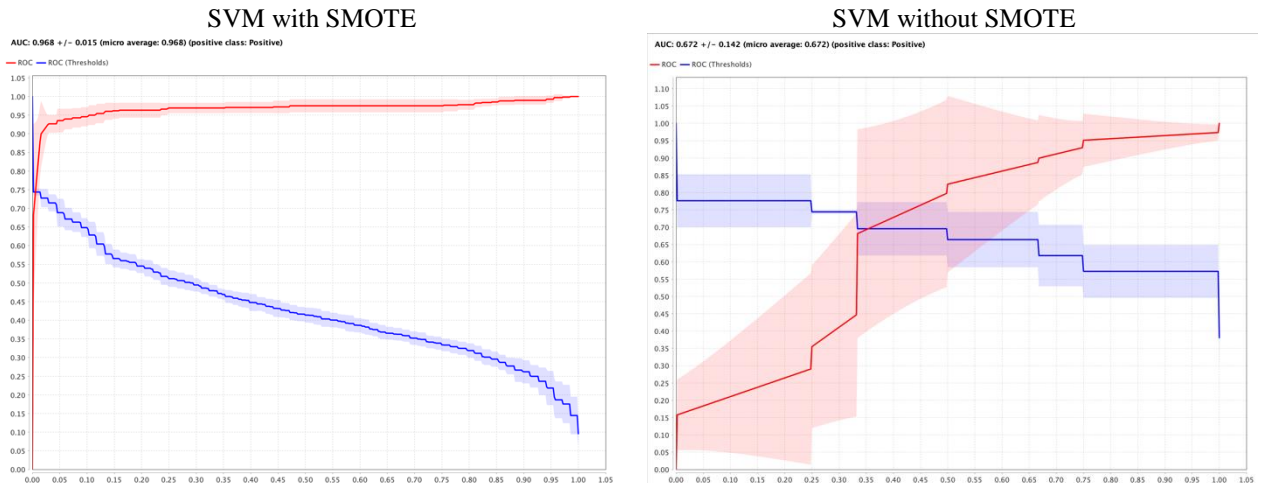


Figure 11. Area Under Curve of SVM

Figure 11 shows the AUC of SVM in sentiment classification. Evaluating the AUC values for the SVM algorithm with and without SMOTE reveals significant insights into sentiment classification performance. With SMOTE, the optimistic, regular, and pessimistic AUC values are 0.971, 0.968, and 0.966, respectively, indicating a high capability of the model to distinguish between positive and negative sentiments. Without SMOTE, the AUC values are notably lower, with the optimistic, regular, and pessimistic AUCs at 0.761, 0.672, and 0.584, respectively. This comparison highlights that using SMOTE substantially enhances the algorithm's performance in identifying sentiment polarity. Consequently, incorporating SMOTE into the sentiment classification process improves the model's accuracy and reliability, making it a crucial step for achieving optimal results in sentiment analysis.

After understanding the sentiments of hotel guests, this information can be correlated with visitor data related to the month and year of stay to analyze the optimal seasons for pricing adjustments and promotional package offers. By examining the patterns of guest stays alongside their feedback, hotels can identify peak seasons and periods of higher demand. This analysis allows for strategic pricing and tailored promotional offers to maximize occupancy and revenue. Consequently, leveraging sentiment data with seasonal trends provides a robust framework for optimizing marketing and pricing strategies, ensuring enhanced guest satisfaction and improved financial performance.

Table 3. Month and Year of Stay

Month	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	Grand Total
April		2	12	7	8	1	2	4	13	11	60
August		13	10	4	5	7	1	11	4		55
December	1	15	12	10	9	5	11	31	20		114
February		1	16	7	5	6		2	7	32	76
January			19	5	7		3	6	17	25	82
July		7	7	1	3	3		13	10		44
June		3	7	9	6	1	2	11	8	6	53
March		3	17	17	9	1	5	4	6	12	74
May		3	9	7	4		4	14	7	15	63
November	1	10	10	4	2	4	4	8	13		56
October		15	12	5	6	3	8	8	11		68
September		27	15	5	5	1	3	6	9		71
Grand Total	2	99	146	81	69	32	43	118	125	101	816

Table 3 shows the visitor data related to the month and year of visit. Based on the month and year of stay data, the hotel can optimize its products and services by preparing promotional pricing tailored to room types and guest profiles to boost sales during typically low occupancy periods. Analysis of the data reveals that December has the highest number of stays (114), indicating a peak period, while months like July (44) and August (55) show lower occupancy. Specific promotional strategies can be targeted for these slower months to attract more guests. Additionally, understanding the distribution of stays across different years helps identify trends and adjust marketing strategies accordingly. Consequently, leveraging this data for targeted promotions can enhance occupancy rates and overall sales performance during off-peak periods.

The complexity of guest data used as a reference for marketing strategies of hotel products and services necessitates a digital system that facilitates decision-making. Therefore, optimizing the visualization results can be achieved by designing a report system based on Oracle Apex. Such a system allows for efficient data management and real-time insights, enabling hotel management to make informed decisions quickly. By leveraging Oracle Apex, data can be

presented in an interactive and user-friendly manner, enhancing the accuracy and effectiveness of marketing strategies. Consequently, implementing this system is crucial for improving the hospitality industry's strategic planning and operational efficiency.

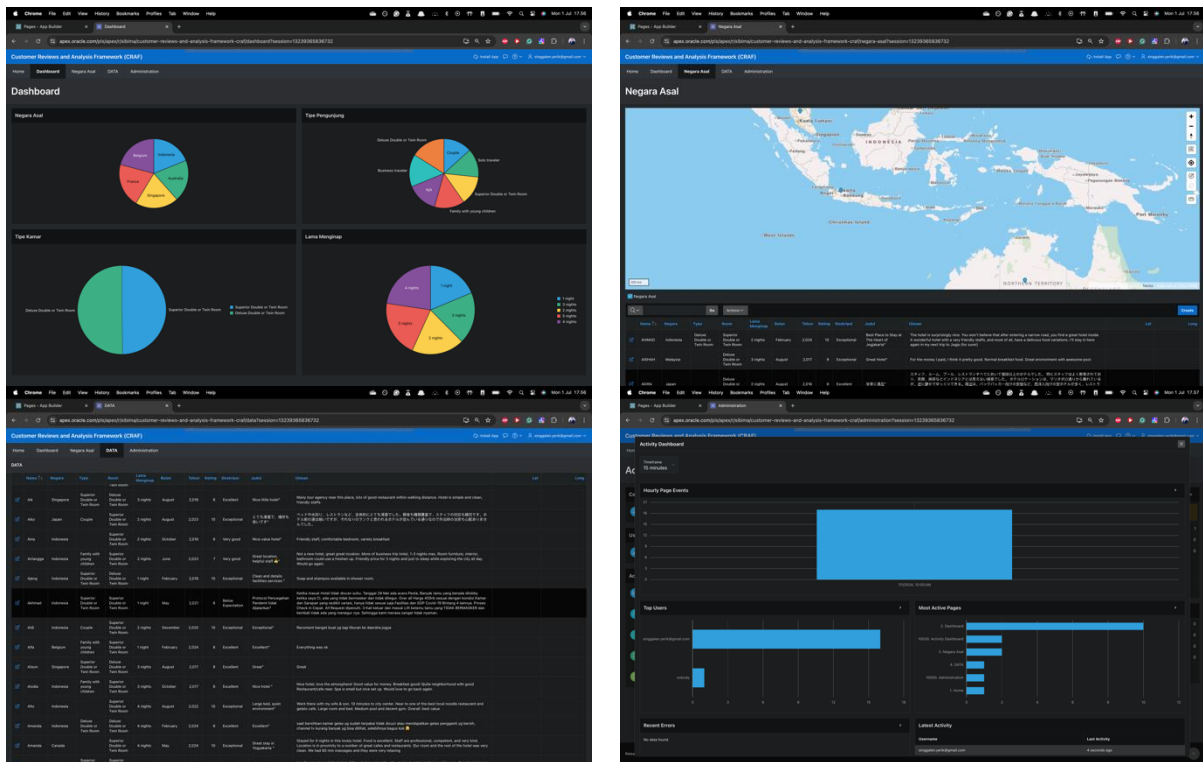


Figure 12. Information System Design for Hotel Customer Segmentation and Analysis

Figure 12 shows the information system design for hotel customer segmentation and analysis. An information system for hotel customer segmentation and analysis is designed using Oracle Apex to facilitate decision-makers in establishing marketing programs for hotel products and services. This system provides an integrated platform for collecting, analyzing, and visualizing customer data, enabling a comprehensive understanding of market segments. The system uses Oracle Apex to ensure data accuracy and real-time updates, crucial for making informed marketing decisions. Consequently, implementing this information system significantly enhances the efficiency and effectiveness of marketing strategies, leading to improved customer satisfaction and increased revenue for the hotel.

Thus, the Customer Reviews and Analysis Framework (CRAF) emerges as a solution for optimizing marketing strategies in the hospitality industry. By leveraging advanced data analytics, CRAF enables the extraction of valuable insights from customer reviews, which are crucial for understanding guest preferences and behavior. This comprehensive analysis allows hotels to tailor their marketing efforts more precisely, ensuring targeted and effective promotional campaigns. Consequently, adopting CRAF enhances marketing strategies' overall efficiency and impact, increasing customer satisfaction and improving business performance.

3.3 Discussion

Marketing strategies can be established by analyzing sales history and hotel guest behavior data. By examining past sales data, trends, and patterns can be identified, providing insights into peak seasons, popular room types, and successful promotions (Kim & Velthuis, 2021; Makanyeza, Svtwa, & Jaiyeoba, 2021). Additionally, understanding guest behavior, such as booking preferences and feedback, allows for creating personalized marketing campaigns that cater to specific guest segments (Thirumaran, Mohammadi, Azzali, Eijdenberg, & Donough-Tan, 2023). This data-driven approach ensures targeted and effective marketing efforts, enhancing guest satisfaction and optimizing revenue. Consequently, leveraging historical sales and behavioral data is essential for formulating strategic marketing initiatives in the hospitality industry.

Market segmentation is essential for identifying the purchasing behavior of hotel guests to enhance sales. By categorizing guests into segments based on demographics, booking patterns, and preferences, hotels can gain deeper insights into each group's specific needs and behaviors (Singh, Mir, & Nazki, 2024). This segmentation allows for developing tailored marketing strategies that resonate more effectively with each target audience (Mariani & Borghi, 2020). Such personalized approaches improve customer satisfaction and drive higher conversion rates and revenue (Wu, Wang, Xia, Li, & Pan, 2024). Consequently, understanding and leveraging market segments is crucial for optimizing sales strategies in the hospitality industry.

The research findings underscore the importance of segmentation data derived from guest reviews for enhancing the quality of hotel products and services. By analyzing detailed feedback, hotels identify areas of strength and opportunities for improvement (Martel–Escobar, González-Martel, & Vázquez-Polo, 2023). This valuable insight allows for the customization of offerings to meet guest expectations better, ultimately leading to higher satisfaction and loyalty (Moliner-Velázquez, Fuentes-Blasco, & Gil-Saura, 2023). Implementing changes based on this data-driven approach ensures that the enhancements are relevant and impactful. Consequently, utilizing guest review segmentation is pivotal for continuous improvement and maintaining a competitive edge in the hospitality industry.

The limitations of this research are noteworthy and merit consideration. Firstly, the data analyzed is confined to reviews from a specific hotel, which may not fully represent the broader hospitality industry. Additionally, the sentiment analysis relies heavily on the algorithms' accuracy, which, despite high-performance metrics, may still encounter challenges with nuanced language and context. Another constraint is excluding external factors such as economic conditions and competitive actions that could influence guest behavior and perceptions. Consequently, while the findings provide valuable insights, they should be interpreted with an understanding of these inherent limitations.

Recommendations for further research include expanding the scope of data collection to encompass multiple hotels across various regions to enhance the generalizability of the findings. Additionally, incorporating advanced natural language processing techniques can improve the accuracy of sentiment analysis by capturing more nuanced aspects of guest feedback. Exploring the impact of external factors, such as economic trends and competitor strategies, on guest behavior can provide a more comprehensive understanding of the market dynamics. Consequently, these recommendations aim to address current limitations and contribute to a more robust and detailed analysis of customer segmentation and sentiment in the hospitality industry.

4. CONCLUSION

In conclusion, implementing the Customer Reviews and Analysis Framework (CRAF) is a pivotal advancement for the hospitality industry, particularly optimizing marketing strategies. CRAF provides deep insights into guest preferences and behaviors through meticulous data analysis and sentiment evaluation. This enables hotels to craft highly targeted marketing campaigns, aligning their services and promotional efforts with their customers' needs and desires. The comparative analysis of SVM algorithms with and without SMOTE highlights the importance of data balancing techniques in enhancing classification accuracy and reliability, with the former achieving an accuracy of 85.19% and the latter 93.86%. Furthermore, integrating Oracle Apex for data visualization and decision support underscores the value of sophisticated digital tools in strategic planning. Ultimately, these methodologies collectively contribute to refined marketing strategies, improved customer satisfaction, and enhanced operational efficiency, driving competitive advantage and sustainable growth in the hospitality sector.

ACKNOWLEDGMENT

Thanks to the Tourism Study Program, Faculty of Business Administration and Communication, Atma Jaya Catholic University of Indonesia, LPPM, and Center of Digital Transformation and Tourism Development for the support and invaluable contributions to this research publication.

REFERENCES

- Claudio-Quiroga, G., de Villanueva, C. F., & Gil-Alana, L. A. (2024). Tourism and covid-19 in three European cities. Change in persistence. *Cogent Social Sciences*, 10(1). <https://doi.org/10.1080/23311886.2024.2358159>
- de Vries, M. (2022). What can hotels learn content wise from online travel agency websites? *Research in Hospitality Management*, 12(3), 294–298. <https://doi.org/10.1080/22243534.2023.2202489>
- Deng, M., Shi, W., Chen, H., & Li, X. (2022). The Effects of Power on Consumer Decision Strategies: The Mediating Role of Behavioral Approach-Inhibition Tendency and the Moderating Role of Gender. *Psychology Research and Behavior Management*, 15, 3931–3953. <https://doi.org/10.2147/PRBM.S386917>
- Fazal-e-Hasan, S. M., Mortimer, G., Ahmadi, H., Adil, M., & Sadiq, M. (2024). Examining the impact of tourists' hope, knowledge and perceived value on online hotel booking intentions. *Asia Pacific Journal of Tourism Research*, 29(6), 719–735. <https://doi.org/10.1080/10941665.2024.2343058>
- Gazi, M. A. I., Islam, M. A., Masud, A. Al, Senathirajah, A. R. bin S., Biswas, S., & Shuvro, R. A. (2024). The moderating impacts of COVID-19 fear on hotel service quality and tourist satisfaction: Evidence from a developing country. *Cogent Social Sciences*, 10(1). <https://doi.org/10.1080/23311886.2024.2331079>
- Han, H. (2021). Consumer behavior and environmental sustainability in tourism and hospitality: a review of theories, concepts, and latest research. *Journal of Sustainable Tourism*, 29(7), 1021–1042. <https://doi.org/10.1080/09669582.2021.1903019>
- Ho, R. C., Withanage, M. S., & Khong, K. W. (2020). Sentiment drivers of hotel customers: a hybrid approach using unstructured data from online reviews. *Asia-Pacific Journal of Business Administration*, 12(3/4), 237–250. <https://doi.org/10.1108/APJBA-09-2019-0192>
- Huang, H. Te, & Tsaih, L. S. J. (2022). Prioritizing hotel lobby design factors: perspectives of hotel operators in China. *Journal of Asian Architecture and Building Engineering*, 21(5), 1801–1813. <https://doi.org/10.1080/13467581.2021.1966016>
- Issaka, R. W., Bansah, A. K., & Kuuder, C.-J. W. (2022). Hotel managers' perspectives of Airbnb operations in Accra Metropolis, Ghana. *Research in Hospitality Management*, 12(1), 35–43. <https://doi.org/10.1080/22243534.2022.2080944>

- Kim, B., & Velthuis, O. (2021). From reactivity to reputation management: online consumer review systems in the restaurant industry. *Journal of Cultural Economy*, 14(6), 675–693. <https://doi.org/10.1080/17530350.2021.1895280>
- Ledi, K. K., Dumeda, R. B., Bandoma, S., & Ameza-Xemalordzo, E. (2024). Strategic thinking and innovation as panaceas for hotel performance in turbulent business environments. *Cogent Business and Management*, 11(1). <https://doi.org/10.1080/23311975.2024.2328330>
- Liu-Lastres, B., Guo, Y., & Liu, H. (2024). Hotel crisis communication on social media: effects of message appeal. *Anatolia*, 35(1), 56–66. <https://doi.org/10.1080/13032917.2022.2119590>
- Makanyeza, C., Svtowa, T. D., & Jaiyeoba, O. (2021). The effect of consumer rights awareness on attitude and purchase intention in the hotel industry: Moderating role of demographic characteristics. *Cogent Business and Management*, 8(1). <https://doi.org/10.1080/23311975.2021.1898301>
- Mariani, M., & Borghi, M. (2020). Environmental discourse in hotel online reviews: a big data analysis. *Journal of Sustainable Tourism*, 29(5), 829–848. <https://doi.org/10.1080/09669582.2020.1858303>
- Martel-Escobar, M., González-Martel, C., & Vázquez-Polo, F. J. (2023). Managing score heterogeneity between online consumer review websites. *Cogent Social Sciences*, 9(2). <https://doi.org/10.1080/23311886.2023.2267261>
- Moise, M. S., Gil-Saura, I., & Ruiz Molina, M. E. (2021). The importance of green practices for hotel guests: does gender matter? *Economic Research-Ekonomska Istrazivanja*, 34(1), 3508–3529. <https://doi.org/10.1080/1331677X.2021.1875863>
- Moliner-Velázquez, B., Fuentes-Blasco, M., & Gil-Saura, I. (2023). Effect of motivations and engagement with eWOM on hotel queries. *Economic Research-Ekonomska Istrazivanja*, 36(3). <https://doi.org/10.1080/1331677X.2023.2218475>
- Nwokorie, E. C., Igbojekwe, P. A., & Ukabuili, E. N. (2024). Customer intelligence and customer retention in hotels in south-west Nigeria. *Research in Hospitality Management*, 14(1), 43–56. <https://doi.org/10.1080/22243534.2024.2354664>
- Ogada, P., & Lindberg, F. (2022). Immigration blues: understanding market dynamics through consumer acculturation. *Consumption Markets and Culture*, 25(6), 546–566. <https://doi.org/10.1080/10253866.2022.2107512>
- Promnil, N., Srepirote, K., & Poopruksachat, T. (2024). Exploratory factor analysis of hotel employee competencies in Thailand: the evidence after the COVID-19 pandemic. *Cogent Business and Management*, 11(1). <https://doi.org/10.1080/23311975.2024.2364850>
- Singh, R., Mir, M. A., & Nazki, A. A. (2024). Evaluation of tourist behavior towards traditional food consumption: validation of extended Theory of Planned Behaviour. *Cogent Social Sciences*, 10(1). <https://doi.org/10.1080/23311886.2023.2298893>
- Suleri, J., Meijer, R., & Tarus, E. (2021). Exploring hotel identity by focusing on customer experience analysis. *Research in Hospitality Management*, 11(2), 113–120. <https://doi.org/10.1080/22243534.2021.1917178>
- Tata, E., Sharrock, M., & Westerlaken, R. (2023). Generation Z consumer behaviour and hotel branding: exploring the role of values, corporate identity and trust. *Research in Hospitality Management*, 13(1), 63–68. <https://doi.org/10.1080/22243534.2023.2239585>
- Thirumaran, K., Mohammadi, Z., Azzali, S., Eijdenberg, E. L., & Donough-Tan, G. (2023). Transformed landscapes, tourist sentiments: the place making narrative of a luxury heritage hotel in Singapore. *Journal of Heritage Tourism*, 18(2), 243–264. <https://doi.org/10.1080/1743873X.2023.2169151>
- Wang, P. Q. (2024). Personalizing guest experience with generative AI in the hotel industry: there's more to it than meets a Kiwi's eye. *Current Issues in Tourism*, 1–18. <https://doi.org/10.1080/13683500.2023.2300030>
- Wen, J., Zheng, D., Jiang, Y., Hou, H., Phau, I., & Wang, W. (2023). Exploring perceptions of travel-eligible individuals with dementia and hotel operators. *Anatolia*, 00(00), 1–13. <https://doi.org/10.1080/13032917.2023.2282050>
- Whalen, E. A., & Sisson, A. D. (2022). Co-branding hotel owners and operators to increase willingness to pay. *Research in Hospitality Management*, 12(2), 151–159. <https://doi.org/10.1080/22243534.2022.2133172>
- Wu, Y., Wang, J., Xia, Y., Li, Q., & Pan, Y. (2024). Sensing hotel customers distribution and their sentiment variations using online travel agent data: a case of Shanghai star-rated hotels. *Annals of GIS*, 00(00), 1–21. <https://doi.org/10.1080/19475683.2024.2335976>
- Yang, Q., Zhu, B., Liao, H., & Wu, X. (2023). Learning consumer preferences from online textual reviews and ratings based on the aggregation-disaggregation paradigm with attitudinal Choquet integral. *Economic Research-Ekonomska Istrazivanja*, 36(1), 3059–3086. <https://doi.org/10.1080/1331677X.2022.2106282>