International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. #, No.# (Mon-20..)

Detection of Violations on Public Reporting to Police with CNN-BiLSTM Hybrid Architecture Development

Ade Oktarino[1], Sarjon Defit[2] and Yuhandri[3]

[1] *Faculty of Engineering and Computer Science, Adiwangsa Jambi University, Jambi City, Indonesia*
[2,3] *Faculty of Computer Science, Putra Indonesia University "YPTK", Padang, Indonesia*

E-mail address: adeoktarino@unaja.ac.id, sarjon_defit@upiyptk.ac.id, yuyui@upiyptk.ac.id

Received ## Mon. 20##, Revised ## Mon. 20##, Accepted ## Mon. 20##, Published ## Mon. 20##

Abstract: The Jambi Regional Police Department (Polda Jambi) has introduced an innovative system for managing public complaints via WhatsApp. However, this implementation faces challenges due to the manual documentation process, leading to poorly documented and inaccurate complaints. To address these issues, this study proposes the development of a text mining-based system utilizing deep learning to facilitate the accurate categorization of public complaints, thereby streamlining the police's processing of these reports. Deep learning, specifically through Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BiLSTM), offers a robust framework for learning patterns from complaint data. This study develops and optimizes a CNN-BiLSTM architecture, involving the adjustment of layer configurations and the application of early stopping techniques to prevent overfitting. The proposed architecture is tested on two datasets: public complaints submitted to the Indonesian National Police via WhatsApp and Tweets from social media X. Experimental results indicate high performance across both datasets, with the architecture achieving a peak accuracy of 99% on the police data and 79% on the Twitter data. The highest-performing model is then integrated into a graphical user interface (GUI) using Streamlit, enabling the efficient and accurate classification of public complaints. This system demonstrates significant potential for enhancing the efficiency and accuracy of complaint management processes within law enforcement agencies. The findings suggest that integrating advanced deep learning techniques into public complaint systems can substantially improve the documentation and categorization of complaints, providing a scalable solution for law enforcement operations.

Keywords: Deep Learning Architectures, CNN-BiLSTM, Early Stopping, WhatsApp, Tweet.

1. INTRODUCTION

The policy of the Indonesian National Police to implement an online complaint system via WhatsApp is commendable. This system enables the public to report crimes or incidents more swiftly. However, the police face challenges in processing these complaints. The incoming complaints are not pre-filtered, necessitating manual filtering to categorize them appropriately. Therefore, a more efficient system is needed to address these issues, enhancing the effectiveness and efficiency of police operations.

To solve the problem, one attempt is to use data science. Data science is the study of the volume of data, using modern techniques to find invisible patterns, obtain meaningful information, and make business decisions with that information [1]. The data science applied can help in the classification of the type of public report. Deep learning is a part of machine learning with the function of training basic human instincts on computers using computer algorithms [2][3]. Deep learning has two approaches: generative architecture and discriminative architecture [4]. The proposed research will employ discriminatory architecture, as it necessitates the labeling of datasets prior to classification [5]. This research has a foundation in the development of deep learning architecture, which is based on some previous research.

The CNN (Convolutional Neural Network) algorithm is used to classify public sentiment; the labels used to do the classification are negative, positive, and neutral. The results of this study obtained an accuracy of 87% [6]. Then another algorithm, BiLSTM (Bidirectional Long Short-Term Memory), is also used to carry out sentiment analysis using two labels, namely positive and negative. The results obtained in this algorithm with LSTM are 84% [7]. Several researchers then hybridised these two algorithms to create

E-mail:author's email

CNN-BiLSTM. This hybrid was carried out to analyze sentiment analysis using word embedding Glove. The labels used are positive, negative, and neutral. The results using this hybrid give an accuracy of 95.65 [8]. Then the research also carried out sentiment analysis using CNN-BiLSTM. The resulting accuracy is quite good, namely, 90.66% [9].

This research, which is based on several previous studies, will carry out various tests on WhatsApp data from community reports. The researchers developed CNN, BiLSTM, and hybrid CNN-BiLSTM. In this study, word2vec is used as the word embedding. Apart from that, this research also uses SMOTE (Synthetic Minority Oversampling Technique) to overcome the problem of class imbalance [10]. The research then uses preprocessing to better prepare the data [11].

The development involves adding several layers to increase accuracy. In addition, this research employs early stopping as a strategy to combat overfitting [12]. With the development carried out, it is hoped that it will be able to increase the accuracy of both the CNN, BiLSTM, and hybrid CNN - BiLSTM algorithms.

2. Related Study

The research carried out is a development of previous research. Each research on CNN-BiLSTM has a different architecture. Table 1 is the architecture developed by previous research.

TABLE I. Differences With Previous And Developed Architecture

Researcher	Architecture	Accuracy
[13]	Input → Word Embedding → BiLSTM → CNN → Max-Pooling → CNN → Max-Pooling → CNN → Max-Pooling → Dense → Dense → Output	77.49%
[14]	Input Convolution + Relu → Convolution + Relu →Max Pool → Flatten → BiLSTM → Fully Connected Dense Layer → Output	98.64%
[15]	Input → Preprocessing → Word embedding → CNN → Max Pooling → Fully Connected Layer → BiLSTM → Dense → Sigmoid → Accuracy	92.85%
[16]	Input → Word Embedding → CNN → Max Pooling1d → BiLSTM→ Dense → Dropout → Dense → Output	87%
[17]	Input → Word Embedding → Conv → Pool → Conv → Pool → Conv → Pool → Fully Connected → Output	70.2%
This Study	Input → Labelling → SMOTE → Pre-Processing → Word Embedding (word2vec) → 1D Conv → Max-Polling → 1D Conv → Max-Polling → Flatten → Dropout → Fully Connected → BiLSTM → Dropout → BiLSTM → Dropout → Fully Connected → softmax → early Stopping	-

The architectural differences highlighted in Table 1 show the evolution of CNN-BiLSTM models over various studies. Each architecture incorporates unique configurations and layer sequences to enhance performance and address specific challenges in processing and analyzing sequential data, such as text.

Research [13] utilized a combination of BiLSTM and multiple CNN layers with max-pooling, followed by dense layers. This approach achieved an accuracy of 77.49%. Research [14] implemented a dual convolutional layer with ReLU activation, followed by max pooling, flattening, and a BiLSTM layer. This model reached an accuracy of 98.64%. Research [15] incorporated preprocessing, word embedding, CNN with max pooling, a fully connected layer, and a BiLSTM layer with a sigmoid activation function. This study achieved 92.85% accuracy. Research [16] employed a straightforward approach with word embedding, CNN, max pooling, BiLSTM, and dropout layers, achieving an accuracy of 87%. Research [17] used multiple convolution and pooling layers followed by a fully connected output layer, resulting in 70.2% accuracy.

In contrast, this study introduces an elaborate architecture, integrating multiple stages of 1D convolution, max-pooling, and BiLSTM layers with dropout and fully connected layers. Additionally, early stopping is implemented to enhance generalization and prevent overfitting. This architecture is designed to leverage the strengths of both CNN and BiLSTM, providing robust feature extraction and capturing long-term dependencies in sequential data.

The comprehensive approach in this study, including advanced preprocessing techniques like SMOTE for handling imbalanced data and sophisticated layer combinations, aims to push the boundaries of model performance. By addressing the limitations of previous architectures and introducing novel elements, this study seeks to achieve superior accuracy and reliability in text analysis tasks.

3. Methodology

This section will explain in detail the steps and methods used in creating a police report classification system using the CNN - BiLSTM architecture. Figure 1 illustrates the progression of this research.

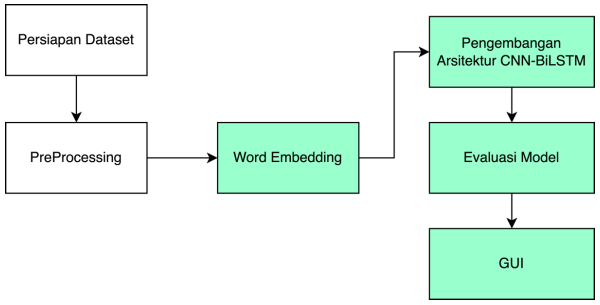

Figure 1. Research Flow

Here's an explanation of the course of this penalty methodology.

A. Dataset

The dataset used in this study includes public complaints collected through the police WhatsApp channel, specifically from Polda Jambi. In addition, data from Twitter is also included to provide a more comprehensive analysis. The Twitter data includes public tweets related to the same issues, providing a broader perspective on public sentiment and complaints. By combining these two sources, the study aims to provide a more accurate and holistic understanding of the issues reported, allowing for better analysis and insight into public concerns.

B. Text Preprocessing

The police complaint dataset that has been collected from the relevant regional police cannot be used directly. The data we have is still dirty and full of symbols, text and icons that are not needed or have no meaning in the further research process. Text preprocessing aims to prepare data so that it can be further processed in word embedding. The process or steps of text preprocessing can be seen in Figure 2.

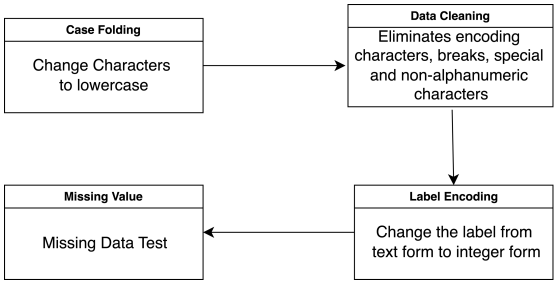

Figure 2. Pre-Processing Data

The diagram represents the preprocessing steps involved in preparing textual data for analysis. Initially, the text data undergoes case folding, which involves converting all characters to lowercase. This step ensures uniformity in the data by eliminating differences in case that do not affect the meaning of the text, such as treating "Text" and "text" as the same word [18]. Following case folding, data cleaning is performed to remove unwanted characters, including encoding characters, line breaks, special characters, and non-alphanumeric characters [19]. The purpose of this step is to eliminate noise that could interfere with the analysis, making the text more consistent and easier to process in subsequent steps.

After cleaning, label encoding is applied to transform categorical labels from text form to integer form [20]. This is essential because many machine learning algorithms require numerical input. By converting text labels to integers, the data becomes suitable for these algorithms. For instance, labels like "spam" and "not spam" might be converted to 0 and 1, respectively. Finally, a missing value check is conducted to identify and handle any missing data within the dataset. Missing data can pose problems for analysis and model training, so this step ensures that any missing values are appropriately addressed, either by filling them in or removing incomplete records, thereby maintaining the dataset's integrity and readiness for analysis [21].

C. Word Embedding

Word embedding is a technique that represents textual data as vectors of real numbers, enabling machines to understand and process natural language [22]. Experts have conducted research to evaluate the application and performance of word embeddings in various domains [23]. They used direct intrinsic word embedding evaluation tasks to assess the relatedness of philosophical terms, and they found that these tasks can be effective in evaluating word embeddings for specific languages [24]. In addition, experts have proposed methods to learn multiple vectors for each entity, capturing different aspects of the domain, and using a mixture of experts' formulations to jointly learn more specific word embeddings [25]. Another approach involves enriching pretrained word embeddings with domain-specific information, improving prediction accuracy and expert knowledge in classification tasks [26]. These studies highlight the importance of evaluating and improving word embedding to improve its performance and applicability in specific domains.

Word embedding in this research uses Word2Vec, which is a method for producing word vector representations from text [27]. This algorithm maps words into a high-dimensional vector space, where words with similar meanings or usage tend to have close vector representations [28].

The advantages of word embedding using Word2Vec are its ease of implementation and good performance for natural language processing tasks [29]. Word2vec has become one of the most commonly used methods for generating word representations in natural language processing and other related applications [30].

D. Algorithm Development Testing

Table 2 is a test of the model developed in this research.

TABLE II. TESTING ALGORITHM DEVELOPMENT

Dataset	Model
Public Report Public Complaint Dataset	CNN
	CNN + Early Stopping
	BiLSTM
	BiLSTM + Early Stopping
	Proposed Hybrid CNN - BiLSTM
	Proposed Hybrid CNN – BiLSTM + Early Stopping
Twitter Dataset	CNN
	CNN + Early Stopping
	BiLSTM
	BiLSTM + Early Stopping
	Proposed Hybrid CNN - BiLSTM
	Proposed Hybrid CNN – BiLSTM + Early Stopping

After testing, testing is then carried out for the highest accuracy to be used as a model for streamlit. Streamlit is a test that uses a simple GUI to test new data [31]. Figure 3 shows the proposed architecture carried out in this research.

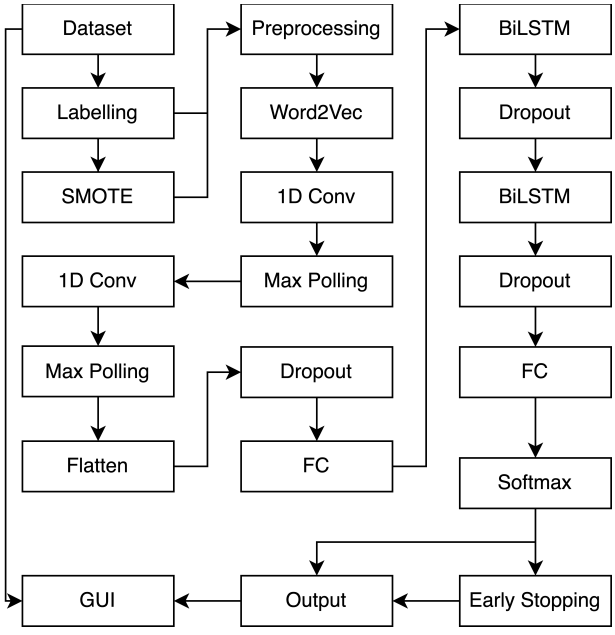

Figure 3. Development of CNN-BiLSTM Architecture

Using a combination of Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) for sequential data analysis, such as text, provides several significant advantages. CNN is effective at extracting local features from data, such as patterns or n-grams in text, by applying filters that detect various features at a local level [32]. This enables CNN to recognize important phrases or keywords. However, CNN has limitations in understanding long-term context due to its focus on localized data regions [33].

On the other hand, BiLSTM is designed to capture long-term dependencies in sequential data [34]. BiLSTM processes data in both forward and backward directions, allowing it to understand the full context of the sequence [35]. This is particularly useful in natural language processing (NLP), where the meaning of a word often depends on the context provided by surrounding words [36]. By combining CNN and BiLSTM, the model can leverage the strengths of both architectures: CNN for detecting local features and BiLSTM for capturing global context.

Adding multiple layers to this architecture aims to enhance the model's ability to capture and process more complex and abstract features from the data. In the case of CNN, additional convolutional layers allow the model to progressively recognize higher-level features, from simple patterns to more complex structures [37]. Pooling and dropout layers are added to reduce data dimensionality and overfitting, thereby improving the model's generalization [38]. In BiLSTM, adding layers strengthens the model's understanding of sequential context and captures longer and more complex dependencies within the data [39]. Fully connected (FC) layers are used to combine the extracted features and map them to the target classes, while the softmax layer ensures that the final predictions are presented as probabilities.

The primary strength of combining CNN and BiLSTM is the ability to merge local feature detection with long-term context understanding. CNN excels at capturing local patterns, while BiLSTM captures sequential information involving broader context [40]. This combination allows the model to comprehend data at various levels of complexity, which is crucial in text analysis. By harnessing the strengths of both architectures, the model can provide more accurate and reliable predictions, improving performance in tasks such as sentiment analysis, text classification, and other NLP tasks. Therefore, the combination of CNN and BiLSTM is often used to maximize analytical capabilities and enhance overall model performance.

4. RESULT AND DISCUSSION

The following are the results and discussions carried out in this research.

A. *Dataset Using SMOTE*

The dataset that has been taken from the Jambi regional police has obtained 5166 records, which are divided into 4 classifications. After classification, the dataset is unbalanced. Therefore, it is necessary to balance the data using SMOTE. The technique used is oversampling. This technique is used because minority data will be multiplied by the SMOTE algorithm so that the data will not be the same as the majority data [41]. We took this action to ensure the research did not destroy any existing data. Following the implementation of the class balancing process, the dataset expanded to 10296 records. Apart from that, this research also uses other data to determine the effectiveness of the deep learning model being developed. The developed model preprocesses the balanced data before testing it.

B. *Algorithm Testing*

Table 3 is a test carried out using data from public reports to the Jambi regional police.

TABLE III. TESTING WITH COMMUNITY REPORTED DATA

No	Model That Used	Accuracy
1	CNN	91%
2	CNN + Early Stopping	68%
3	BiLSTM	98%
4	BiLSTM + Early Stopping	73%
5	Proposed Hybrid CNN - BiLSTM	99%
6	Proposed Hybrid CNN – BiLSTM + Early Stopping	73%

Table 3 displays the testing results using data from community reports to the Jambi regional police. This evaluation compares various deep learning models, including CNN, BiLSTM, and a hybrid combination of CNN - BiLSTM, both with and without early stopping.

The results reveal that the hybrid CNN - BiLSTM model achieves the highest accuracy of 99%, highlighting the superiority of the hybrid approach over single models. Despite early stopping being intended to prevent overfitting, in this test, models with early stopping did not show significant accuracy improvements and, in fact, had lower accuracy compared to models without early stopping.

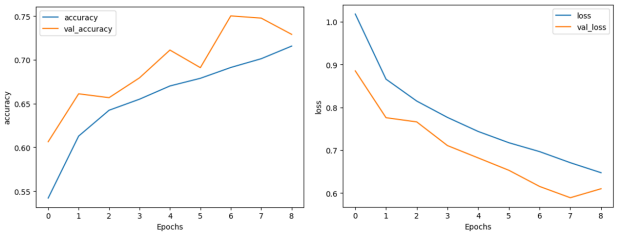

Figure 4. Plotting Results Using Early Stopping

Figure 4 illustrates that the epoch process halts at the 10th epoch due to early stopping. This indicates that continuing training beyond this point could lead to overfitting, which would degrade performance on the test data.

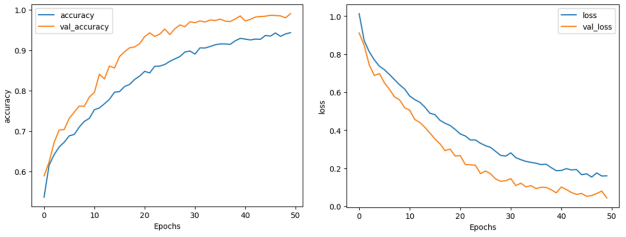

Figure 5. Plotting Results Without Using Early Stopping

Figure 5 shows the complete epoch process extending up to 50 epochs without early stopping. As a result, the accuracy reaches an impressive 99%, demonstrating that the model can effectively learn from the data without overfitting.

This explanation emphasizes that while early stopping can prevent overfitting, in this case, full training without early stopping yielded more optimal results. The hybrid CNN - BiLSTM approach proves to be exceptionally effective in handling community report data, delivering outstandingly high accuracy.

To evaluate the robustness of the proposed system, this study also utilized a different dataset, specifically tweets from Twitter. Table 4 shows the results of utilizing Twitter data.

TABLE IV. TESTING WITH TWITTER DATA

No	Model That Used	Accuracy
1	CNN	72%
2	CNN + Early Stopping	46%
3	BiLSTM	59%
4	BiLSTM + Early Stopping	47%
5	Proposed Hybrid CNN - BiLSTM	79%
6	Proposed Hybrid CNN – BiLSTM + Early Stopping	51%

From Table 4, it can be observed that the use of Twitter data in this study led to a significant decrease in accuracy compared to previous datasets. However, the hybrid model's accuracy remains the highest at 79%. This reduction in accuracy can be attributed to the high complexity of Twitter data, which frequently employs abbreviations and informal language. This contrasts with community report data, which is typically more formal and structured.

C. Comparison With Previous Research

To further contextualize the findings, a comparison with previous studies that employed hybrid CNN - BiLSTM models was conducted. Table 5 presents the comparative results.

TABLE V. COMPARISON

Research	Model	Dataset	Accuracy
[42]	CNN - BiLSTM	Review Product	88%
[43]	CNN - BiLSTM - TE	Dataset of Comment	93.73%
[44]	CNN - BiLSTM – ATT	The Stanford Sentiment Treebank (SST)	90,4%
[45]	CNN - BiLSTM	Twitter	85%
This Study	Proposed CNN - BiLSTM	Community Report	99%

Table 5 compares the performance of various deep learning models used for text analysis on different datasets. The referenced research highlights the efficacy of combining Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) models for sentiment analysis tasks. Study [38] achieved an accuracy of 88% using a CNN - BiLSTM model on a product review dataset. Study [39] combined CNN, BiLSTM, and Transformation Ensemble (TE) on a comment dataset, resulting in a higher accuracy of 93.73%. In the following study [40], the combination of CNN - BiLSTM with an Attention Mechanism (ATT) was applied to The Stanford Sentiment Treebank (SST) dataset, achieving an accuracy of 90.4%. Meanwhile, study [41] applied the CNN - BiLSTM model to a Twitter dataset, obtaining an accuracy of 85%. The current study, which employed the proposed CNN - BiLSTM model on community reports, demonstrated superior performance with an accuracy reaching 99%.

This comparison underscores the superior performance of the proposed hybrid model in handling community report data, achieving the highest accuracy among the datasets tested. The results validate the robustness and efficacy of the hybrid CNN-BiLSTM approach in processing diverse types of textual data with varying levels of complexity and formality.

This explanation illustrates that the proposed model in this study significantly outperforms previous models, showcasing its potential in improving sentiment analysis accuracy across various datasets. Following the testing phase using the proposed CNN-BiLSTM, the subsequent step involves automating the prototype process for detecting public complaints. The prototype is developed using a simple graphical user interface (GUI) with Streamlit. Figure 6 illustrates the detection of public complaints in the Indonesian language.

Figure 6. Results of Public Complaint Detection

The image labeled as Figure 6. "Results of Public Complaint Detection" depicts an example of a report detected by the system regarding a public complaint. Here is the explanation of the content within the image:

The title of the report is "Violation Detection In The Jambi Police." The body of the report includes a message from a user, identified by the handle @febi1231, who is reporting suspicious activities in their neighborhood. The message reads:

"@febi1231, I would like to report suspicious activities in my complex. Several individuals frequently enter and leave a house at 123 Merdeka Street. I suspect they are involved in drug dealing due to their unusual activities late at night. Please investigate this matter. Thank you. #StopNarkoba"

The report is marked with a detection tag indicating the nature of the complaint. In this case, it is labeled under "Detection" with a note specifying that the detection results categorize the report as a legal issue related to "Kriminal Narkoba" (Drug Crime).

The green text at the bottom confirms that the detection results of the report have been categorized appropriately, identifying the issue as related to drug-related criminal activities.

This figure illustrates the system's capability to identify and categorize public complaints about violations, specifically focusing on drug-related crimes in this instance.

5. CONCLUSION

The development of the CNN-BiLSTM architecture demonstrated superior performance in analyzing public complaints data submitted to the Indonesian National Police, achieving an accuracy of 99%. This significantly outperformed the accuracy obtained from tweet data. Furthermore, the BiLSTM algorithm attained a peak accuracy of 98%, whereas the CNN algorithm reached 91%. The implementation of early stopping effectively

prevented overfitting, contributing to the model's robustness. However, it did not yield significant improvements in accuracy for either the basic or hybrid algorithm developments in this study. These findings underscore the potential of the CNN-BiLSTM architecture in processing and categorizing public complaints with high precision. Future research should explore additional techniques to further enhance model accuracy and efficiency. The success of this model suggests its applicability in other domains requiring precise text classification and highlights the need for continuous refinement of deep learning models to address specific dataset characteristics. Further development of the proposed CNN-BiLSTM architecture is essential to address potential issues when applied to other datasets with complex characteristics. Future research should emphasize hyperparameter tuning and optimization methods, as these strategies are expected to enhance model accuracy. Furthermore, exploring alternative machine learning algorithms for object classification, beyond the softmax function used in this study, could yield effective solutions. Continuous refinement and the adoption of advanced techniques are crucial for improving the model's performance and ensuring its versatility across various datasets. This approach will not only bolster the model's robustness but also expand its applicability in diverse real-world scenarios, thereby maximizing its utility and effectiveness in different domains. Ultimately, these efforts will contribute to the advancement of deep learning methodologies and their implementation in complex data analysis tasks.

ACKNOWLEDGMENT

Thank you to the universitas Adiwangsa Jambi and the UPI YPTK Padang university for facilitating and marking this research.

REFERENCES

[1] I. H. Sarker, "Data Science and Analytics: An Overview from Data-Driven Smart Computing, Decision-Making and Applications Perspective," *SN Comput Sci*, vol. 2, no. 5, Sep. 2021, doi: 10.1007/s42979-021-00765-8.

[2] M. A. Khan *et al.*, "A deep learning-based intrusion detection system for mqtt enabled iot," *Sensors*, vol. 21, pp. 1–25, Nov. 2021, doi: 10.3390/s21217016.

[3] S. Agarwal, "Deep Learning-based Sentiment Analysis: Establishing Customer Dimension as the Lifeblood of Business Management," *Global Business Review*, vol. 23, no. 1, pp. 119–136, Feb. 2022, doi: 10.1177/0972150919845160.

[4] I. H. Sarker, "Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions," *SN Computer Science*, vol. 2, no. 6. Springer, Nov. 01, 2021. doi: 10.1007/s42979-021-00815-1.

[5] R. Senthilkumar, A. Bharathi, and S. D. M. Raja, "Discriminative convolution neural network architecture for diagnosis of diabetic retinopathy through classification and progression prediction of lesions grading in color fundus images of retina," *Int J Health Sci (Qassim)*, vol. 6, no. S2, pp. 10224–10242, May 2022, doi: 10.53730/ijhs.v6ns2.7731.

[6] A. A. I. A. Maharani, S. S. Prasetiyowati, and Y. Sibaroni, "Classification of Public Sentiment on Fuel Price Increases Using CNN," *Sinkron*, vol. 8, no. 3, pp. 1630–1637, Jul. 2023, doi: 10.33395/sinkron.v8i3.12609.

[7] R. H. Yahya, W. Maharani, and R. Wijaya, "Disaster Management Sentiment Analysis Using the BiLSTM Method," *JURNAL MEDIA INFORMATIKA BUDIDARMA* , vol. 7, no. 1, pp. 501–508, 2023, doi: 10.30865/mib.v7i1.5573.

[8] L. Xiaoyan, R. C. Raga, and S. Xuemei, "GloVe-CNN-BiLSTM Model for Sentiment Analysis on Text Reviews," *J Sens*, vol. 2022, pp. 1–12, 2022, doi: 10.1155/2022/7212366.

[9] M. Rhanoui, M. Mikram, S. Yousfi, and S. Barzali, "A CNN-BiLSTM Model for Document-Level Sentiment Analysis," *Mach Learn Knowl Extr*, vol. 1, no. 3, pp. 832–847, Sep. 2019, doi: 10.3390/make1030048.

[10] M. K. Anam, T. A. Fitri, Agustin, Lusiana, M. B. Firdaus, and A. T. Nurhuda, "Sentiment Analysis for Online Learning using The Lexicon-Based Method and The Support Vector Machine Algorithm," *ILKOM Jurnal Ilmiah*, vol. 15, no. 2, pp. 290–302, 2023, doi: 10.33096/ilkom.v15i2.1590.290-302.

[11] M. K. Anam, M. I. Mahendra, W. Agustin, Rahmaddeni, and Nurjayadi, "Framework for Analyzing Netizen Opinions on BPJS Using Sentiment Analysis and Social Network Analysis (SNA)," *Intensif*, vol. 6, no. 1, pp. 2549–6824, 2022, doi: 10.29407/intensif.v6i1.15870.

[12] M. Vilares Ferro, Y. Doval Mosquera, F. J. Ribadas Pena, and V. M. Darriba Bilbao, "Early stopping by correlating online indicators in neural networks," *Neural Networks*, vol. 159, pp. 109–124, Feb. 2023, doi: 10.1016/j.neunet.2022.11.035.

[13] G. Wiedemann, C. Biemann, E. Ruppert, and R. Jindal, "Transfer Learning from LDA to BiLSTM-CNN for Offensive Language Detection in Twitter," in *Conference on Natural Language*

Processing, 2018, pp. 85–94. doi: 10.48550/arXiv.1811.02906.

[14] R. Halder and R. Chatterjee, "CNN-BiLSTM Model for Violence Detection in Smart Surveillance," *SN Comput Sci*, vol. 1, no. 4, pp. 1–9, Jul. 2020, doi: 10.1007/s42979-020-00207-x.

[15] M. Lestandy and Abdurrahim, "Effect of Word2Vec Weighting with CNN-BiLSTM Model on Emotion Classification," *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, vol. 12, no. 1, pp. 99–107, Mar. 2023, doi: 10.23887/janapati.v12i1.58571.

[16] L. N. Aqilla, Y. Sibaroni, and S. S. Prasetiyowati, "Word2vec Architecture in Sentiment Classification of Fuel Price Increase Using CNN-BiLSTM Method," *Sinkron*, vol. 8, no. 3, pp. 1654–1664, Jul. 2023, doi: 10.33395/sinkron.v8i3.12639.

[17] H. Kim and Y. S. Jeong, "Sentiment classification using Convolutional Neural Networks," *Applied Sciences (Switzerland)*, vol. 9, no. 11, pp. 1–14, Jun. 2019, doi: 10.3390/app9112347.

[18] A. Angdresey and G. Saroinsong, "The Decision Tree Algorithm on Sentiment Analysis: Russia and Ukraine War," *Jurnal Sisfotenika*, vol. 13, no. 2, pp. 192–200, 2023, doi: 10.30700/jst.v13i2.1397.

[19] M. Gagolewski, "stringi: Fast and Portable Character String Processing in R," *J Stat Softw*, vol. 103, no. 2, 2022, doi: 10.18637/jss.v103.i02.

[20] M. K. Dahouda and I. Joe, "A Deep-Learned Embedding Technique for Categorical Features Encoding," *IEEE Access*, vol. 9, pp. 114381–114391, 2021, doi: 10.1109/ACCESS.2021.3104357.

[21] T. Emmanuel, T. Maupong, D. Mpoeleng, T. Semong, B. Mphago, and O. Tabona, "A survey on missing data in machine learning," *J Big Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-021-00516-9.

[22] D. S. Asudani, N. K. Nagwani, and P. Singh, "Impact of word embedding models on text analytics in deep learning environment: a review," *Artif Intell Rev*, vol. 56, no. 9, pp. 10345–10425, Sep. 2023, doi: 10.1007/s10462-023-10419-1.

[23] R. E. Ramos-Vargas, I. Román-Godínez, and S. Torres-Ramos, "Comparing general and specialized word embeddings for biomedical named entity recognition," *PeerJ Comput Sci*, vol. 7, pp. 1–22, 2021, doi: 10.7717/peerj-cs.384.

[24] A. Onan and M. A. Toçoğlu, "Weighted word embeddings and clustering-based identification of question topics in MOOC discussion forum posts," *Computer Applications in Engineering Education*, vol. 29, no. 4, pp. 675–689, Jul. 2021, doi: https://doi.org/10.1002/cae.22252.

[25] M. Esposito, E. Damiano, A. Minutolo, G. De Pietro, and H. Fujita, "Hybrid query expansion using lexical resources and word embeddings for sentence retrieval in question answering," *Inf Sci (N Y)*, vol. 514, pp. 88–105, 2020, doi: https://doi.org/10.1016/j.ins.2019.12.002.

[26] S. Wang, W. Zhou, and C. Jiang, "A survey of word embeddings based on deep learning," *Computing*, vol. 102, no. 3, pp. 717–740, 2020, doi: 10.1007/s00607-019-00768-7.

[27] M. A. H. Wadud, M. F. Mridha, and M. M. Rahman, "Word Embedding Methods for Word Representation in Deep Learning for Natural Language Processing," *Iraqi Journal of Science*, vol. 63, no. 3, pp. 1349–1361, 2022, doi: 10.24996/ijs.2022.63.3.37.

[28] P. G. Shivakumar and P. Georgiou, "Confusion2Vec: Towards enriching vector space word representations with representational ambiguities," *PeerJ Comput Sci*, vol. 2019, no. 6, 2019, doi: 10.7717/peerj-cs.195.

[29] A. Desai *et al.*, "Word2vec Word Embedding-Based Artificial Intelligence Model in the Triage of Patients with Suspected Diagnosis of Major Ischemic Stroke: A Feasibility Study," *Int J Environ Res Public Health*, vol. 19, no. 22, Nov. 2022, doi: 10.3390/ijerph192215295.

[30] S. Al-Saqqa and A. Awajan, "The Use of Word2vec Model in Sentiment Analysis: A Survey," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Dec. 2019, pp. 39–43. doi: 10.1145/3388218.3388229.

[31] N. Aendikov and A. Azayeva, "Integration of GIS and machine learning analytics into Streamlit application," in *Procedia Computer Science*, Elsevier B.V., 2024, pp. 691–696. doi: 10.1016/j.procs.2023.12.160.

[32] S. Soni, S. S. Chouhan, and S. S. Rathore, "TextConvoNet: a convolutional neural network based architecture for text classification," *Applied Intelligence*, vol. 53, no. 11, pp. 14249–14268, Jun. 2023, doi: 10.1007/s10489-022-04221-9.

[33] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J Big Data*, vol. 8, no. 1, pp. 1–72, Dec. 2021, doi: 10.1186/s40537-021-00444-8.

[34] S. Arslan, "Application of BiLSTM-CRF model with different embeddings for product name extraction in unstructured Turkish text," *Neural*

Comput Appl, vol. 36, no. 15, pp. 8371–8382, May 2024, doi: 10.1007/s00521-024-09532-1.

[35] M. Marjani, M. Mahdianpari, and F. Mohammadimanesh, "CNN-BiLSTM: A Novel Deep Learning Model for Near-Real-Time Daily Wildfire Spread Prediction," *Remote Sens (Basel)*, vol. 16, no. 8, pp. 1–17, Apr. 2024, doi: 10.3390/rs16081467.

[36] R. Egger and E. Gokce, "Natural Language Processing (NLP): An Introduction: Making Sense of Textual Data," in *Tourism on the Verge*, vol. Part F1051, Springer Nature, 2022, pp. 307–334. doi: 10.1007/978-3-030-88389-8_15.

[37] X. Zhao, L. Wang, Y. Zhang, X. Han, M. Deveci, and M. Parmar, "A review of convolutional neural networks in computer vision," *Artif Intell Rev*, vol. 57, no. 4, pp. 1–43, Apr. 2024, doi: 10.1007/s10462-024-10721-6.

[38] I. Salehin and D. K. Kang, "A Review on Dropout Regularization Approaches for Deep Neural Networks within the Scholarly Domain," *Electronics (Switzerland)*, vol. 12, pp. 1–23, Jul. 2023, doi: 10.3390/electronics12143106.

[39] B. Abimbola, E. de La Cal Marin, and Q. Tan, "Enhancing Legal Sentiment Analysis: A Convolutional Neural Network–Long Short-Term Memory Document-Level Model," *Mach Learn Knowl Extr*, vol. 6, no. 2, pp. 877–897, Apr. 2024, doi: 10.3390/make6020041.

[40] R. Bharal and O. V Vamsi Krishna, "Social Media Sentiment Analysis Using CNN-BiLSTM," *International Journal of Science and Research*, vol. 10, no. 9, pp. 656–661, 2020, doi: 10.21275/SR21913110537.

[41] J. H. Joloudari, A. Marefat, M. A. Nematollahi, S. S. Oyelere, and S. Hussain, "Effective Class-Imbalance Learning Based on SMOTE and Convolutional Neural Networks," *Applied Sciences (Switzerland)*, vol. 13, no. 6, Mar. 2023, doi: 10.3390/app13064006.

[42] S. Susandri, S. Defit, and M. Tajuddin, "Enhancing Text Sentiment Classification with Hybrid CNN-BiLSTM Model on WhatsApp Group," *Journal of Advances in Information Technology*, vol. 15, no. 3, pp. 355–363, 2024, doi: 10.12720/jait.15.3.355-363.

[43] Y. Zeng, R. Zhang, L. Yang, and S. Song, "Cross-Domain Text Sentiment Classification Method Based on the CNN-BiLSTM-TE Model," *Journal of Information Processing Systems*, vol. 17, no. 4, pp. 818–833, Aug. 2021, doi: 10.3745/JIPS.04.0221.

[44] L. Deng, T. Yin, Z. Li, and Q. Ge, "Analysis of the Effectiveness of CNN-LSTM Models Incorporating Bert and Attention Mechanisms in Sentiment Analysis of Data Reviews," in *ICBDIE*, 2024, pp. 821–829. doi: 10.2991/978-94-6463-238-5_106.

[45] Abdurrahim and D. H. Fudholi, "Kinetik: Game Technology, Information System," *Mental Health Prediction Model on Social Media Data Using CNN-BiLSTM*, vol. 9, no. 1, pp. 29–44, 2019, doi: 10.22219/kinetik.v9i1.1849.

Ade Oktarino is a 2010 alumni of the S1 Informatics Engineering degree at Surakarta University, a 2013 graduate of the Masters in Information Systems at the Dinamika Bangsa Jambi College of Computer Science and is currently pursuing the Doctoral Information Technology Program at Putra Indonesia University YPTK Padang in In 2022, previously served as a permanent lecturer at Jambi Polytechnic until 2014 and is now actively teaching at Adiwangsa University Jambi in the Information Technology Study Program, Faculty of Engineering and Computer Science. Previously, Lecturer on Additional Duties as Deputy Chancellor II of Adiwangsa University for the period 2017 – 2023 and currently Lecturer on Additional Duties as Dean of the Teaching and Education Faculty, Adiwangsa University, Jambi. Active in the Jambi Province ICT Volunteer organization as one of the administrators and also active as an IT consultant at the Jambi Regional Police, Jambi City Environmental Service and Jambi Provincial Immigration Office. Research concentration on the topics of artificial intelligence, machine learning, deep learning in the fields of computer vision and natural language processing.

Sarjon Defit received the B.Sc. degree in computer science from Putra Indonesia University YPTK Padang, Indonesia, the M.Sc. degree in computing from the University Technology Malaysia., and the Ph.D. degree in computing science from The University Technology Malaysia. He is currently a Professor at Putra Indonesia University YPTK Padang (UPI YPTK Padang). He is actively involved as a lecturer in the Information Technology Doctoral Program at UPI YPTK Padang. His research focus lies in the field of data science.

Yuhandri holds the distinguished title of Professor at Universitas Putra Indonesia YPTK Padang, contributing as a lecturer in the information technology study program within the Faculty of Computer Science. He earned his undergraduate degree (S1) in the computer systems study program and completed his Master's (S2) in the informatics engineering study program at UPI YPTK Padang, West Sumatra, Indonesia. Additionally, he obtained his doctoral degree from Gunadharma University, Jakarta, Indonesia. Presently, Professor Yuhandri is recognized as a leading expert in the field of image processing at Universitas Putra Indonesia YPTK Padang. His pedagogical repertoire encompasses courses in the information technology study program, encompassing subjects like image processing, artificial intelligence, computer networks. For inquiries, Professor Yuhandri can be reached via email at yuyu@upiyptk.ac.id.