Content-Based Filtering in Recommendation Systems Culinary Tourism Based on Twitter (X) Using Bidirectional Gated Recurrent Unit (Bi-GRU)

Adhyasta Naufal Faadhilah, Erwin Budi Setiawan
Telkom University, Jl. Terusan Buah Batu, Bandung 40257, Indonesia

ARTICLE INFO

Article history:

Received June 06, 2024
Revised July 04, 2024
Published July 15, 2024

Keywords:

Content-Based Filtering;
Bidirectional Gated Recurrent Unit;
Culinary Recommendation System;
TF-IDF;
Twitter (X)

ABSTRACT

To address the challenge of information overload in the rapidly expanding culinary sector, a recommendation system using Content-Based Filtering (CBF) and the Bidirectional Gated Recurrent Unit (Bi-GRU) algorithm was developed. This system can help users to suggest culinary options based on user profiles and preferences. Twitter (X) is frequently used to gather culinary reviews in Bandung, forming the foundation for developing recommendation systems. This research contributes to integrating CBF and Bi-GRU to enhance the relevance of culinary recommendations. The system uses Term Frequency-Inverse Document Frequency (TF-IDF) for feature extraction and Cosine Similarity for item matching. Research adapting CBF and Bi-GRU methods specifically for culinary recommendations, especially in Bandung, remains limited. This study focuses on evaluating the performance of a culinary recommendation system. Data collected from Twitter (X) and PergiKuliner includes 2,645 reviews from 44 Twitter (X) accounts and on 200 culinary places. The culinary recommendation model, using CBF with TF-IDF and Cosine Similarity, achieved a Mean Absolute Error (MAE) of 0.254 and Root Mean Square Error (RMSE) of 0.425, indicating high accuracy in rating predictions compared to previous studies. From the experiments conducted, the third experiment using Bi-GRU, SMOTE, and the Nadam algorithm showed the best improvement with a learning rate of 0.014563484775012459, achieving an accuracy of 86.8%, precision of 86.3%, recall of 85.2%, and an F1-Score of 85.5%, with a 16.2% increase in accuracy from the baseline. Thus, this system effectively helps users with culinary recommendations in Bandung, providing good performance based on user preferences.

Corresponding Author:

Erwin Budi Setiawan, Telkom University, Jl. Terusan Buah Batu, Bandung 40527, Indonesia
Email: erwinbudisetiawan@telkomuniversity.ac.id

1. INTRODUCTION

The development of social media has increased user access to various information. Platforms like Twitter (X), with over 336 million active users per month and 500 million tweets per day, play a significant role in recommending culinary tours in Bandung [1]. Users can find information about culinary offerings through tweets or reviews in Bandung [2]. Additionally, according to the Badan Pusat Statistika (BPS), the restaurant and café sector has been growing by approximately 15% annually, with the highest increase recorded in 2018, reaching up to 1,000 new restaurants and cafes in Bandung based on data from 2013 to 2021 [3]. The culinary sector in the city is experiencing rapid growth, posing a challenge for visitors to choose culinary options that match their preferences amidst abundant information.

To address this challenge, this study utilizes a dataset of tweets about culinary offerings from PergiKuliner, which includes ratings, types of cuisine, descriptions, and prices [4]. The aim is to develop a culinary recommendation system that facilitates users in finding suitable culinary options. The recommendation system emerges as an optimal solution by identifying information, products, or services that align with the user's preferences [5]. The recommendation system itself is divided into three types that have been identified. One of the types used in this study is Content-Based Filtering (CBF) [6]. CBF is chosen over Collaborative Filtering (CF) due to potential issues such as cold start and data sparsity when dealing with new users who have not yet rated items [7]. CBF is a recommendation system method that utilizes user profiles to suggest new items similar to those previously liked or visited by the user [8]. One of its primary advantages lies in providing more personalized recommendations [9], [10]. This study employs a CBF approach using the Term Frequency-Inverse Document Frequency (TF-IDF) algorithm for term weighting and Cosine Similarity to measure the similarity of these values [11].

This research [12] utilized TF-IDF and Cosine Similarity in Content-Based Filtering for recommending movies to find similarities between movies based on their synopsis, title, and genres for features, achieving an accuracy of 82.32% for single queries and 75.05% for multiple seed queries. The subsequent research [13] focuses on providing movie recommendations using Content-Based Filtering with a TF-IDF matrix. The dataset used comes from Netflix and Disney+ with a total of 854 movies and 34,086 film reviews from Twitter (X) to find similarities based on description and genre features. The evaluation results of this method show an MAE of 0.28 and an RMSE of 0.67. In the following study, the research [14] focuses on conducting a deep sentiment analysis on Twitter (X) using the Bi-GRU model. In this study, we developed and compared several models such as MLP, CNN, LSTM, GRU, and Bi-GRU using 20% of the Twitter dataset for training and testing. The evaluation results show that the Bi-GRU model achieved a relatively high accuracy of 82.04% compared to the GRU, MLP, LSTM, and CNN models. The classification evaluation metrics with Word2Vec obtained are Precision: 86%, F1-Score: 87%, and Recall: 87%. This approach has proven effective in improving accuracy across various domains. For instance, previous research [15] showcased that their Bi-GRU model achieved an accuracy rate of 98.55% in biometric electrocardiogram classification, compared to the unidirectional GRU model, which only reached an accuracy of 91.2%. Based on previous research, the use of Bi-GRU is beneficial for classification models because of its better results despite the increased complexity compared to GRU and LSTM, because Bi-GRU captures two-way semantic dependencies by utilizing past and future contexts. [16]–[18].

This research contributes to the development of a culinary tourism recommendation model using Content-Based Filtering (CBF) integrated with the Bidirectional Gated Recurrent Unit (Bi-GRU) algorithm. The study focuses on evaluating the performance of a culinary recommendation system in Bandung, built upon the Term Frequency-Inverse Document Frequency (TF-IDF) and Cosine Similarity algorithms for feature extraction and item matching, enhanced by Bi-GRU to improve classification accuracy. To the author's knowledge, research specifically adapting the CBF and Bi-GRU methods to the culinary recommendation domain, particularly in Bandung, is still limited. This study utilizes optimization algorithms to achieve optimal predictions in recommending dishes aligned with user preferences. The primary motivation is to enhance the accuracy of culinary recommendations by integrating classification methods and semantic feature extraction. The combined approach of CBF and Bi-GRU in the context of culinary tourism represents a significant contribution, applying previous research findings to develop a personalized, relevant, and high-performance culinary recommendation system.

This research is organized into several sections. Section 2 provides a summary of related studies that explore similar research topics and includes sub-sections pertinent to system design. Section 3 delves into the proposed methodology, with a primary focus on developing a recommendation system using Content-Based Filtering and Bi-GRU classification, and includes various experiments as part of the discussion. Finally, Section 4 presents the development of the system using the proposed methods and classification techniques, along with the accuracy results obtained.

2. METHODS

This research designs a culinary tourism recommendation system using Content-Based Filtering (CBF) with feature extraction via TF-IDF and Cosine Similarity to calculate similarity, combined with Bidirectional Gated Recurrent Unit (Bi-GRU). The design steps for the system can be seen in Fig. 1.

2.1. Crawling Data

The dataset collection involved gathering reviews and tweets about culinary tourism in Bandung from social media platforms Twitter (X) and PergiKuliner. Twitter (X) offers an Application Programming Interface (API) to facilitate data retrieval, accessible by applying for a developer account [19]. The initial data collection used the Tweet-Harvest library to obtain user IDs, culinary places, and tweets, totaling 2,645 reviews from 39 accounts, along with 5 rating accounts from GrabFood, Gojek, TripAdvisor, PergiKuliner, and Google Maps, resulting in a total of 44 accounts. To ensure the representativeness of the first dataset, tweets were crawled using specific keywords related to culinary topics from the second dataset. However, we acknowledge limitations in selecting relevant accounts from diverse demographic backgrounds. The first dataset results are presented in Table 1. The second dataset was obtained from PergiKuliner, encompassing information about 200 culinary places in Bandung, including names, types of cuisine, addresses, ratings, descriptions, and price ranges. The data collected from PergiKuliner is shown in Table 2. Subsequently, both datasets were compiled into Comma Separated Values (CSV) files.

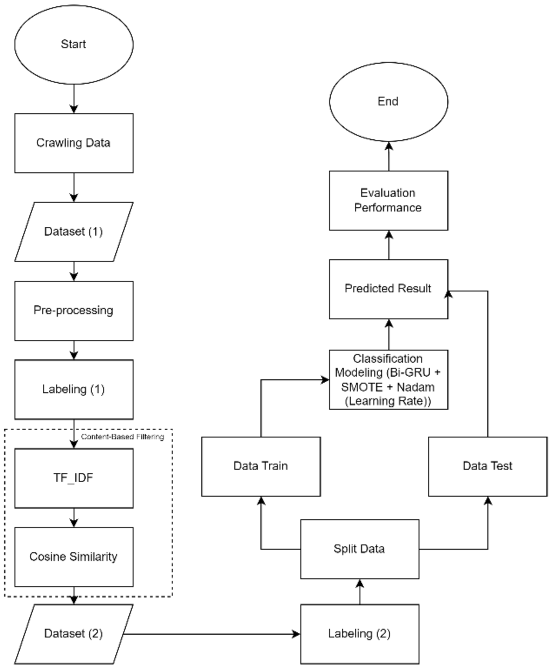

Fig. 1. Flowchart Design Culinary Recommendation System

Table 1. Example of First Crawling Data

Culinary	Cuisine	Price	Description
150 Coffee and Garden	Cafe	Rp 50.000 – Rp 100.000	150 Coffee and Garden is a café located in Taman Hutan Raya Ir. H. Djuanda. It offers a tranquil atmosphere with beautiful forest views, serving a variety of coffee, tea, and light snacks.
Yoshinoya	Japan	< Rp 50.000	Yoshinoya is the perfect choice for those who want to enjoy authentic Japanese flavors at affordable prices. The restaurant offers a variety of delicious gyudon dishes in generous portions.

Table 2. Example of Second Crawling Data

Username	Culinary	Tweets
aarrddyyee_95	de.u Coffee	@juaramageran Bandung mah banyak. Ini aku kasih versi kopi susu yg enakâ€ di bandung - pintu biru - makmur jaya - kopi cante - truk kuning - jati kopi - kata anja - kino Kimi backyard - De.u - kopi senin
papapicko	Lomie dan Bakmie Lombok	@_anggihasibuan_ Cobain Lomie lombok/imam bonjol deket DU sama soto Bandung Pak Simon/Ojolali di Sudirman kuahnya bening gurih ringan sedep

2.2. Data Preprocessing

Before being integrated into the system, the collected dataset needs to undergo preprocessing [20]. Data preprocessing aims to resolve issues with raw data that may be incorrect, incomplete, or otherwise unsuitable, ensuring the resulting dataset is useful [21]. During this process, text cleaning is essential by removing characters such as "RT", emojis, hashtags, links, and alphanumeric symbols [22]. This step is crucial because these results will later undergo a labeling assessment. The steps in data preprocessing include:

- **Data Cleaning:** This step removes punctuation, numbers, URLs, and hashtags from tweet sentences as they do not influence the document's information content.
- **Case Folding:** This process converts all text to lowercase to ensure data consistency.
- **Stop Word Removal:** This step eliminates irrelevant words or frequently occurring words that do not add meaningful value.
- **Stemming:** This process reduces words to their root form by removing prefixes and suffixes.
- **Tokenizing:** This step splits text into smaller units or tokens using punctuation marks as separators.

2.3. Labeling

In the labeling stage, as shown in Fig. 1, the process is conducted twice. Labeling 1 involves processing data by incorporating ratings from culinary reviews on Twitter (X). The culinary reviews from Twitter (X) in Dataset 1 are assigned ratings ranging from 0 to 5. This process begins with using the Python Deep Translator library to translate the culinary review data from user tweets from Indonesia into English using Google Translator into a rating. Once the data is translated, the research employs the Python TextBlob Library to calculate the Polarity Score for each tweet. A polarity score close to 0 indicates a negative sentiment, while a score close to 5 indicates a positive sentiment [23]. TextBlob is a Python library that calculates the polarity score after the dataset is automatically, which is then converted into a rating on a scale of 0-5 [24]. Labeling 1 generates data with 1663 rows, each with a Polarity Score. In Labeling 2, the predicted ratings from CBF on Dataset 2 are converted into labels, with 0 indicating not recommended and 1 indicating recommended.

2.4. Content-Based Filtering

The Content-Based Filtering (CBF) approach is a recommendation method where the items recommended are similar to those the user is interested in, by matching information between the user's profile and item attributes to generate relevant recommendations [25]. To build user profiles in CBF, this research gathers user preference data regarding types of cuisine, culinary descriptions, and price ranges of culinary places. Item descriptions are needed to identify and recommend similar items to users [26]. Once all features are collected, data is extracted from the available culinary place reviews to compare with the descriptions and attributes of culinary places using TF-IDF. This comparison yields similarity scores, with the highest-scoring culinary places being recommended to users through top-n recommendations.

2.5. Term Frequency – Inverse Document Frequency (TF-IDF)

Term Frequency – Inverse Document Frequency (TF-IDF) is a widely used numerical method in information retrieval to measure the importance of a term within one of many documents [27]. This process is used because the algorithm combines the frequency of a word's occurrence in a document (Term Frequency) with the inverse of the document frequency (Inverse Document Frequency), which measures how rarely the word appears across the document set [28]. TF-IDF aims to count the word occurrences and their frequency, with the results used to determine the most relevant keywords in a document [29]. This research uses TF-IDF for the extraction of text features into numerical vectors, producing a matrix that reflects the weight values of each word in the text according to the dataset's size [30]. The formula for TF-IDF is shown in equation (1):

$$W_{i,j} = tf_{i,j} \times \log(\frac{N}{df_i}) \quad (1)$$

In equation (1), where $tf_{i,j}$ is the frequency of word i in document j, df_i is the number of documents containing the word i and N is the total number of documents. The formula combines the frequency of a word's occurrence in a document with the weight of words that rarely appear across the entire collection to reflect the importance of the word [31]. After calculating TF-IDF for the document set, the next step is to compute Cosine Similarity to measure the similarity between documents. Cosine Similarity assesses the degree of similarity between two vectors in the inner product space in text classification [32]. This approach is highly effective because it efficiently and accurately captures the similarity patterns between documents.

2.6. Bidirectional Gated Recurrent Unit (Bi-GRU)

A Bidirectional Gated Recurrent Unit (Bi-GRU) is a sequence processing model that consists of two GRUs, one processing the input in a forward direction and the other in a backward direction. This bidirectional structure allows the model to capture information from both past and future contexts, which is highly beneficial for tasks involving sequential data [33], [34]. As shown in Fig. 2, the Bi-GRU structure combines two GRUs operating in opposite directions, both receiving the same input. The hidden states of two GRU inputs are calculated, and the cell output is determined by combining these hidden states [35]. This model is built using TensorFlow, a framework developed by Google for machine learning needs because it can generalize vectors and matrices to various dimensions [36]. At the Bi-GRU approach labels are set to 1 if the value is greater than 2.5, and 0 if less than 2.5. Additionally, the model compared with ratios 10%, 20%, 30%, and 40%. Includes a dropout layer with a rate of 0.3 to counter overfitting, a dense layer with a 1D sigmoid activation commonly used in neural networks for computation, batch normalization with a size of 32 for input normalization, and it is trained for 50 epochs [37]. Below is an illustration of the basic architecture of Bi-GRU:

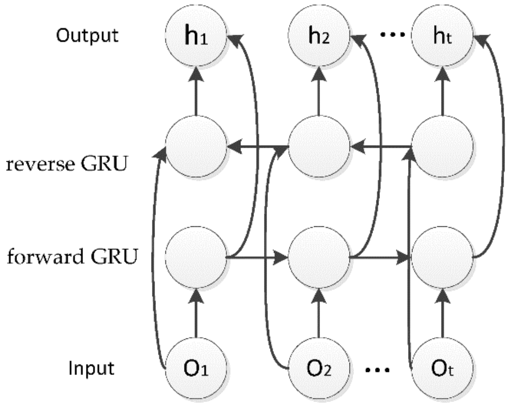

Fig. 2. Bi-GRU Model Architecture

2.7. Evaluation

In this research, model evaluation is conducted using various metrics to ensure optimal performance in the Content-Based Filtering recommendation and classification system with Bi-GRU. The recommendation system model is evaluated using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) metrics for Content-Based Filtering (CBF) [38], [39]. The MAE evaluation measures how close the model's predictions are to the actual values. Lower MAE values indicate better model performance in making predictions that are close to the actual values. RMSE, on the other hand, measures the accuracy of the model's predictions. A lower RMSE indicates more accurate model predictions. The use of this approach has resulted in evaluation outcomes for MAE and RMSE that are better than those in previous research. The formulas for MAE and RMSE are shown in equations (2) and (3):

$$MAE = \frac{1}{n}\sum_{i-1}^{n} |yi - xi| \tag{2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i-1}^{n} (yi - \hat{y}i)^2} \tag{3}$$

In equation (2), yi represents the actual value, xi is the predicted value, and n is the total number of predictions. In equation (3), yi is the actual value, $\hat{y}i$ is the predicted value, and n is the total number [40], [41]. After employing the Bi-GRU classification, the evaluation utilizes a confusion matrix to visualize the algorithm's classification performance. The confusion matrix is a table used to evaluate the performance of systems, including True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) values [42]. This matrix compares the predicted results with the actual classifications, providing a comprehensive overview of the system's performance. Using the confusion matrix, we can calculate evaluation metrics such as accuracy, F1-Score, precision, and recall, which assess the model's validity. The evaluation performance results obtained are slightly lower compared to the previous study, with only a few percentage points difference. However, the method used offers computational efficiency and ease of implementation without the need for embedding. This evaluation is the final process to determine the performance results of the prediction system [43], [44]. The formulas for these performance metrics can be seen in equations (4), (5), (6), and (7):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

$$Accuracy = 2\frac{Recall \times Precision}{Recall + Precision} \tag{7}$$

3. RESULTS AND DISCUSSION

3.1. Data Preparation

In the data preparation process for the content-based culinary tourism recommendation system (Content-Based Filtering), the first step is to prepare the "culinary dataset" containing information on culinary places in Bandung from the PergiKuliner website. The information used includes culinary places, types of cuisine, price ranges, ratings, and descriptions for a total of 200 entries. The next dataset consists of culinary tourism reviews that have been crawled in the form of tweets from Twitter (X) with library Tweet Harvest found by keywords specific to data, based on the names of the culinary places in the "culinary dataset". This process yielded relevant accounts, totaling 2,645 reviews from 39 accounts, along with 5 rating accounts from GrabFood, Gojek, TripAdvisor, PergiKuliner, and Google Maps, resulting in a total of 44 collected accounts.

The next step is preprocessing, which involves removing unwanted characters such as "RT," emojis, mentions, hashtags, links, non-alphanumeric characters, and new lines. Additionally, the text is converted to lowercase for consistency. This process is essential for selecting pertinent information and ensuring accuracy in subsequent processing steps. After cleaning, the text is translated into English using Google Translator from the deep_translator library to facilitate rating analysis, as shown in Table 3. Sentiment analysis is then performed by calculating the Polarity Score of the translated culinary tourism reviews using TextBlob automatically. This Polarity Score is normalized on a scale of 0-5, as shown in Table 4. This process is applied to each row of data, including text cleaning, translation, and polarity score calculation, with the results stored in a new data frame.

Table 3. Cleaned dataset

Username	Culinary	Cleaned Text
aarrddyyee_95	de.u Coffee	bandung, there's a lot of this, i'll give you a delicious version of coffee with milk in bandung, blue doors, makmur jaya, cante coffee, yellow truck, jati coffee, kata anja, kino kimi backyard de u, monday coffee
….	….	….
papapicko	Lomie dan Bakmie Lombok	try lomie lombok, imam bonjol, near du and soto bandung, sir simon ojolali in sudirman, the soup is clear, delicious, light and delicious

Table 4. Polarity Score

Username	Culinary	Polarity Score
aarrddyyee_95	de.u Coffee	3.333333
….	….	….
papapicko	Lomie dan Bakmie Lombok	3.8

The preprocessed data is exported to a CSV file, which is then downloaded for further analysis. Next, the processed data is grouped by account name and culinary place, and the average Polarity Score is calculated for each group. The final results are exported to a new CSV file, which is also downloaded. This data preparation process ensures that the data used for the content-based recommendation system is properly cleaned, translated, and processed, allowing the recommendation system to provide accurate and relevant results.

3.2. Content-Based Filtering Result

Before reaching the Content-Based Filtering step, this research includes several preliminary processes. The dataset is checked for missing values, and an empty template dataset is created with columns that will be filled with predicted ratings from the recommendation system. The datasets are then merged by combining the ratings as values with the empty template.

In the Content-Based Filtering step, the culinary dataset is reorganized with relevant columns such as description, type of cuisine, and price range, which are combined into a single content feature for calculating item similarity. For text preprocessing, the study employs a TF-IDF vectorizer to convert textual features into numerical vectors, facilitating content similarity analysis across features like culinary places, types of cuisine, price ranges, ratings, and descriptions from the "culinary dataset". Item similarity is computed using Cosine Similarity based on the TF-IDF matrix, resulting in a similarity matrix used to provide recommendations, as shown in Table 5. However, TF-IDF has limitations as it cannot capture the meaning of words and is highly dependent on the data used [45]. Culinary recommendations are made by selecting the top-n most similar items based on the similarity matrix to generate predicted ratings. Predicted ratings are calculated for each user and item without a rating, using the similarity matrix results, as shown in Table 6. This involves iterating through each user and item and using highly similar items to estimate ratings, thereby filling in missing values in the dataset and allowing the model to provide more accurate recommendations. The evaluation results indicate that the model has a Mean Absolute Error (MAE) with a small average prediction error of 0.254 and a Root Mean Square Error (RMSE) with a moderate prediction error rate of 0.425.

Table 5. Cosine Similarity Matrix TF-IDF Result

	150 Coffee and Garden	Ambrogio Patisserie	….	Yoshinoya
150 Coffee and Garden	1.000000	0.036612	….	0.061429
Ambrogio Patisserie	0.036612	1.000000	….	0.064436
….	….	….	….	….
Yoshinoya	0.061429	0.064436	….	1.000000

Table 6. CBF TF-IDF Predictions Result

Name	BaseBDG	….	grabfood
150 Coffee and Garden	3.37	….	4.2
….	….	….	….
Yoshinoya	2.98	….	4.8

3.3. Classification Result

In this classification process, Dataset 2, which contains predictions from Content-Based Filtering (CBF), is used. This dataset is then converted into binary labels (0 and 1). Labels are set to 1 if the value is greater than 2.5, and 0 if less than 2.5. This conversion simplifies the classification process. The total number of labels with a value of 0 is 7,431, and the total with a value of 1 is 1,569. To address this imbalance, SMOTE (Synthetic Minority Over-sampling Technique) is used because it can improve classification performance [46]. After conversion, Dataset 2 is used for the Bi-GRU classification model, chosen for its ability to handle sequential data and make accurate predictions. To make predictions, the research explores three scenarios. The first scenario applies SMOTE to the baseline to handle unbalanced data by performing a split ratio to determine optimal results. The second scenario uses SMOTE on the most optimal baseline result and adds optimization. The third scenario uses SMOTE on the most optimal baseline result and adds optimization with learning rate tuning. The classification stage will employ the Bi-GRU classification method.

3.3.1. Bi-GRU Baseline Model

In the first scenario, the Bi-GRU model is applied to each column in the "Dataset_CBF_Klasifikasi" data frame with labels converted to binary format (0 and 1). Each iteration involves separating features and labels, where features are taken from columns other than the one being tested. The data is then split into training and testing sets with stratification to ensure balanced class distribution. The training set uses the Synthetic Minority Over-sampling Technique (SMOTE) to address class imbalance by creating synthetic examples of the minority class [47], [48]. Next, the Bi-GRU model is built, followed by a dropout layer of 0.3 and a dense layer with sigmoid activation for binary prediction.

The model is trained for 50 epochs with a batch size of 32, comparing ratios of 90:10, 80:20, 70:30, and 60:40, and its performance is evaluated using the test set. According to Table 7, the highest accuracy of 74.7% is achieved with a test size of 20%. After training, predictions are made and evaluated using additional metrics such as Accuracy, Precision, Recall, and F1-Score. The most optimal results, with a test size of 20%, are saved for further analysis and will be used in the next scenario to provide insights into the effectiveness of the Bi-GRU model in classifying the culinary recommendation system data.

Table 7. Performance of Baseline Model Result

Test Size (%)	Performance Metrics			
	Accuracy	Precision	Recall	F1-Score
10	74.2%	91.4%	73%	75.2%
20	**74.7%**	**89.7%**	**76%**	**81%**
30	74.2%	90%	75%	81%
40	74.3%	89.7%	74%	79%

3.3.2. Bi-GRU Model Optimization

In the second scenario, the model uses a test size of 20%, based on the first scenario where a 20% test size yielded the most optimal results. This scenario also employs SMOTE and optimization. Optimization in classification plays a crucial role as it can enhance performance [49], [50]. This experiment aims to determine the performance improvement after adding optimization and to identify which optimization method works best for this model. The optimizers used in this experiment are Adam and Nadam, both set with a default learning rate parameter of 0.001. The results of this scenario can be seen in Table 8.

Table 8. Performance of Each Optimizer Result

Optimizer	Performance Metrics			
	Accuracy	Precision	Recall	F1-Score
Adam	80 (+7%)	89.7 (+0%)	82 (+7.9%)	84.8 (+4.7%)
Nadam	**80.2 (+7.3%)**	**89 (-0.7%)**	**80 (+5.2%)**	**83.4 (+3%)**

Based on the results in Table 8, each optimization shows a significant increase in accuracy from the baseline after using the default learning rate. Table 8 indicates that the Nadam algorithm with a default learning rate of 0.001 achieves the highest accuracy of 80.2%, showing an improvement of 7.3% over the baseline. In this experiment, the Nadam optimization proves to be more effective, demonstrating the highest performance on the Bi-GRU model compared to the Adam optimization. However, using the default learning rate parameter of 0.001 can be further optimized by tuning the learning rate to achieve the best results for each optimization.

3.3.3. Bi-GRU Model Optimization with Learning Rate

In the third scenario, the model also employs a 20% test size, SMOTE, and optimization. This experiment focuses on evaluating performance improvements when optimization is combined with varying learning rates. The process involves testing learning rate values ranging from 1e-10 to 1e0. Iterations are consistently performed over 50 epochs to gain better insights into the performance of different learning rates. The results of the best learning rate search and the accuracy achieved by each optimizer are shown in Table 9.

Table 9. Performance of Each Optimizer Result with Learning Rate

Optimizer	Performance Metrics				
	Learning Rate	Accuracy	Precision	Recall	F1-Score
Adam	0.018430699693267165	85 (+13.8%)	88 (-1.9%)	85 (+11.8%)	86 (+6.1%)
Nadam	**0.014563484775012459**	**86.8 (+16.2%)**	**86.3 (-3.8%)**	**85.2 (+12.1%)**	**85.5 (+5.5%)**

Table 9 shows the performance results of each optimizer after tuning the learning rate. According to these results, the Adam algorithm with a learning rate of 0.018430699693267165 achieved an accuracy of 85%, representing a 13.8% improvement over the baseline. Similarly, in scenario 2, the Nadam algorithm with a learning rate of 0.014563484775012459 also reached the highest accuracy of 86.8%, marking a 16.2% improvement from scenario 1. Additionally, the Nadam algorithm demonstrated better performance in terms of precision and recall compared to the Adam algorithm. The learning rate, which determines the step size in each iteration of learning, is widely used to enhance testing [51], [52]. In this experiment, the effective use of the learning rate significantly improved the performance of the Bi-GRU model. Using an optimal learning rate is crucial for achieving the best results in model optimization.

3.4. Discussion

This research developed a culinary tourism recommendation system that combines Content-Based Filtering (CBF) with the Bidirectional Gated Recurrent Unit (Bi-GRU) algorithm, achieving a high accuracy increase of 87%. The system uses TF-IDF and Cosine Similarity for feature extraction and item matching, enhanced by Bi-GRU to improve classification accuracy. The CBF results show that this system can provide rating predictions with an MAE of 0.254 and an RMSE of 0.425. Previous research using the same method with a TF-IDF matrix in the context of movie recommendations from Netflix and Disney+ datasets achieved an MAE of 0.28 and an RMSE of 0.67 [13]. Based on a comparison with previous studies, the use of CBF in the context of culinary recommendations demonstrates a significant performance improvement compared to previous research in the context of movie recommendations.

For the classification process with Bi-GRU and SMOTE, the labeling results were converted into binary labels (0 and 1) to simplify the classification process. SMOTE was used to handle imbalanced data after labeling and normalization to (0 and 1), with 7,640 instances of label 0 and 1,360 instances of label 1. In the first scenario, the Bi-GRU model was applied using the SMOTE technique, comparing ratios of 90:10, 80:20, 70:30, and 69:40 with 50 epochs. The results showed that the 80:20 ratio achieved optimal performance, which was then used for subsequent experiments. The model achieved the highest accuracy of 74.7%, with a precision of 89.7%, recall of 76%, and an F1-Score of 81%, as detailed in Table 7. While these results demonstrated good performance, there was room for further improvement with additional optimization. In the second scenario, we optimized the Bi-GRU model combined with SMOTE using the Adam and Nadam algorithms. This resulted in a significant increase in accuracy and other metrics. The Nadam algorithm demonstrated the best performance with an accuracy of 80.2%, a precision of 89%, a recall of 80%, and an F1-Score of 83.4% as detailed in Table 8. This optimization provided a significant improvement in model performance compared to the first scenario, underscoring the importance of optimization in classification.

The third scenario involved tuning the learning rate to further enhance model performance with the baseline, SMOTE, and optimizer to Bi-GRU classification. Each optimizer underwent learning rate tuning over 50 epochs to determine the best learning rate. The results showed that the Nadam algorithm, utilizing a learning rate of 0.014563484775012459, achieved the highest accuracy of 86.8%, marking a 16.2% improvement from the baseline, as seen in Table 9. In addition to improved accuracy, the Nadam algorithm also showed better performance in terms of accuracy performance compared to the Adam algorithm. These results indicate that tuning the learning rate is a crucial step in model optimization to achieve the best performance. Based on the results from the three tested scenarios, the third scenario proved to be the best. This scenario not only achieved the highest accuracy of 86.8% but also demonstrated superior performance in other metrics such as precision and recall, particularly when using the Nadam algorithm with an optimal learning rate. Nadam optimization with learning rate significantly improved performance compared to using default optimization alone, as shown in Fig. 3.

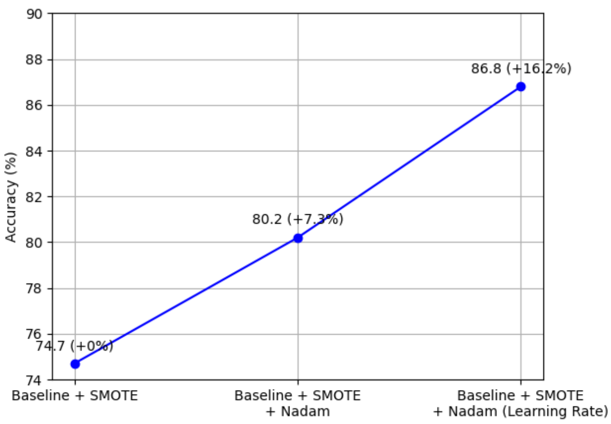

Fig. 3. Accuracy Increase Graph

Previous research has shown that the Bi-GRU model achieves higher accuracy compared to GRU, MLP, LSTM, and CNN models, with an accuracy of 82.04% for sentiment analysis on Twitter [14]. This study highlights the advantages of using the Bi-GRU classification model. It demonstrates the superiority of combining Bi-GRU, SMOTE, and Nadam optimization with an optimal learning rate compared to previous studies. The contribution of this research is highlighting the importance of finding the most optimal learning rate in machine learning models to achieve the best results in the context of culinary recommendations. Additionally, the use of SMOTE was highly effective in addressing class imbalance, a major challenge in this dataset. Overall, the combination of SMOTE, algorithm optimization, and learning rate tuning has proven its ability to enhance the performance of the Bi-GRU model for classification, especially accuracy in the content-based filtering (CBF) culinary recommendation system. The main advantage of this approach is its ability to address challenges in culinary recommendations influenced by diverse user preferences and unstructured data. However, there are limitations, such as dependence on the quality and quantity of available input data and the complexity of optimizing model parameters. Despite TF-IDF having limitations, it is more efficient in analyzing information and yields superior classification performance compared to the use of Word2Vec in previous studies.

4. CONCLUSION

This study demonstrates that the proposed culinary recommendation system based on Content-Based Filtering (CBF) using feature extraction TF-IDF and Bidirectional Gated Recurrent Unit (Bi-GRU) provides accurate and relevant recommendations. To the best of the author's knowledge, the use of feature extraction for word representation and weight assessment approaches to generate numerical results can still be improved. The system uses data from 2,645 Twitter (X) reviews by crawling tweets using keywords of culinary places from PergiKuliner and 44 accounts. Culinary tweet review data is needed to recommend items based on user preferences. Additionally, this study utilizes a second dataset from PergiKuliner, which includes information on descriptions, types, prices, and ratings of 200 culinary places in Bandung. The CBF method with TF-IDF and Cosine Similarity achieved an MAE of 0.254 and an RMSE of 0.425. From the three experiments conducted, with limited data, the Bi-GRU model achieved the best performance after optimization using the Nadam algorithm with an optimal learning rate of 0.014563484775012459. This resulted in a 16.2% accuracy increase from the baseline, with an accuracy of 86.3%, precision of 85.2%, and recall of 85.5%. This study outperforms previous research by adapting the method to recommend culinary options based on relevant user preferences [13], [14]. It provides a theoretical contribution by demonstrating that integrating Bi-GRU into CBF, optimized with a learning rate tuning algorithm, can significantly enhance the performance model, particularly in terms of accuracy and relevance of culinary recommendations. Thus, this recommendation system effectively helps users find culinary spots in Bandung by providing optimal prediction results based on their preferences in real-world applications. This research is expected to be further developed by integrating Collaborative Filtering or Hybrid Filtering methods, conducting experiments with different classifications, adding semantic features, and using larger datasets.

REFERENCES

[1] N. Ben-Lhachemi, E. H. Nfaoui, and J. Boumhidi, “Hashtag Recommender System Based on LSTM Neural Recurrent Network,” *3rd Int. Conf. Intell. Comput. Data Sci. ICDS 2019*, pp. 1–6, 2019, https://dx.doi.org/10.1109/icds47004.2019.8942380.

[2] M. N. Ilyas and E. B. Setiawan, “Weight-Based Hybrid Filtering in a Movie Recommendation System Based on Twitter with LSTM Classification,” *Jurnal Media Informatika Budidarma* vol. 7, pp. 1838–1849, 2023, http://dx.doi.org/10.30865/mib.v7i4.6668.

[3] A. A. Hermawanti, “Influential Elements Identification of Memorable Culinary Tourism in Bandung City,” *Int. J. Recent Trends Bus. Tour.*, vol. 7, no. 2, pp. 13–30, 2023, https://doi.org/10.31674/ijrtbt.2023.v07i02.002.

[4] A. Suciati and I. Budi, “Aspect-based Opinion Mining for Code-Mixed Restaurant Reviews in Indonesia,” *Proc. 2019 Int. Conf. Asian Lang. Process. IALP,* pp. 59–64, 2019, https://doi.org/10.1109/IALP48816.2019.9037689.

[5] H. I. Lee, I. Y. Choi, H. S. Moon, and J. K. Kim, “A multi-period product recommender system in online food market based on recurrent neural networks,” *Sustain.*, vol. 12, no. 3, 2020, https://doi.org/10.3390/su12030969.

[6] D. Roy and M. Dutta, “A systematic review and research perspective on recommender systems,” *J. Big Data*, vol. 9, no. 1, 2022, https://doi.org/10.1186/s40537-022-00592-5.

[7] Y. M. Arif, H. Nurhayati, S. M. S. Nugroho, and M. Hariadi, “Destinations Ratings Based Multi-Criteria Recommender System for Indonesian Halal Tourism Game,” *Int. J. Intell. Eng. Syst.*, vol. 15, no. 1, pp. 282–294, 2022, http://dx.doi.org/10.22266/ijies2022.0228.26.

[8] K. Al Fararni, F. Nafis, B. Aghoutane, A. Yahyaouy, J. Riffi, and A. Sabri, “Hybrid recommender system for tourism based on big data and AI: A conceptual framework,” *Big Data Min. Anal.*, vol. 4, no. 1, pp. 47–55, 2021, https://doi.org/10.26599/BDMA.2020.9020015.

[9] B. Walek and V. Fojtik, "A hybrid recommender system for recommending relevant movies using an expert system," *Expert Syst. Appl.*, vol. 158, p. 113452, 2020, https://doi.org/10.1016/j.eswa.2020.113452.
[10] L. El Youbi El Idrissi, I. Akharraz, and A. Ahaitouf, "Personalized E-Learning Recommender System Based on Autoencoders," *Appl. Syst. Innov.*, vol. 6, no. 6, 2023, https://doi.org/10.3390/asi6060102.
[11] O. Kunde, O. Gaikwad, P. Kelgandre, R. Damodhar, and Prof. Mrs. M. M. Swami, "The Movie Recommendation System using Content Based Filtering with TF-IDF¬¬-Vectorization and Levenshtein Distance," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 2, no. 2, pp. 257–263, 2022, http://dx.doi.org/10.48175/IJARSCT-3648.
[12] A. H. J. Permana and A. T. Wibowo, "Movie Recommendation System Based on Synopsis Using Content-Based Filtering with TF-IDF and Cosine Similarity," *Intl. J. ICT*, vol. 9, no. 2, pp. 1–14, 2023, https://doi.org/10.21108/ijoict.v9i2.747.
[13] A. Nilla and E. B. Setiawan, "Film Recommendation System Using Content-Based Filtering and the Convolutional Neural Network (CNN) Classification Methods," *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 10, no. 1, p. 17, 2024, http://dx.doi.org/10.26555/jiteki.v9i4.28113.
[14] P. Rakshit and A. Sarkar, "A supervised deep learning-based sentiment analysis by the implementation of Word2Vec and GloVe Embedding techniques," *Multimed. Tools Appl.*, pp. 1-34, 2024, http://dx.doi.org/10.1007/s11042-024-19045-7.
[15] H. M. Lynn, S. B. Pan, and P. Kim, "A Deep Bidirectional GRU Network Model for Biometric Electrocardiogram Classification Based on Recurrent Neural Networks," *IEEE Access*, vol. 7, pp. 145395–145405, 2019, https://doi.org/10.1109/ACCESS.2019.2939947.
[16] S. Liu, W. Lin, Y. Wang, D. Z. Yu, Y. Peng, and X. Ma, "Convolutional Neural Network-Based Bidirectional Gated Recurrent Unit–Additive Attention Mechanism Hybrid Deep Neural Networks for Short-Term Traffic Flow Prediction," *Sustain.*, vol. 16, no. 5, 2024, https://doi.org/10.3390/su16051986.
[17] Q. Zhou, C. Zhou, and X. Wang, "Stock prediction based on bidirectional gated recurrent unit with convolutional neural network and feature selection," *PLoS One*, vol. 17, no. 2, pp. 1–20, 2022, https://doi.org/10.1371/journal.pone.0262501.
[18] S. Yang, X. Yu, and Y. Zhou, "LSTM and GRU Neural Network Performance Comparison Study: Taking Yelp Review Dataset as an Example," *Proc. - Int. Work. Electron. Commun. Artif. Intell. IWECAI*, pp. 98–101, 2020, https://doi.org/10.1109/IWECAI50956.2020.00027.
[19] A. Karami, M. Lundy, F. Webb, and Y. K. Dwivedi, "Twitter and Research: A Systematic Literature Review through Text Mining," *IEEE Access*, vol. 8, pp. 67698–67717, 2020, https://doi.org/10.1109/ACCESS.2020.2983656.
[20] Z. Romadhon, E. Sediyono, and C. E. Widodo, "Various Implementation of Collaborative Filtering-Based Approach on Recommendation Systems using Similarity," *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, vol. 4, no. 3, pp. 179–186, 2020, https://doi.org/10.22219/kinetik.v5i3.1062.
[21] V. Çetin and O. Yıldız, "A comprehensive review on data preprocessing techniques in data analysis," *Pamukkale Univ. J. Eng. Sci.*, vol. 28, no. 2, pp. 299–312, 2022, https://dx.doi.org/10.5505/pajes.2021.62687.
[22] I. H. Arsytania, E. B. Setiawan, and I. Kurniawan, "Movie Recommender System with Cascade Hybrid Filtering Using Convolutional Neural Network," *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 9, no. 4, pp. 1262–1274, 2023, http://dx.doi.org/10.26555/jiteki.v9i4.28146
[23] B. M. G. Al Awienoor and E. B. Setiawan, "Movie Recommendation System Based on Tweets Using Switching Hybrid Filtering with Recurrent Neural Network," *Int. J. Intell. Eng. Syst.*, vol. 17, no. 2, pp. 277–293, 2024, https://doi.org/10.22266/ijies2024.0430.24.
[24] G. Ramadhan and E. B. Setiawan, "Collaborative Filtering Recommender System Based on Memory Based in Twitter Using Decision Tree Learning Classification (Case Study: Movie on Netflix)," *ICACNIS Int. Conf. Adv. Creat. Networks Intell. Syst. Blockchain Technol. Intell. Syst. Appl. Hum. Life, Proceeding*, pp. 1-6, 2022, https://doi.org/10.1109/ICACNIS57039.2022.10055248.
[25] X. Bai, M. Wang, I. Lee, Z. Yang, X. Kong, and F. Xia, "Scientific paper recommendation: A survey," *IEEE Access*, vol. 7, pp. 9324–9339, 2019, https://doi.org/10.1109/ACCESS.2018.2890388.
[26] S. Natarajan, S. Vairavasundaram, S. Natarajan, and A. H. Gandomi, "Resolving data sparsity and cold start problem in collaborative filtering recommender system using Linked Open Data," *Expert Syst. Appl.*, vol. 149, 2020, https://doi.org/10.1016/j.eswa.2020.113248.
[27] A. Mee, E. Homapour, F. Chiclana, and O. Engel, "Sentiment analysis using TF–IDF weighting of UK MPs' tweets on Brexit[Formula presented]," *Knowledge-Based Syst.*, vol. 228, p. 107238, 2021, https://doi.org/10.1016/j.knosys.2021.107238.
[28] A. R. Lubis, M. K. M. Nasution, O. S. Sitompul, and E. M. Zamzami, "The effect of the TF-IDF algorithm in times series in forecasting word on social media," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 22, no. 2, pp. 976–984, 2021, http://doi.org/10.11591/ijeecs.v22.i2.pp976-984.
[29] L. Cheng, Y. Yang, K. Zhao, and Z. Gao, *Research and improvement of tf-idf algorithm based on information theory*, vol. 905. Springer International Publishing, 2020, http://dx.doi.org/10.1007/978-3-030-14680-1_67.
[30] N. Valentino and E. B. Setiawan, "Movie Recommender System on Twitter Using Weighted Hybrid Filtering and GRU," *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, vol. 4, no. 2, pp. 159–172, 2024, https://doi.org/10.22219/kinetik.v9i2.1941.
[31] M. Li, "Recommendation System Building based on CNN and TF-IDF Approaches," *Highlights in Science, Engineering and Technology*, vol. 92, pp. 178-187, 2024, https://doi.org/10.54097/633gjj39.

[32] K. Park, J. S. Hong, and W. Kim, "A Methodology Combining Cosine Similarity with Classifier for Text Classification," *Appl. Artif. Intell.*, vol. 34, no. 5, pp. 396–411, 2020, https://doi.org/10.1080/08839514.2020.1723868.

[33] Y. Sun, J. Zhang, Z. Yu, Y. Zhang, and Z. Liu, "The Bidirectional Gated Recurrent Unit Network Based on the Inception Module (Inception-BiGRU) Predicts the Missing Data by Well Logging Data," *ACS Omega*, vol. 8, no. 30, pp. 27710–27724, 2023, https://doi.org/10.1021/acsomega.3c03677.

[34] M. S. Başarslan and F. Kayaalp, "MBi-GRUMCONV: A novel Multi Bi-GRU and Multi CNN-Based deep learning model for social media sentiment analysis," *J. Cloud Comput.*, vol. 12, no. 1, 2023, https://doi.org/10.1186/s13677-022-00386-3.

[35] S. Meng, X. Q. Jiang, Y. Gao, H. Hai, and J. Hou, "Performance Evaluation of Channel Decoder based on Recurrent Neural Network," *J. Phys. Conf. Ser.*, vol. 1438, no. 1, 2020, http://dx.doi.org/10.1088/1742-6596/1438/1/012001.

[36] O. A. Montesinos López, A. Montesinos López, and J. Crossa, *Multivariate Statistical Machine Learning Methods for Genomic Prediction*. 2022, http://dx.doi.org/10.1007/978-3-030-89010-0.

[37] H. Yadav, P. Shah, N. Gandhi, T. Vyas, A. Nair, S. Desai, L. Gohil, S. Tanwar, R. Sharma, V. Marina, and M. S. Raboaca, "CNN and Bidirectional GRU-Based Heartbeat Sound Classification Architecture for Elderly People," *Mathematics*, vol. 11, no. 6, pp. 1–25, 2023, https://doi.org/10.3390/math11061365.

[38] C. V. M. Krishna, G. A. Rao, and S. Anuradha, "Analysing the impact of contextual segments on the overall rating in multi-criteria recommender systems," *J. Big Data*, vol. 10, no. 1, 2023, http://dx.doi.org/10.1186/s40537-023-00690-y.

[39] X. Xu, "Revolutionizing Education: Advanced Machine Learning Techniques for Precision Recommendation of Top-Quality Instructional Materials," *Int. J. Comput. Intell. Syst.*, vol. 16, no. 1, pp. 1–14, 2023, http://dx.doi.org/10.1007/s44196-023-00361-z.

[40] Z. Fayyaz, M. Ebrahimian, D. Nawara, A. Ibrahim, and R. Kashef, "Recommendation systems: Algorithms, challenges, metrics, and business opportunities," *Appl. Sci.*, vol. 10, no. 21, pp. 1–20, 2020, https://doi.org/10.3390/app10217748.

[41] T. R. Mahesh, V. Vinoth Kumar, and S. J. Lim, "UsCoTc: Improved Collaborative Filtering (CFL) recommendation methodology using user confidence, time context with impact factors for performance enhancement," *PLoS One*, vol. 18, no. 3, pp. 1–22, 2023, https://doi.org/10.1371/journal.pone.0282904.

[42] T. Schick and H. Schütze, "BERTRAM: Improved word embeddings have big impact on contextualized model performance," *Proc. Annu. Meet. Assoc. Comput. Linguist.*, pp. 3996–4007, 2020, https://doi.org/10.18653/v1/2020.acl-main.368.

[43] D. H. Riaji and E. B. Setiawan, "Film Recommendation Using Content-Based Using Artificial Neural Network Method And Adam Optimization," *Jurnal Teknik Informatika*, vol. 5, no. 1, pp. 291–300, 2024, https://doi.org/10.52436/1.jutif.2024.5.1.1729.

[44] H. S. Ramadhan and E. B. Setiawan, "Social Media Based Film Recommender System (Twitter) on Disney+ with Hybrid Filtering Using Support Vector Machine," *Sinkron*, vol. 8, no. 4, pp. 2215–2225, 2023, https://doi.org/10.33395/sinkron.v8i4.12876.

[45] A. Jalilifard, V. F. Caridá, A. F. Mansano, R. S. Cristo, and F. P. C. da Fonseca, "Semantic Sensitive TF-IDF to Determine Word Relevance in Documents," *Lect. Notes Electr. Eng.*, vol. 736, pp. 327–337, 2021, http://dx.doi.org/10.1007/978-981-33-6987-0_27.

[46] C. Kaope and Y. Pristyanto, "The Effect of Class Imbalance Handling on Datasets Toward Classification Algorithm Performance," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 22, no. 2, pp. 227–238, 2023, https://doi.org/10.30812/matrik.v22i2.2515.

[47] A. Z. R. Adam and E. B. Setiawan, "Social Media Sentiment Analysis using Convolutional Neural Network (CNN) dan Gated Recurrent Unit (GRU)," *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 9, no. 1, pp. 119–131, 2023, http://dx.doi.org/10.26555/jiteki.v9i1.25813.

[48] J. Qin, Z. Liu, L. Ran, and R. Xie, "Synthetic aperture radar automatic target recognition based on cost-sensitive awareness generative adversarial network for imbalanced data," *IET Radar, Sonar Navig.*, no. April, 2024, http://dx.doi.org/10.1049/rsn2.12583.

[49] M. G. M. Abdolrasol, S. M. S. Hussain, T. S. Ustun, M. R. Sarker, M. A. Hannan, R. Mohamed, J. A. Ali, S. Mekhilef, and A. Milad, "Artificial neural networks based optimization techniques: A review," *Electron.*, vol. 10, no. 21, 2021, http://dx.doi.org/10.3390/electronics10212689.

[50] T. Hillel, M. Bierlaire, M. Z. E. B. Elshafie, and Y. Jin, "A systematic review of machine learning classification methodologies for modeling passenger mode choice," *J. Choice Model.*, vol. 38, p. 100221, 2021, https://doi.org/10.1016/j.jocm.2020.100221.

[51] N. Fatima, "Enhancing Performance of a Deep Neural Network: A Comparative Analysis of Optimization Algorithms," *ADCAIJ Adv. Distrib. Comput. Artif. Intell. J.*, vol. 9, no. 2, pp. 79–90, 2020, http://dx.doi.org/10.14201/ADCAIJ2020927990.

[52] Y. Wu, L. Liu, J. Bae, K. Chow, A. Iyengar, C. Pu, W. Wei, L. Yu, and Q. Zhang, "Demystifying Learning Rate Policies for High Accuracy Training of Deep Neural Networks," *Proc. IEEE Int. Conf. Big Data, Big Data,* pp. 1971–1980, 2019, https://doi.org/10.1109/BigData47090.2019.9006104.

BIOGRAPHY OF AUTHORS

Adhyasta Naufal Faadhilah is an undergraduate student of the Faculty of Informatics at Telkom University, Bandung, Indonesia. Email: adhyastanf@student.telkomuniversity.ac.id.

Erwin Budi Setiawan is a senior lecturer at the Faculty of Informatics at Telkom University, Bandung. With over 10 years of experience in research and teaching in the field of Informatics, he currently holds the position of Associate Professor. His research interests include machine learning, people analytics, and social media analytics.
Email: erwinbudisetiawan@telkomuniversity.ac.id.