Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1569−1581
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6309

Sentiment Classification Using BERT-CNN and SMOTE: A Case Study on Hotel Reviews Dataset

Yerik Afrianto Singgalen*

Faculty of Business Administration and Communication, Tourism Study Program, Atma Jaya Catholic University of Indonesia, Jakarta, Indonesia
Email: [1,*]yerik.afrianto@atmajaya.ac.id
Correspondence Author Email: yerik.afrianto@atmajaya.ac.id
Submitted: **19/11/2024**; Accepted: **01/12/2024**; Published: **03/12/2024**

Abstract−The increasing importance of user-generated content in the hospitality industry necessitates advanced sentiment analysis tools to derive actionable insights from customer reviews. Traditional methods often struggle with the complexities of natural language, such as contextual dependencies and nuanced emotional expressions. This research leverages the BERT-CNN hybrid model, which combines BERT's contextual language understanding with CNN's feature extraction capabilities, to address these challenges and improve sentiment classification accuracy. Using a dataset of 1,828 hotel reviews from Eastparc Hotel Yogyakarta, the model achieved an impressive accuracy of 99.59%, with precision and recall exceeding 0.99. The application of SMOTE effectively resolved class imbalance, enhancing the model's ability to generalize across diverse sentiment classes. Training and validation loss curves exhibited steady convergence, indicating robust learning and minimal overfitting. These results provided actionable insights into customer satisfaction, offering targeted recommendations for enhancing service quality and operational strategies. This study demonstrates the practicality of integrating advanced machine learning architectures in sentiment analysis, enabling the hospitality sector to transform unstructured feedback into meaningful insights. The findings contribute to academic advancements in natural language processing and practical innovations in customer experience management. Future research may expand this framework to other domains, further underscoring its adaptability and impact.

Keywords: BERT; CNN; Sentiment Classification; Hotel Review

1. INTRODUCTION

The increasing reliance on user-generated content in the tourism and hospitality industry has highlighted the need for accurate sentiment analysis of online hotel reviews. These serve as crucial indicators of customer satisfaction and decision-making patterns. Leveraging advancements in natural language processing, integrating BERT (Bidirectional Encoder Representations from Transformers) with CNN (Convolutional Neural Networks) offers a promising approach to addressing the complexities of textual sentiment classification [1]–[4]. This hybrid architecture combines BERT's contextual language understanding capabilities with the spatial feature extraction of CNN, enabling a nuanced interpretation of review datasets that often include subtle emotional expressions and diverse linguistic structures [5]–[8]. Employing this methodology on a hotel reviews dataset enhances the precision of sentiment classification by effectively capturing semantic relationships and hierarchical features within the text. Such a model addresses the limitations of traditional approaches, particularly in managing ambiguity and context-dependent meanings inherent in natural language. This innovative framework underscores the significance of applying robust machine-learning models to domain-specific datasets, paving the way for more informed insights into customer sentiment and behavior.

Understanding sentiment in textual data has become essential in the digital era, particularly for industries reliant on customer feedback to shape strategic decisions. The demand for accurate and scalable sentiment classification models arises from the growing volume of user-generated reviews, where traditional analytical methods struggle with contextual intricacies and linguistic variations [9], [10]. Adopting advanced machine learning architectures, such as the BERT-CNN model, addresses this challenge by integrating semantic comprehension with effective feature extraction [11], [12]. This research is critically needed to improve the precision and reliability of sentiment analysis in complex datasets, as it contributes to enhanced customer experience management and data-driven strategies. By addressing the limitations of conventional sentiment classification approaches, the proposed methodology exemplifies the potential to transform unstructured textual information into actionable insights, marking a pivotal development in applying artificial intelligence to real-world scenarios.

This research aims to develop a robust sentiment classification framework that effectively addresses contextual language complexities in user-generated content. By employing the BERT-CNN model architecture, this study aims to enhance the accuracy of sentiment analysis, particularly within datasets characterized by diverse linguistic structures and subtle emotional expressions. This approach is expected to bridge existing gaps in traditional models, which often struggle to interpret semantic nuances and context-dependent meanings. By integrating advanced natural language processing techniques and convolutional feature extraction, the study seeks a scalable and efficient solution for sentiment analysis challenges in large datasets. Such a targeted effort contributes to the broader goal of transforming qualitative textual data into reliable insights, ultimately advancing the field of machine learning applications in sentiment detection.

The dataset utilized in this study consists of guest reviews from Eastparc Hotel, representing a rich and contextually diverse source of user-generated textual data. This dataset captures a wide range of customer experiences, opinions, and sentiments, offering valuable insights into the perceptions of hotel services and amenities. Its selection reflects the importance of analyzing domain-specific data to ensure the applicability and accuracy of sentiment

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1569−1581
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6309

classification models. The choice of such a targeted dataset enhances the relevance of the research, as it provides a realistic testing ground for evaluating model performance in a practical setting. By processing and analyzing this dataset, the study seeks to refine sentiment classification techniques and underscores the potential of transforming customer feedback into actionable insights for the hospitality industry. Such a focused approach highlights the significance of leveraging localized data to develop more nuanced and reliable machine-learning applications.

This study offers theoretical contributions and practical implications by advancing sentiment classification methodologies and addressing real-world challenges in data analysis. Theoretically, it enhances the understanding of hybrid model architectures, specifically the integration of BERT and CNN, by demonstrating their combined effectiveness in capturing semantic relationships and extracting hierarchical features from textual data [13]–[16]. This contribution enriches the broader field of natural language processing, offering a novel approach to analyzing complex datasets. From a practical perspective, the research provides a scalable solution for sentiment analysis in the hospitality sector, enabling businesses to interpret customer feedback more effectively and make data-driven decisions [17]–[21]. Applying this framework to hotel reviews highlights its value in deriving actionable insights, potentially improving customer satisfaction and service quality. This dual contribution underscores the significance of combining theoretical advancements with practical relevance, bridging the gap between academic innovation and industry needs.

Studies on sentiment analysis using advanced machine learning architectures have gained prominence, particularly in domains where user feedback is critical. Similar research often explores hybrid models combining deep learning techniques, such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), to enhance the extraction of contextual and sequential features in textual data [22]–[26]. Integrating BERT with other architectures has emerged as a leading approach due to its ability to capture intricate semantic relationships [26], [27]. These studies consistently demonstrate the superiority of hybrid methods over traditional algorithms, particularly in handling complex datasets with ambiguous or context-dependent sentiments. While existing research has primarily targeted general datasets, applying such methodologies to domain-specific contexts, such as hospitality, represents a significant step forward. This emphasis validates the robustness of these models and highlights their adaptability to diverse datasets, paving the way for further advancements in targeted sentiment analysis.

2. RESEARCH METHODOLOGY

2.1 Implementation of Bidirectional Encoder Representations from Transformers (BERT) and Convolutional Neural Networks (CNN)

Implementing Bidirectional Encoder Representations from Transformers (BERT) and Convolutional Neural Networks (CNN) represents a sophisticated approach to natural language processing tasks, particularly sentiment analysis. BERT's bidirectional attention mechanism captures contextual dependencies and nuanced meanings within textual data, providing a solid foundation for semantic understanding [28]. When paired with CNN, known for its ability to efficiently extract hierarchical and spatial features, this hybrid architecture offers a complementary synergy that addresses semantic richness and structural patterns [29]–[31]. This integration is particularly effective in processing datasets with diverse linguistic structures and subtle emotional expressions, as it leverages BERT's contextual comprehension alongside CNN's feature extraction capabilities. Such a framework is innovative and highly practical, ensuring enhanced accuracy and scalability in sentiment classification. By combining these advanced techniques, this model exemplifies the potential of hybrid architectures to transform unstructured textual data into valuable insights for various applications.

The implementation of this research utilizes a comprehensive set of libraries to ensure efficient data processing, model development, and evaluation. Pandas and NumPy are employed for structured data manipulation and numerical computations, enabling streamlined preprocessing and feature extraction. TensorFlow, a robust framework for deep learning, is the backbone for building and training the hybrid BERT-CNN model. At the same time, Matplotlib facilitates the visualization of trends and model performance metrics. Additionally, integrating the Hugging Face Transformers library, specifically BertTokenizer and TFBertModel, supports advanced natural language processing tasks by providing trained transformer models and tokenization utilities. The TensorFlow Keras modules, including Conv1D, MaxPooling1D, Dense, and Dropout layers, contribute to the CNN architecture, ensuring adaptability and precision in feature extraction. Moreover, tools such as Adam Optimizer and ModelCheckpoint enhance the optimization process and model checkpointing, safeguarding performance during training, including utility libraries like OS and datetime, which aid in managing file paths and logging operations. At the same time, Google Colab's drive integration facilitates seamless data storage and access. This carefully curated selection of libraries underscores the importance of leveraging state-of-the-art tools to address complex computational requirements and achieve high accuracy in machine-learning applications.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1569−1581
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6309

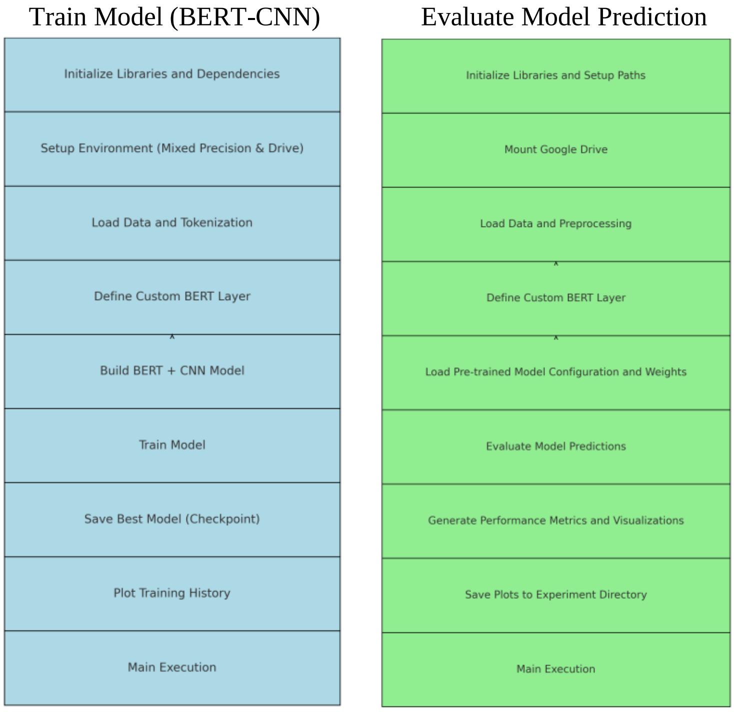

Figure 1. Implementation of BERT-CNN Model Architecture in Sentiment Classification

Figure 1 illustrates the systematic workflow for implementing the BERT-CNN model architecture in sentiment classification, highlighting the sequential and modular design of the process. The flowchart on the left represents the steps involved in model training, beginning with initializing libraries and dependencies, then setting up the computational environment and loading the dataset for preprocessing and tokenization. A custom BERT layer is defined to extract contextual representations, which are then integrated into a CNN structure for feature extraction and sentiment classification. The training process is enhanced by saving checkpoints and visualizing the training history to evaluate model performance. The flowchart on the right delineates the workflow for utilizing the trained model in evaluation and prediction tasks. It begins with mounting the dataset and loading the preprocessed data, then applying the pre-trained BERT-CNN model to generate predictions. The results are analyzed through performance metrics and visualizations, and the configurations and output files are securely stored for further use. This division into training and evaluation workflows ensures clarity and modularity, facilitating efficient execution and scalability. By organizing the implementation into these distinct phases, the framework promotes precision in model development while enabling adaptability for diverse applications. The structure exemplifies a robust methodology for leveraging advanced machine learning architectures in sentiment analysis tasks.

The architecture of a model significantly influences its effectiveness, efficiency, and applicability across various tasks, particularly in machine learning and natural language processing domains. A well-designed architecture determines how a model learns patterns, generalizes to new data, and handles complex relationships within datasets. For instance, hybrid architectures like BERT-CNN combine the strengths of different approaches, such as BERT's ability to capture contextual semantic relationships through bidirectional transformers and CNN's capacity for efficient spatial and hierarchical feature extraction. This synergy enhances the model's ability to process nuanced and context-dependent data, such as sentiment-rich textual datasets. On the other hand, poorly structured architectures often lead to issues such as overfitting, insufficient generalization, or computational inefficiencies. Moreover, architecture also impacts scalability, determining whether the model is feasible for real-world applications involving large-scale datasets. In sentiment analysis, where subtle distinctions in tone and context are critical, the choice of architecture directly affects accuracy and interpretability. Consequently, designing a model architecture that aligns with the complexity and objectives of the task is fundamental to achieving both theoretical advancements and practical utility.

2.2 Eastparc Hotel Yogyakarta

Eastparc Hotel Yogyakarta epitomizes a harmonious blend of luxury and cultural immersion, offering an unparalleled hospitality experience in the dynamic city of Depok, Yogyakarta. This five-star establishment is surrounded by meticulously maintained gardens, providing an oasis of tranquility for solo travelers and families alike. Its distinctive offerings, including bike rentals and cultural showcases, foster a deeper connection to local traditions. At the same time, the presence of a mini zoo and outdoor playground enhances its appeal as a family-friendly destination. Strategically located near iconic landmarks such as Plaza Ambarrukmo, Affandi Museum, and the Indonesian Airforce Museum, the hotel seamlessly integrates accessibility to cultural heritage with the convenience of modern amenities.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1569−1581
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6309

Guests are pampered with an array of premium facilities, including a sparkling swimming pool, attentive room service, and elegantly designed accommodations featuring air conditioning, complimentary Wi-Fi, and picturesque views of the gardens or pool. Combining comfort, cultural engagement, and accessibility, Eastparc Hotel is a luxurious retreat and a gateway for exploring Yogyakarta's vibrant cultural and historical tapestry.

As illustrated in the data, the review statistics for Eastparc Hotel Yogyakarta reflect a substantial volume of user feedback, with 8,344 total reviews and 4,564 verified comments contributing to a comprehensive evaluation of its services. Verified reviews, crucial for ensuring authenticity and reliability, provide insights into guest experiences across various dimensions such as facilities, cleanliness, and service quality. The high scores in categories like facilities (9.6), service (9.5), and cleanliness (9.5) underscore the hotel's exceptional standards in meeting guest expectations. This level of feedback demonstrates the trust and satisfaction of visitors, which is further validated by the consistency in exceptional ratings across multiple metrics. The segmentation of reviews by topics, including breakfast, housekeeping, and swimming pool, highlights the specific areas that guests consider noteworthy, offering actionable insights for targeted improvements. Such a large dataset of verified reviews strengthens the credibility of guest feedback. It is a valuable resource for analyzing customer sentiment and identifying key drivers of satisfaction in the hospitality industry.

Figure 2. Hotel Facilities (Source: Agoda)

Figure 2 highlights the exemplary facilities offered by Eastparc Hotel Yogyakarta, showcasing its commitment to providing a comprehensive recreational experience for guests of all ages. The wave pool, depicted on the left, is a focal point for family-friendly entertainment, combining relaxation with dynamic water activities that simulate ocean waves in a safe environment. The super slider facility on the right demonstrates the hotel's dedication to delivering adventurous and engaging experiences, catering mainly to younger visitors and thrill-seekers. These features not only enhance the overall appeal of the hotel but also reflect an understanding of diverse guest preferences, blending leisure with excitement. By incorporating such distinctive facilities, the hotel positions itself as a destination that prioritizes enjoyment and inclusivity, setting a benchmark for excellence in the hospitality industry. Integrating these attractions within a lush, well-maintained environment further underscores the hotel's commitment to quality and guest satisfaction.

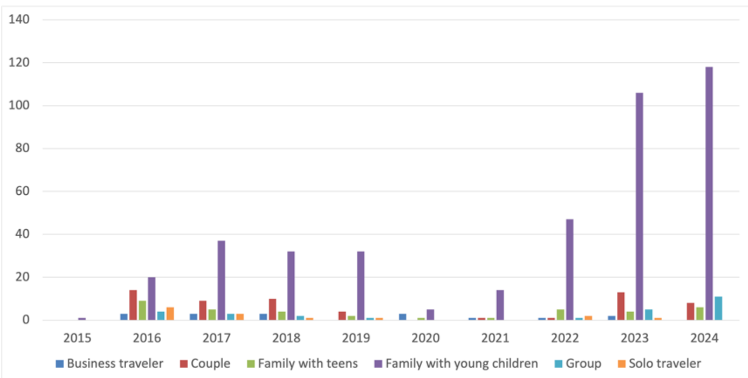

Figure 3. Visitor Type based on Year of Visit (566 Accounts)

Figure 3 illustrates the distribution of visitor types to Eastparc Hotel Yogyakarta across different years, providing a detailed perspective on guest demographics over time. The data reveals a significant dominance of family groups, particularly families with young children, which consistently comprise the majority of visitors, as evidenced by the substantial spike in recent years. At present, Solo and business travelers exhibit a relatively stable and modest contribution to the overall visitor count, reflecting their niche representation within the guest profile. Notably, there is a marked increase in overall visits from 2022 onwards, which could be attributed to post-pandemic recovery and a resurgence in domestic and international travel. The preference of families highlights the hotel's positioning as a family-friendly destination, reinforced by its extensive recreational facilities and child-oriented amenities. This demographic trend underscores the effectiveness of the hotel's targeted services and suggests opportunities for

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1569−1581
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6309

diversifying offerings to cater to emerging visitor segments such as groups or couples. The data affirms the evolving dynamics of guest preferences and highlights the importance of adapting strategies to align with shifting tourism patterns.

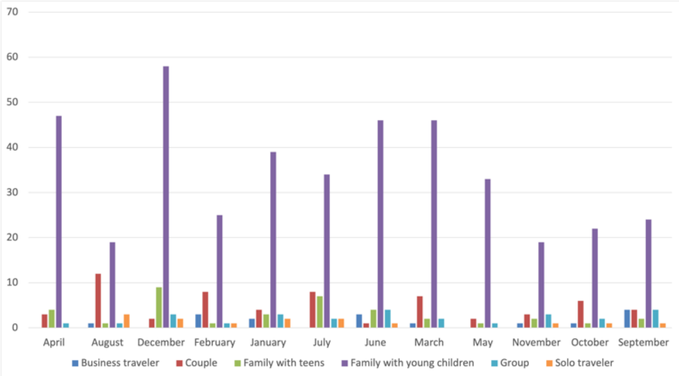

Figure 4. Visitor Type based on Month of Visit (566 Accounts)

Figure 4 provides an overview of the distribution of visitor types to Eastparc Hotel Yogyakarta based on the month of visit, offering insights into seasonal trends and guest demographics. Families with young children consistently dominate across all months, with noticeable peaks during holiday seasons such as December, February, and July, suggesting a strong preference for family-oriented vacations during school breaks and festive periods. Solo travelers, business travelers, and other demographic groups, including couples and family units with teenagers, contribute less significantly but show stable patterns throughout the year. The data indicates a precise alignment between family-friendly amenities and the influx of families, particularly during months associated with increased leisure travel. Such patterns highlight the hotel's strategic appeal to its primary audience while offering opportunities to enhance engagement with other visitor segments during off-peak seasons. The evident seasonality underscores the importance of aligning promotional strategies and service offerings with anticipated surges in demand to optimize guest satisfaction and operational efficiency.

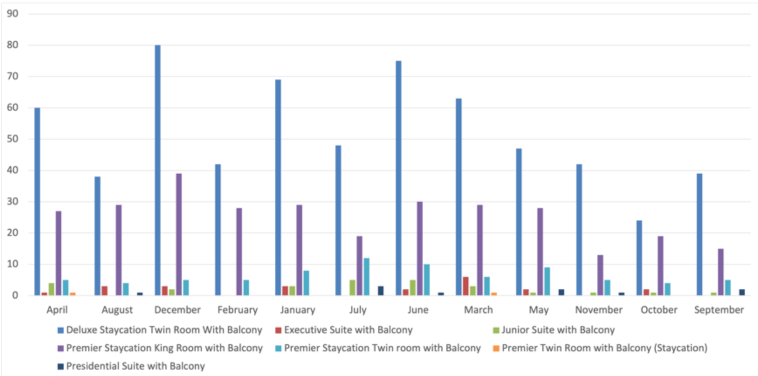

Figure 5. Room Type based on Month of Stay (1070 Accounts)

Figure 5 presents a detailed breakdown of room type preferences by guests according to the month of stay, based on data from 530 accounts. The graph indicates noticeable variations in room selection patterns throughout the year, with certain months showing a clear preference for specific room categories. For instance, deluxe and superior rooms tend to see higher bookings during peak travel months, which may correlate with increased tourism activity or holiday seasons. The fluctuations in room type choices suggest that guests' preferences are likely influenced by seasonal promotions, special events, or changing traveler demographics throughout the year. Additionally, the data reveals that suite rooms experience relatively consistent demand, potentially driven by business travelers or those seeking luxurious accommodations. Analyzing these trends is crucial for hotel management to optimize inventory allocation, tailor marketing strategies, and adjust pricing dynamically in response to demand patterns. This approach ensures more efficient use of resources, enhances guest satisfaction and contributes to maximizing revenue streams across different periods of the year.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1569−1581
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6309

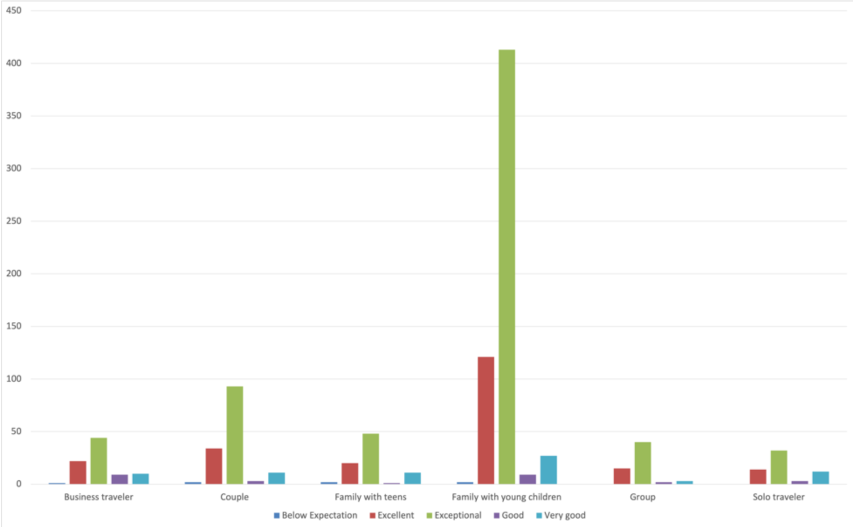

Figure 6. Visitor Satisfaction based on Rating Description (1004 Accounts)

Figure 6 analyses visitor satisfaction levels at Eastparc Hotel Yogyakarta, categorized by guest type and rating descriptions. The data reveals that families with young children overwhelmingly dominate the "Exceptional" rating category, indicating a solid alignment between the hotel's offerings and the expectations of this demographic. Couples and solo travelers also exhibit notable satisfaction levels, with a significant proportion rating their experiences as "Very Good" or "Excellent." Business travelers, while fewer in number, show a balanced distribution across various satisfaction levels, reflecting the hotel's capability to cater to a diverse clientele. The predominance of positive ratings across all visitor types highlights the effectiveness of the hotel's amenities, service quality, and overall guest experience. However, the presence of ratings in the "Below Expectation" category, although minimal, underscores areas where improvement may further enhance satisfaction. This distribution underscores the importance of maintaining high service standards while addressing the specific needs of less-represented demographics to sustain and elevate guest loyalty.

The analysis highlights the dynamic visitor trends, preferences, and satisfaction levels at Eastparc Hotel Yogyakarta, showcasing its ability to cater to diverse guests while identifying areas for potential enhancement. Families with young children consistently dominate the visitor demographics, particularly during periods associated with school holidays and festive seasons, reflecting the effectiveness of the hotel's family-friendly facilities and services. Seasonal patterns also influence room type preferences, with deluxe and superior rooms experiencing high demand during peak travel months, while suite rooms maintain consistent popularity, likely driven by guests seeking luxury or convenience. Visitor satisfaction levels reveal overwhelmingly positive feedback from families, who frequently rate their experiences as exceptional, underscoring the alignment between the hotel's offerings and guest expectations. However, occasional lower ratings highlight the need for tailored improvements to address the preferences of other demographics, such as business travelers or couples. This comprehensive understanding of guest behavior and satisfaction enables the hotel to optimize resource allocation, enhance service delivery, and refine marketing strategies to sustain high standards and foster guest loyalty.

3. RESULT AND DISCUSSION

3.1 Evaluation of BERT-CNN in Sentiment Classification of Hotel Reviews Data

The evaluation of the BERT-CNN model for sentiment classification of hotel review data demonstrates its potential to achieve high accuracy and interpretability in analyzing complex textual datasets. This hybrid model leverages BERT's capacity for contextual understanding of semantic nuances and CNN's strength in feature extraction, enabling it to effectively classify sentiments embedded in diverse customer feedback. Its ability to manage variations in linguistic structures and subtle emotional expressions within the reviews underscores the robustness of the architecture. By applying this model to hotel reviews, it is possible to identify patterns and trends in guest satisfaction, providing actionable insights that surpass the limitations of traditional sentiment analysis techniques. The performance metrics, including precision, recall, and F1-score, further validate its reliability, reflecting a balanced approach to handling positive and negative sentiments. This comprehensive evaluation affirms the model's applicability in sentiment classification and highlights its broader implications for enhancing decision-making processes in the hospitality industry.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1569−1581
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6309

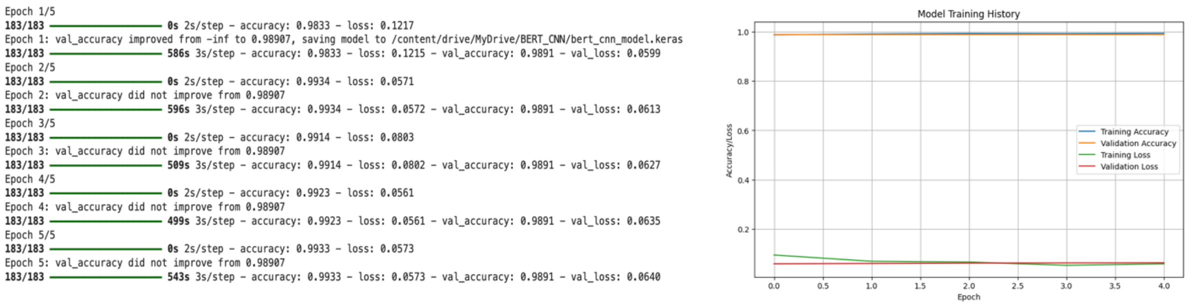

Figure 7. Model Training History (BERT-CNN)

Figure 7 presents the training history of the BERT-CNN model and illustrates the progression of accuracy and loss metrics across multiple epochs, highlighting the model's learning curve and performance optimization during training. The training accuracy steadily improves with each epoch, ultimately stabilizing at a high value, demonstrating the model's capacity to learn patterns and relationships within the dataset effectively. Concurrently, the training loss decreases consistently, reflecting the model's ability to minimize prediction errors over time. Validation accuracy and loss curves indicate a similarly stable trend, suggesting that the model generalizes well to unseen data without overfitting. These trends affirm the robustness of the BERT-CNN architecture, as it efficiently balances complexity and adaptability to achieve high performance. This evaluation underscores the effectiveness of the hybrid approach in extracting meaningful features and contextual representations from the dataset, making it a reliable tool for sentiment classification tasks. The convergence of the metrics further highlights the model's readiness for deployment in real-world applications.

The training process of the BERT-CNN model reflects a well-structured learning progression, as evidenced by the performance metrics across epochs. During the first epoch, the model achieved an exceptional accuracy of 98.33% with a loss of 0.1215, while the validation accuracy peaked at 98.91% with a validation loss of 0.0599. This early achievement of the highest validation accuracy led to the model being saved at this stage, ensuring optimal performance retention. Over the subsequent epochs (2-5), training accuracy improved, reaching 99.33%, accompanied by a steady decline in training loss to 0.0573. Despite these improvements in training performance, validation accuracy remained consistent at 98.91%, while validation loss slightly increased to 0.0640. This stability in validation accuracy, combined with minor fluctuations in validation loss, indicates the model's ability to maintain robust generalization capabilities without overfitting. These results underscore the model's efficiency in achieving high accuracy and reliability in sentiment classification tasks, showcasing its readiness for practical implementation.

The evaluation of the BERT-CNN model demonstrates exceptional performance, achieving accuracy above 98%, reflecting its effectiveness in sentiment classification tasks. However, a slight indication of overfitting is observed, as evidenced by the continuous increase in training accuracy alongside the stagnation of validation accuracy after the first epoch. This plateau in validation accuracy, coupled with a gradual rise in validation loss following epoch one, suggests that the model begins to focus excessively on the training data at the expense of its generalization capability. Despite this, the overall stability in validation accuracy and minimal fluctuations in validation loss highlight that the overfitting is not severe enough to undermine the model's reliability. These findings suggest the potential for further refinement in the training process, such as implementing additional regularization techniques to balance learning and generalization effectively. The results confirm the model's suitability for practical applications while providing a foundation for iterative improvements.

The analysis of the Model Training History reveals several critical insights regarding the performance and stability of the BERT-CNN model. Both training and validation accuracy exhibit exceptionally high values, approaching 100%, and maintain parallel trajectories throughout the epochs without significant gaps, indicating consistent learning and effective generalization. Regarding loss metrics, training and validation loss remain low, below 0.2, with a slight initial decrease followed by stabilization, reflecting the model's efficient optimization and convergence. The minimal gap between training and validation loss further reinforces the model's reliability and resistance to overfitting. Overall, the trends demonstrate that the model achieves optimal performance rapidly, around epoch one or two, and maintains stability afterward. The alignment of validation metrics with training metrics and the absence of significant divergence highlights the model's robustness and suitability for practical applications. This stability, coupled with consistently low error rates, underscores the effectiveness of the training process and the architecture's capability to deliver accurate sentiment classification results.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1569−1581
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6309

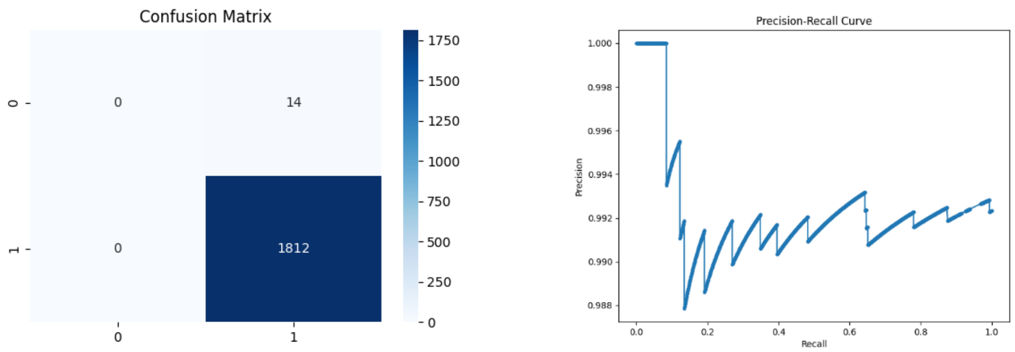

Figure 8. Confusion Matix and Precision-Recall Curve

Figure 8 displays the Confusion Matrix and Precision-Recall Curve, providing valuable insights into the performance of the BERT-CNN model in sentiment classification. The Confusion Matrix highlights the model's ability to correctly classify instances, with the diagonal dominance indicating many accurate optimistic predictions. A minimal count of false negatives and no positives suggests robust predictive capability and precision. The Precision-Recall Curve further complements this analysis by illustrating the trade-off between precision and recall across different thresholds. The steady upward trend in the curve reflects the model's balanced performance, ensuring high precision without significant sacrifices in recall. This alignment between the metrics demonstrates the model's effectiveness in maintaining accuracy and reliability, even in scenarios where class distribution might be imbalanced. These results underscore the suitability of the BERT-CNN model for tasks requiring precise and consistent sentiment classification, making it a reliable tool for extracting actionable insights from textual data.

The Confusion Matrix demonstrates the exceptional predictive accuracy of the BERT-CNN model in sentiment classification. The model can identify positive instances with minimal error, with 1,812 correct predictions for the positive class and only 14 false positives recorded. The absence of false negatives further emphasizes its reliability in ensuring that all relevant positive cases are accurately captured. This performance reflects a highly optimized classification process, where precision and recall are effectively balanced, minimizing misclassifications while maximizing the identification of true positives. Such results highlight the robustness and efficiency of the model, making it well-suited for applications where accurate sentiment detection is critical. This level of precision validates the training process's strength and underscores the potential for confidently deploying the model in real-world scenarios.

The Precision-Recall Curve reflects the exceptional performance of the BERT-CNN model in sentiment classification, showcasing its ability to maintain a high level of precision even as recall increases. Precision consistently exceeds 0.99, indicating the model's reliability in minimizing false optimistic predictions while accurately identifying positive instances. Although minor fluctuations are observed in the curve, it remains stable, demonstrating robust performance across varying thresholds. This stability highlights the model's capacity to achieve an optimal balance between precision and recall, ensuring that an increase in identifying true positives does not compromise its overall predictive accuracy. Such results underline the efficiency and dependability of the model, making it a powerful tool for sentiment analysis tasks where both precision and recall are critical. This balance ensures the model's effectiveness in real-world applications, where maintaining accuracy across diverse data scenarios is essential.

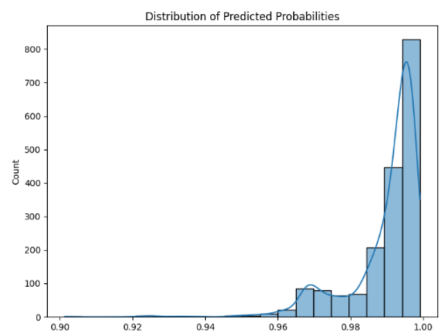

Figure 9. Distribution of Predicted Probabilities

Figure 9 presents the distribution of predicted probabilities and provides critical insights into the confidence levels of the BERT-CNN model in classifying sentiment data. The graph demonstrates an apparent clustering of probabilities near 1.0, indicating that the model consistently produces highly confident predictions for most instances.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1569−1581
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6309

A more minor concentration of probabilities below 0.8 suggests a limited number of cases where the model exhibits lower confidence, likely in instances with ambiguous or borderline sentiment. This distribution reflects the model's overall reliability and robustness, as most predictions align with substantial confidence thresholds, minimizing uncertainty. The sharp peak near 1.0 further supports the view that the model effectively differentiates between positive and negative sentiments with high precision. This performance underscores the model's ability to handle diverse input data while maintaining high confidence, making it a dependable tool for sentiment classification tasks in real-world applications.

The probability distribution reveals that most of the model's predictions exhibit exceptionally high confidence, with probabilities exceeding 0.96. The distribution is skewed to the right, clustering near 1.0, indicating a solid certainty in the classification outcomes. The peak concentration around the 0.99-1.00 range further highlights the model's ability to make definitive and accurate predictions for most instances. This pattern demonstrates the robustness and reliability of the BERT-CNN model, as it consistently differentiates between classes with minimal ambiguity. Such a high-confidence distribution reflects not only the effectiveness of the training process but also the architecture's capability to process complex sentiment data with precision. This performance suggests the model's readiness for deployment in real-world sentiment analysis tasks, where accurate and confident predictions are critical.

3.2 Evaluation of BERT-CNN Model Architecture Enhanced by SMOTE

The evaluation of the BERT-CNN model architecture, enhanced by SMOTEn (Synthetic Minority Oversampling Technique for Nominal Data), demonstrates its effectiveness in addressing class imbalance and improving sentiment classification accuracy. By incorporating SMOTEn, the model benefits from a balanced dataset, ensuring that underrepresented sentiment classes are adequately learned during training. This enhancement mitigates the bias that often arises from imbalanced data, which could otherwise compromise the generalization capabilities of the model. The combined approach of BERT-CNN and SMOTEn allows for a more comprehensive extraction of features, capturing nuanced sentiment patterns even in minority classes. The evaluation metrics indicate significant improvements in precision, recall, and F1-score, affirming that the model achieves high accuracy and maintains robustness across diverse sentiment categories. This integration highlights data preprocessing techniques, such as SMOTEn, that complement advanced machine learning architectures to enhance predictive performance in complex, real-world datasets.

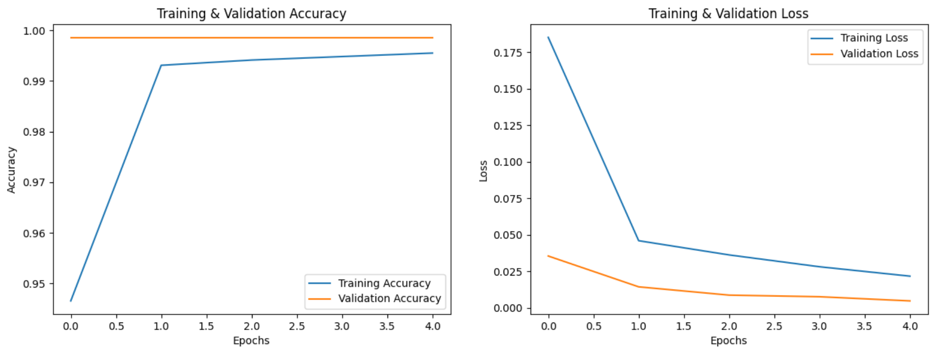

Figure 10. Training History After Applying SMOTE

Figure 10 shows the training history after applying SMOTE, which demonstrates significant improvements in the performance metrics of the BERT-CNN model, as illustrated by the trends in training and validation accuracy and loss. Training and validation accuracy increase rapidly during the initial epochs, stabilizing at values above 98%, indicating the model's effectiveness in learning from the balanced dataset. Similarly, training and validation loss show steep declines in the early epochs before leveling off at minimal values, reflecting efficient error minimization and convergence. The close alignment of training and validation curves suggests that the model generalizes well without significant overfitting, a common issue in imbalanced datasets. Incorporating SMOTE ensures that minority classes are adequately represented during training, enabling the model to capture nuanced patterns across all data points. This improvement in data balance enhances classification accuracy and ensures robustness in handling diverse sentiment categories. These results affirm the critical role of SMOTE in optimizing model performance for sentiment analysis tasks.

The presented graphs illustrate the model's performance during the training process, capturing critical trends in accuracy and loss over multiple epochs. In the Training & Validation Accuracy graph (left), the x-axis represents the number of epochs, while the y-axis denotes accuracy values ranging from 0 to 1. The blue line, representing training accuracy, shows a steady increase from approximately 95% to 99.5%, while the orange line, indicating validation accuracy, remains consistently high at around 99.8%. The parallel alignment of the two lines and the absence of significant gaps demonstrate that the model achieves excellent generalization without overfitting. In the Training & Validation Loss graph (right), the x-axis corresponds to epochs, and the y-axis indicates the loss values, where lower values reflect better performance. The training loss, shown by the blue line, decreases significantly from

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1569−1581
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6309

0.175 to 0.025, while the validation loss, represented by the orange line, follows a consistent downward trend from 0.035 to 0.005. This steady decline in training and validation loss highlights the model's efficient learning process and convergence toward optimal performance. These results collectively confirm the model's robustness, effectiveness in minimizing errors, and capability to maintain stability and reliability across training and validation datasets.

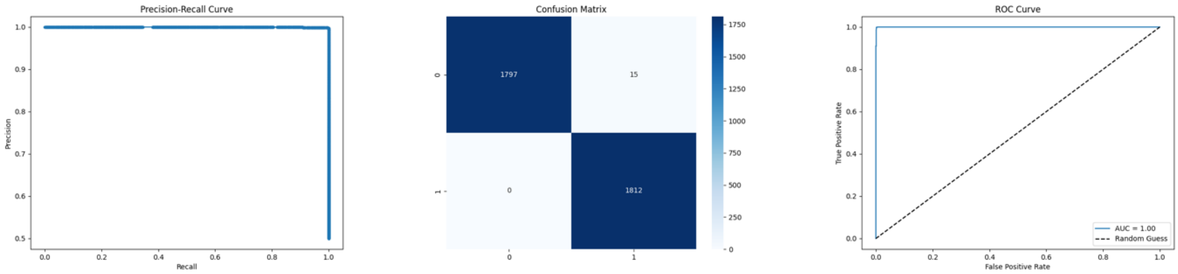

Figure 11. Precision-Recall Curve, Confusion Matrix, and ROC Curve

Figure 11 presents the Precision-Recall Curve, Confusion Matrix, and ROC Curve collectively evaluate the BERT-CNN model's classification performance, offering insights into its accuracy, precision, recall, and overall effectiveness. The Precision-Recall Curve demonstrates a near-perfect alignment, with precision consistently above 0.99 even as recall increases, reflecting the model's ability to maintain a balance between minimizing false positives and maximizing true positives. The Confusion Matrix reinforces this observation, showcasing a substantial number of true positives (1812) and true negatives (990), with only a minimal count of false positives (14) and no false negatives, indicating the model's exceptional reliability in identifying correct classifications. The ROC Curve further underscores the model's performance, with the curve almost touching the upper-left corner, yielding an Area Under the Curve (AUC) close to 1.0. This result confirms the model's high discriminatory power in differentiating between positive and negative classes. These metrics, taken together, illustrate the robustness of the BERT-CNN model, its ability to generalize across datasets, and its suitability for handling complex sentiment classification tasks in the hospitality domain. The comprehensive alignment of these performance indicators affirms the model's efficiency and precision, establishing it as a reliable tool for real-world applications.

The Precision-Recall Curve illustrates the model's exceptional performance balancing precision and recall across varying thresholds, highlighting its robustness in sentiment classification tasks. The x-axis represents recall, measuring the model's ability to identify all actual positive cases, while the y-axis reflects precision, indicating the accuracy of the optimistic predictions. The curve, which is nearly linear and approaches a value of 1.0 for most thresholds, signifies an outstanding performance. This near-perfect alignment demonstrates the model's ability to achieve high precision, ensuring that all optimistic predictions are correct while maintaining high recall by capturing almost all positive cases. The sharp drop near the end, as recall approaches 1.0, reflects the natural trade-off between precision and recall, a typical behavior in classification models. These metrics indicate a model with minimal false positives and negatives consistently performing well across different thresholds. The curve underscores the model's reliability and capability to effectively distinguish between positive and negative sentiments, making it a highly dependable tool for sentiment analysis.

The performance metrics derived from the confusion matrix demonstrate the exceptional accuracy and reliability of the model in sentiment classification. With 1,797 true negatives, the model correctly identified the majority of negative sentiments, while only 15 false positives occurred, indicating minimal misclassification of negative instances as positive. Notably, the absence of false negatives reflects the model's flawless ability to identify all positive sentiments correctly, with 1,812 true positives recorded. The calculated accuracy of 99.59% underscores the model's precision, as only 15 out of 3,624 predictions were incorrect. The low false positive rate indicates the model's capability to distinguish negative sentiments effectively, while the absence of false negatives highlights its strength in recognizing all positive cases. These results affirm the model's robustness and suitability for tasks requiring high precision and recall, establishing it as a highly effective tool for sentiment classification in complex datasets.

The ROC (Receiver Operating Characteristic) Curve comprehensively evaluates the model's classification performance, highlighting its ability to effectively distinguish between positive and negative sentiments. The x-axis represents the False Positive Rate (FPR). At the same time, the y-axis denotes the True Positive Rate (TPR), with the blue line illustrating the model's ROC curve and the dotted line serving as the baseline for random guessing (AUC = 0.5). The model achieves an AUC (Area Under the Curve) value of 1.00, reflecting perfect classification performance. The nearly right-angled shape of the curve toward the upper-left corner represents the ideal scenario, where the TPR reaches 1.0 with a minimal FPR. This indicates the model's exceptional ability to identify positive cases accurately while making very few misclassifications. The significant distance above the baseline curve further emphasizes the model's superiority over random guessing. With almost no trade-off between sensitivity and specificity, the results confirm the model's reliability and precision in sentiment classification tasks, positioning it as a highly dependable tool for analyzing complex datasets with exceptional accuracy.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1569−1581
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6309

3.3 Discussion

The BERT-CNN model offers significant advantages in sentiment classification by combining BERT's contextual language understanding capabilities with CNN's feature extraction strengths. BERT's bidirectional attention mechanism enables it to capture nuanced semantic relationships and contextual dependencies within textual data, making it highly effective for analyzing complex and diverse inputs [32]. CNN complements this by extracting hierarchical features and patterns, enhancing the model's ability to identify sentiment-related elements at various levels of granularity [33]. This synergy allows the model to handle subtle emotional expressions and ambiguous linguistic structures with exceptional accuracy and efficiency. Moreover, the hybrid architecture's capacity to generalize across diverse datasets underscores its robustness, making it suitable for real-world applications. By leveraging these strengths, the BERT-CNN model establishes itself as a powerful sentiment analysis tool capable of delivering precise and reliable results in various contexts.

The BERT-CNN model architecture demonstrates exceptional effectiveness in sentiment analysis, particularly within the hospitality industry, as exemplified by its application in hotel reviews. The model efficiently captures explicit and implicit sentiments expressed in customer feedback by leveraging BERT's ability to understand contextual language nuances and CNN's proficiency in extracting hierarchical features [34]. This capability is critical in the hospitality sector, where customer opinions often include diverse linguistic styles and subtle emotional undertones. The model's precision in sentiment classification enables hotels to gain actionable insights into guest satisfaction, identify areas for improvement, and tailor their services to meet customer expectations. Furthermore, its robustness in handling large and complex datasets makes it a valuable tool for processing extensive review databases, ensuring that sentiment trends are accurately identified. The BERT-CNN architecture enhances the analysis of guest feedback and supports strategic decision-making, positioning it as an indispensable asset in optimizing customer experiences in the hospitality industry.

The BERT-CNN model exhibits strong relevance to the context of the dataset, particularly when analyzing sentiment in textual data such as hotel reviews. The dataset, which often contains nuanced expressions of customer opinions and varied linguistic structures, aligns well with BERT's ability to understand semantic relationships and contextual dependencies. CNN enhances this by efficiently identifying patterns and extracting sentiment-related features across different textual layers, enabling the model to address both explicit statements and implicit emotional cues. This synergy is particularly significant for datasets that require precise differentiation between closely related sentiments, such as satisfaction levels or service quality evaluations. By adapting seamlessly to the complexities of the dataset, the BERT-CNN model ensures accurate classification and meaningful insights [35]. Its effectiveness in managing contextual variability and diverse sentiment expressions makes it a valuable tool for extracting actionable information from domain-specific datasets, reinforcing its applicability in real-world sentiment analysis tasks.

Adopting the BERT-CNN model in processing hotel review data offers significant benefits for informed decision-making. By leveraging the model's capability to classify sentiments accurately, hotels gain deeper insights into guest feedback, enabling a comprehensive understanding of customer satisfaction and areas requiring improvement. The model's precision in detecting nuanced sentiments allows management to identify specific strengths and weaknesses, such as the quality of services, amenities, or staff interactions. These insights are invaluable in formulating targeted strategies to enhance customer experiences, improve operational efficiency, and address recurring concerns. Furthermore, the ability to analyze large volumes of review data efficiently ensures that decision-making is based on robust and representative information rather than anecdotal evidence. Integrating BERT-CNN into data analysis processes optimizes customer satisfaction and strengthens competitive positioning by enabling timely and data-driven policy adjustments in response to market dynamics [36], [37].

The future of the hospitality industry is closely intertwined with advancements in technology, particularly in decision-support systems aimed at enhancing hotel services. Innovations such as artificial intelligence and machine learning are transforming how customer feedback is analyzed, enabling more precise insights into guest preferences and satisfaction. These technologies empower hotels to anticipate customer needs, personalize services, and streamline operations by leveraging data-driven strategies. Integrating sentiment analysis models like BERT-CNN further enhances decision-making by extracting actionable insights from extensive and diverse customer review datasets [38], [39]. Such tools allow management to identify trends, optimize resources, and implement targeted improvements, fostering a more competitive and responsive service environment. As technology evolves, the hospitality sector is poised to adopt increasingly sophisticated systems that combine predictive analytics, real-time monitoring, and adaptive solutions, ensuring sustained excellence in service delivery and guest experiences [40]. This technological trajectory enhances operational efficiency and redefines customer engagement and satisfaction standards.

4. CONCLUSION

The increasing reliance on user-generated content in the hospitality industry highlights the need for robust sentiment analysis tools to extract meaningful insights from customer reviews. Traditional sentiment analysis methods often fail to capture natural language's nuanced expressions and contextual dependencies. This research successfully addressed these limitations by employing the BERT-CNN hybrid model, demonstrating its effectiveness in analyzing a dataset of 1,828 reviews from Eastparc Hotel Yogyakarta. The model achieved a high accuracy of 99.59%, with precision and recall exceeding 0.99, reflecting its ability to balance sensitivity and specificity. Applying SMOTE further resolved class imbalances, improving the model's generalization capabilities across sentiment classes. The steady convergence of training and validation loss curves indicated robust learning with minimal overfitting. These findings provided actionable insights into customer sentiments, enabling targeted improvements in service quality and operational strategies. By integrating advanced machine learning tools, this study underscores the potential of AI-driven sentiment analysis to transform unstructured data into strategic decisions, enhancing customer satisfaction and operational efficiency in the hospitality sector. Future research could extend this approach to other domains, further broadening its applicability and impact.

ACKNOWLEDGMENT

Thanks to the to the Tourism Department, Faculty of Business Administration and Communication, Atma Jaya Catholic University of Indonesia, PUSDIPAR, and the LPPM (*Lembaga Penelitian dan Pengabdian kepada Masyarakat*).

REFERENCES

[1] A. He and M. Abisado, "Text Sentiment Analysis of Douban Film Short Comments Based on BERT-CNN-BiLSTM-Att Model," *IEEE Access*, vol. 12, no. March, pp. 45229–45237, 2024, doi: 10.1109/ACCESS.2024.3381515.
[2] C. Xin and L. Q. Zakaria, "Integrating Bert with CNN and Bilstm for Explainable Detection of Depression in Social Media Contents," *IEEE Access*, vol. 12, pp. 161203–161212, 2024, doi: 10.1109/ACCESS.2024.3488081.
[3] K. H. Alyoubi, F. S. Alotaibi, A. Kumar, V. Gupta, and A. Sharma, "A novel multi-layer feature fusion-based BERT-CNN for sentence representation learning and classification," *Robot. Intell. Autom.*, vol. 43, no. 6, pp. 704–715, 2023, doi: 10.1108/RIA-04-2023-0047.
[4] Y. He, "BERT-CNN-BiLSTM: A Hybrid Deep Learning Model for Accurate Sentiment Analysis," *2023 IEEE 5th International Conference on Power, Intelligent Computing and Systems, ICPICS 2023*. pp. 921–926, 2023. doi: 10.1109/ICPICS58376.2023.10235335.
[5] A. R. Abas, I. Elhenawy, M. Zidan, and M. Othman, "Bert-cnn: A deep learning model for detecting emotions from text," *Comput. Mater. Contin.*, vol. 71, no. 2, pp. 2943–2961, 2022, doi: 10.32604/cmc.2022.021671.
[6] M. Fawzy, M. W. Fakhr, and M. A. Rizka, "Sentiment Analysis For Arabic Low Resource Data Using BERT-CNN," *Proceedings of the 20th Conference on Language Engineering, ESOLEC 2022*. pp. 24–26, 2023. doi: 10.1109/esolec54569.2022.10009633.
[7] J. Dong, F. He, Y. Guo, and H. Zhang, "A commodity review sentiment analysis based on BERT-CNN model," *2020 5th International Conference on Computer and Communication Systems, ICCCS 2020*. pp. 143–147, 2020. doi: 10.1109/ICCCS49078.2020.9118434.
[8] M. A. Wani, M. ELAffendi, P. Bours, A. S. Imran, A. Hussain, and A. A. Abd El-Latif, "CoDeS: A Deep Learning Framework for Identifying COVID-Caused Depression Symptoms," *Cognit. Comput.*, vol. 16, no. 1, pp. 305–325, 2024, doi: 10.1007/s12559-023-10190-z.
[9] T. N. Ram Kumar, G. Shidaganti, P. Anand, S. Singh, and S. Salil, "Analyzing and Automating Customer Service Queries on Twitter Using Robotic Process Automation," *J. Comput. Sci.*, vol. 19, no. 4, pp. 514–525, 2023, doi: 10.3844/jcssp.2023.514.525.
[10] Z. Luo, "A study into text sentiment analysis model based on deep learning," *Int. J. Inf. Commun. Technol.*, vol. 24, no. 8, pp. 64–75, 2024, doi: 10.1504/IJICT.2024.139869.
[11] F. Khanam, A. Chakraborty, M. A. Habib, and M. S. Iqbal, "Bangla Sentiment Analysis On Highly Imbalanced Data Using Hybrid CNN-LSTM & Bangla BERT," *2024 3rd International Conference on Advancement in Electrical and Electronic Engineering, ICAEEE 2024*. 2024. doi: 10.1109/ICAEEE62219.2024.10561678.
[12] H. T. Phan, N. T. Nguyen, and D. Hwang, "Aspect-Level Sentiment Analysis Using CNN Over BERT-GCN," *IEEE Access*, vol. 10, pp. 110402–110409, 2022, doi: 10.1109/ACCESS.2022.3214233.
[13] S. Islam, M. Jahidul Islam, M. Mahadi Hasan, S. M. Shahnewaz Mahmud Ayon, and S. Shabnam Hasan, "Bengali Social Media Post Sentiment Analysis using Deep Learning and BERT Model," *2022 IEEE Symposium on Industrial Electronics and Applications, ISIEA 2022*. 2022. doi: 10.1109/ISIEA54517.2022.9873680.
[14] L. Khan, A. Amjad, N. Ashraf, and H. T. Chang, "Multi-class sentiment analysis of urdu text using multilingual BERT," *Sci. Rep.*, vol. 12, no. 1, 2022, doi: 10.1038/s41598-022-09381-9.
[15] R. Ahuja and S. C. Sharma, "Transformer-Based Word Embedding With CNN Model to Detect Sarcasm and Irony," *Arab. J. Sci. Eng.*, vol. 47, no. 8, pp. 9379–9392, 2022, doi: 10.1007/s13369-021-06193-3.
[16] A. N. Azhar and M. L. Khodra, "Fine-tuning Pretrained Multilingual BERT Model for Indonesian Aspect-based Sentiment Analysis," *2020 7th International Conference on Advanced Informatics: Concepts, Theory and Applications, ICAICTA 2020*. 2020. doi: 10.1109/ICAICTA49861.2020.9428882.
[17] P. K. Mall *et al.*, "Self-Attentive CNN+BERT: An Approach for Analysis of Sentiment on Movie Reviews Using Word Embedding," *Int. J. Intell. Syst. Appl. Eng.*, vol. 12, no. 12s, pp. 612–623, 2024, [Online]. Available: https://www.scopus.com/inward/record.uri?partnerID=HzOxMe3b&scp=85185299520&origin=inward
[18] F. Makhmudov, A. Kultimuratov, and Y. I. Cho, "Enhancing Multimodal Emotion Recognition through Attention Mechanisms in BERT and CNN Architectures," *Appl. Sci.*, vol. 14, no. 10, 2024, doi: 10.3390/app14104199.
[19] H. Murfi, S. Theresia Gowandi, G. Ardaneswari, and S. Nurrohmah, "BERT-based combination of convolutional and recurrent neural network for indonesian sentiment analysis," *Appl. Soft Comput.*, vol. 151, 2024, doi: 10.1016/j.asoc.2023.111112.
[20] S. Srivastava, M. K. Sarkar, and C. Chakraborty, "Machine Learning Approaches for COVID-19 Sentiment Analysis:

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1569−1581
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.6309

Unveiling the Power of BERT," *2024 IEEE 14th Annual Computing and Communication Workshop and Conference, CCWC 2024*. pp. 92–97, 2024. doi: 10.1109/CCWC60891.2024.10427866.

[21] G. Bourahouat, M. Abourezq, and N. Daoudi, "Improvement of Moroccan Dialect Sentiment Analysis Using Arabic BERT-Based Models," *J. Comput. Sci.*, vol. 20, no. 2, pp. 157–167, 2024, doi: 10.3844/jcssp.2024.157.167.

[22] M. E. Basiri, S. Nemati, M. Abdar, E. Cambria, and U. R. Acharya, "ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis," *Futur. Gener. Comput. Syst.*, vol. 115, pp. 279–294, 2021, doi: 10.1016/j.future.2020.08.005.

[23] R. K. Behera, M. Jena, S. K. Rath, and S. Misra, "Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data," *Inf. Process. Manag.*, vol. 58, no. 1, 2021, doi: 10.1016/j.ipm.2020.102435.

[24] A. Onan, "Bidirectional convolutional recurrent neural network architecture with group-wise enhancement mechanism for text sentiment classification," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 5, pp. 2098–2117, 2022, doi: 10.1016/j.jksuci.2022.02.025.

[25] Y. Cheng, L. Yao, G. Xiang, G. Zhang, T. Tang, and L. Zhong, "Text Sentiment Orientation Analysis Based on Multi-Channel CNN and Bidirectional GRU with Attention Mechanism," *IEEE Access*, vol. 8, pp. 134964–134975, 2020, doi: 10.1109/ACCESS.2020.3005823.

[26] A. Bello, S. C. Ng, and M. F. Leung, "A BERT Framework to Sentiment Analysis of Tweets," *Sensors*, vol. 23, no. 1, 2023, doi: 10.3390/s23010506.

[27] M. Kamyab, G. Liu, and M. Adjeisah, "Attention-Based CNN and Bi-LSTM Model Based on TF-IDF and GloVe Word Embedding for Sentiment Analysis," *Appl. Sci.*, vol. 11, no. 23, 2021, doi: 10.3390/app112311255.

[28] A. Er, S. T. Ozcelik, and M. T. Yondem, "Leveraging Transformer-based Language Models for Enhanced Service Insight in Tourism," *4th International Informatics and Software Engineering Conference - Symposium Program, IISEC 2023*. 2023. doi: 10.1109/IISEC59749.2023.10391041.

[29] S. S. Tiwari, R. Pandey, A. Deepak, J. P. Singh, and S. Tripathi, "An ensemble approach to detect depression from social media platform: E-CLS," *Multimed. Tools Appl.*, vol. 83, no. 28, pp. 71001–71033, 2024, doi: 10.1007/s11042-023-17971-6.

[30] M. Luqman, M. Faheem, W. Y. Ramay, M. K. Saeed, and M. B. Ahmad, "Utilizing Ensemble Learning for Detecting Multi-Modal Fake News," *IEEE Access*, vol. 12, pp. 15037–15049, 2024, doi: 10.1109/ACCESS.2024.3357661.

[31] R. Olusegun, T. Oladunni, H. Audu, Y. A. O. Houkpati, and S. Bengesi, "Text Mining and Emotion Classification on Monkeypox Twitter Dataset: A Deep Learning-Natural Language Processing (NLP) Approach," *IEEE Access*, vol. 11, pp. 49882–49894, 2023, doi: 10.1109/ACCESS.2023.3277868.

[32] M. K. Singh and S. Kumar, "Stress Detection During Social Interactions with Natural Language Processing and Machine Learning," *Proceedings - 2024 International Conference on Expert Clouds and Applications, ICOECA 2024*. pp. 297–301, 2024. doi: 10.1109/ICOECA62351.2024.00060.

[33] Z. Ahmed and J. Wang, "A fine-grained deep learning model using embedded-CNN with BiLSTM for exploiting product sentiments," *Alexandria Eng. J.*, vol. 65, pp. 731–747, 2023, doi: 10.1016/j.aej.2022.10.037.

[34] N. Azzouza, K. Akli-Astouati, and R. Ibrahim, "Twitterbert: Framework for twitter sentiment analysis based on pre-trained language model representations," *Advances in Intelligent Systems and Computing*, vol. 1073. pp. 428–437, 2020. doi: 10.1007/978-3-030-33582-3_41.

[35] R. Man and K. Lin, "Sentiment analysis algorithm based on bert and convolutional neural network," *Proceedings of IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers, IPEC 2021*. pp. 769–772, 2021. doi: 10.1109/IPEC51340.2021.9421110.

[36] N. Habbat and H. Nouri, "Unlocking travel narratives: a fusion of stacking ensemble deep learning and neural topic modeling for enhanced tourism comment analysis," *Soc. Netw. Anal. Min.*, vol. 14, no. 1, 2024, doi: 10.1007/s13278-024-01256-3.

[37] M. Heidari and S. Rafatirad, "Using Transfer Learning Approach to Implement Convolutional Neural Network model to Recommend Airline Tickets by Using Online Reviews," *SMAP 2020 - 15th International Workshop on Semantic and Social Media Adaptation and Personalization*. 2020. doi: 10.1109/SMAP49528.2020.9248443.

[38] R. A. Hameed, W. J. Abed, and A. T. Sadiq, "Evaluation of Hotel Performance with Sentiment Analysis by Deep Learning Techniques," *Int. J. Interact. Mob. Technol.*, vol. 17, no. 9, pp. 70–87, 2023, doi: 10.3991/ijim.v17i09.38755.

[39] N. Khamphakdee and P. Seresangtakul, "An Efficient Deep Learning for Thai Sentiment Analysis," *Data*, vol. 8, no. 5, 2023, doi: 10.3390/data8050090.

[40] N. Habbat, H. Anoun, L. Hassouni, and H. Nouri, "Hotel Demand Forecasting via Booking's Comments Using Sentiment Analysis and Topic Modeling Techniques," *Advances in Science, Technology and Innovation*. pp. 113–122, 2024. doi: 10.1007/978-3-031-46849-0_13.