Matrik: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer
Vol. 23, No. 3, July 2024, pp. 663∼676
ISSN: 2476-9843, accredited by Kemenristekdikti, Decree No: 200/M/KPT/2020
DOI: 10.30812/matrik.v23i3.3835

Multiclass Text Classification of Indonesian Short Message Service (SMS) Spam using Deep Learning Method and Easy Data Augmentation

Nurun Latifah , Ramaditia Dwiyansaputra , Gibran Satya Nugraha
University of Mataram, Mataram, Indonesia

Article Info

Article history:

Received January 29, 2024
Revised February 29, 2024
Accepted June 12, 2024

Keywords:

Easy Data Augmentation
Multiclass Classification
Short Message Service Spam
Text Classification

ABSTRACT

The ease of using Short Message Service (SMS) has brought the issue of SMS spam, characterized by unsolicited and unwanted. Many studies have been conducted utilizing machine learning methods to build models capable of classifying SMS Spam to overcome this problem. However, most of these studies still rely on traditional methods, with limited exploration of deep learning-based approaches. Whereas traditional methods have a limitation compared to deep learning, which performs manual feature extraction. Moreover, many of these studies only focus on binary classification rather than multiclass SMS classification, which can provide more detailed classification results. **The aim of this research** was to analyze the deep learning model for multiclass Indonesian SMS spam classification with six categories and to assess the effectiveness of the text augmentation method in addressing data imbalance issues arising from the increased number of SMS categories. **The research method** used was the Indonesian version of the Bidirectional Encoder Representations from Transformers (IndoBERT) model and the Easy Data Augmentation (EDA) technique to address the imbalance dataset issue. The evaluation is conducted by comparing the performance of the IndoBERT model on the dataset and applying EDA techniques to enhance the representation of minority classes. **The result of this research** showed that IndoBERT achieved a 91% accuracy rate in classifying SMS spam. Furthermore, using the EDA technique significantly improved the f1-score, with an average 12% increase in minority classes. Overall model accuracy also improved to 93% after EDA implementation. **This research concluded** that IndoBERT is effective for multiclass SMS spam classification, and the EDA is beneficial in handling imbalanced data, contributing to enhancing model performances.

Corresponding Author:

Ramaditia Dwiyansaputra, +6285737087075
Faculty of Engineering, Departement of Informatics Engineering,
University of Mataram, Mataram, Indonesia,
Email: rama@unram.ac.id

How to Cite:
N. Latifah, R. Dwiyansaputra, and G. S. Nugraha, "Multiclass Text Classification of Indonesian Short Message Service (SMS) Spam using Deep Learning Method and Easy Data Augmentation", *MATRIK: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer*, Vol. 23, No. 3, pp. 663-676, July, 2024.

1. INTRODUCTION

Short Message Service (SMS) spam is an unsolicited and unwanted text message sent to a user's cellular phone and is generally intended for commercial purposes [1]. Based on the observation by the Indonesian Consumer Protection Agency (BPKN), the results show that the spread of prolonged short messages (SMS spam) is growing and rampant, even causing unrest in the community [2]. The data obtained is that there were 890 reports from January to August 2020 [2]. In addition, the 2019 Truecaller Insight Report provides findings that Indonesia is the country ranked 10th largest recipient of SMS spam in the world [2]. SMS users in Indonesia are still high, with 37 million registered customers and 25-28 million active users [3]. In recent years, it has been recorded that the total number of spam messages exceeds the number of spam emails [4]. This can be a factor of how massively SMS spam is spread by irresponsible parties. With the widespread use of SMS as the most important delivery platform, many use it for marketing purposes similar to advertising media, and some even use it for fraud purposes [5]. In addition, due to the affordability and instant nature of SMS for sending text messages between users, fraudsters exploit it to their advantage [4]. They utilize the trust that users place in SMS as a medium to receive important notifications, such as financial alerts, job application information, and the latest offers from network service providers [4]. Consequently, mobile phone users may inadvertently provide their personal information, believing they are responding to a legitimate bank or service provider request. The consequences of such actions can be catastrophic for the deceived user. Based on research [6], it is explained that many victims have received fraud in the Covid-19 pandemic era in the form of SMS spam and calls from people using various methods, typically using the family of the victim as a tactic and unpaid online loan payments as a trap, which has an impact on the number of victims that continues to grow. In addition, in the research [6], there was a respondent who received spam SMS after taking a loan from an online lending site with a message in the form of getting a prize from several big banks, which disturbed the convenience and privacy of SMS users.

There is an initiative taken by the government to accommodate the problem of SMS spam by making regulations through "Article 17 of the Regulation of Minister of Communications and Informatics on the Implementation of Premium Messaging Services and the Delivery of Short Message Services" which explains that senders of short message services are prohibited from sending messages that are contrary to the public interest and violate norms [2]. Referring to this law, it can be interpreted that SMS spam is an activity that violates the law because it is contrary to the public interest and disrupts public security and order [2]. Although a law already regulates it, the problem of SMS spam persists, so additional solutions are needed to overcome it. The solution needed is to implement SMS spam detection to be more effective by classifying messages that belong to the spam or non-spam category. Classification is a stage to find a model by categorizing so that unlabeled classes can be predicted [3]. There is a lot of research related to the classification of SMS spam. The most commonly used methods for SMS spam classification are Support Vector Machine, Naïve Bayes, k-NN, Decision Tree, and Logistic Regression [7–12]. Some research related to Indonesian SMS spam has been done before, such as in [7] and [8] with two categories, namely spam and non-spam [7] and three categories, namely ham, promotional messages, and spam [8]. The amount of data in [7] is 500 data while in [8] is 4125 train data with. Both studies performed classification by applying several machine learning techniques. Research [7] uses algorithms namely Sigmoid Kernel SVM, Linear Kernel SVM and RBF, KNN, and Multinomial Naive Bayes with the highest accuracy achieved by the Sigmoid Kernel SVM with a difference in accuracy with the RBF Kernel of 2.26%, Linear Kernel of 0.09%, KNN of 27.56%, and Multinomial Naïve Bayes of 4.37%. While the algorithms used in [8] are Multinomial Naive Bayes (MNB), Multinomial Logistic Regression (MLR), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Decision Tree (DT), Stochastic Gradient Descent (SGD), XGBoost (XGB), and Random Forest (RF) with 10-fold cross validation obtained the four best models based on accuracy, namely Random Forest (94.62%), Multinomial Logistic Regression (94.57%), Support Vector Machine (94.38%), and XGBoost (94.52%). However, traditional machine learning methods have limitations, especially those involving feature extraction such as Bag-of-Word (BOW) and Term Frequency-Inverse Document Frequency (TF-IDF). These approaches tend to treat words separately without considering their position in the text, the contextual relationship between words in a sentence, and the meaning of the word as a whole. Therefore, semantic and contextual information is often limited [13]. In addition, although traditional machine learning methods provide good performance, they require manual feature extraction [14]. Sometimes, more complex features are required to perform a more in-depth analysis. To overcome this limitation, deep learning methods can be used. Deep learning extracts features directly from the data during the training process, eliminating the need for manual extraction [14].

One of the deep learning methods that can be used in classifying SMS spam is the IndoBERT method. IndoBERT is a Transformer-based model similar to BERT [15] but trained as a masked language model using the Huggingface framework. IndoBERT is one of the state-of-the-art models for performing text analysis using the Indonesian language. The architecture built using the Transformer model in BERT generally uses English. Several studies have used IndoBERT in the context of Natural Language Processing (NLP), such as sentiment analysis and fake news detection, and the results show that IndoBERT obtains a fairly good level of performance when compared to traditional machine learning methods [16, 17]. Another research that has applied IndoBERT to NLP cases is the Analysis of COVID-19 Radiology Reports, which also provides accurate results [18]. Therefore, the IndoBERT method performs well for several NLP tasks [19]. Meanwhile, research related to IndoBERT has not been found in the use of SMS spam classification. Examples of the use of IndoBERT include research conducted in Indonesia to detect fake news [17]. The study obtained data from Twitter with a total of 3465 fake news and 766 real news. The research [17] who compared the performance of the IndoBERT model with several machine learning models, such as TF-IDF + SVM model and TF-IDF + Naïve Bayes, obtained the results that the accuracy of the TF-IDF + SVM model reaches 90%, the accuracy of the TF-IDF + Naïve Bayes model reaches 83% and the highest accuracy is obtained using the IndoBERT method, which is 94.66%. Other research that has used IndoBERT is to conduct sentiment analysis, namely sentiment analysis of Indonesian public opinion about the 2024 general election [20] and sentiment classification on the issue of leakage of cellphone identity card data [21]. Both studies obtained data from Twitter, where research [21] used 957 tweets. The results of [20] showed that the majority of Indonesian people gave a neutral response to the election, and the IndoBERT large-p1 model achieved 83.5% accuracy, outperforming the machine learning model by 14.49%, and the lexicon-based model by 46.49%, in terms of f1-score. Then, in [22], which uses several methods such as Random Forest, Logistic Regression, Support-Vector Machine, and IndoBERT obtained the results that the Support Vector Machine algorithm has the best performance with an f1-score value of 0.81, followed by the Random Forest of 0.78, IndoBERT of 0.76, and the Logistic Regression of 0.74. IndoBERT, as one of the state-of-the-art in NLP, has low performance compared to other algorithms. This is due to the imbalanced class and lack of training data.

One of the problems faced by the IndoBERT model is the imbalance of data [22] and can cause a decrease in the performance of IndoBERT compared to other algorithms due to the imbalance of data [21]. Imbalanced data is a situation where the models cannot rationally learn the features of the minority class because the majority class samples are more than the minority class samples [23]. The problem of imbalanced data also occurs in our research, where the number of minority class samples is less than that of the majority class. One approach that can be attempted to address this issue is the implementation of Easy Data Augmentation (EDA) techniques. EDA is a data augmentation technique to improve performance in text classification [24]. EDA has four simple operations that are very influential, namely synonym replacement, random insertion, random swap, and random deletion [24]. This technique is expected to provide a solution by generating new data variations on the minority class to improve the model's ability to recognize and understand the class context.

Based on the literature reviews described above, some gaps have not been resolved by previous research, namely that many studies have focused on traditional machine learning techniques for SMS spam classification, and there is limited exploration of deep learning-based approaches like IndoBERT. Traditional methods often require manual feature extraction, which can be a limitation in capturing the semantic and contextual nuances of SMS messages. Furthermore, most prior research has categorized SMS messages into only a few broad categories, typically two (spam and non-spam) or three (ham, promotional messages, and spam). This does not capture the diversity and complexity of SMS spam content into broader categories as done by studies [9, 10] and [25]. Our study proposes a more detailed classification into six categories, including ”online loan SMS,” ”fraud SMS,” ”gambling SMS,” ”personal SMS,” ”offer SMS,” and ”operator SMS.” This finer granularity aims to provide a more comprehensive understanding and better handling of various types of SMS content. However, one problem arising from the increased number of SMS categories is the imbalanced dataset. This imbalance can lead to poor performance in recognizing minority classes, especially for the IndoBERT model [22]. Although some research acknowledges this issue, effective solutions like using EDA to balance the dataset have not been fully explored. The potential of combining EDA with IndoBERT to address the data imbalance problem has not been fully examined. While EDA can generate new data variations to improve the representation of minority classes, its impact on the performance of IndoBERT in the context of SMS spam classification has not been empirically validated in prior research.

This research aims to analyze a deep learning model for multiclass Indonesian SMS spam classification with six categories and to assess the effectiveness of the text augmentation method in addressing data imbalance issues arising from the increased number of SMS categories. **The difference between this research** and the previous studies is that it proposes a pre-trained IndoBERT model for multiclass classification. Additionally, this research is to evaluate how effective the EDA augmentation technique is in addressing data imbalance issues. This research makes significant **contributions** in introducing a model for multiclass SMS classification, presenting an innovative approach compared to previous studies. This study advances the comprehension of deep learning-based strategies in SMS spam classification. It provides valuable insights into tackling data imbalance challenges, thereby contributing to the progression of knowledge in natural language processing and cybersecurity.

2. RESEARCH METHOD

This research is conducted in several stages, described as a research flow diagram, starting from the data collection stage to evaluation. The research flow diagram is shown in Figure 1. Based on Figure 1, the research flow consists of data preparation, preprocessing, modeling, Data Augmentation, and Evaluation.

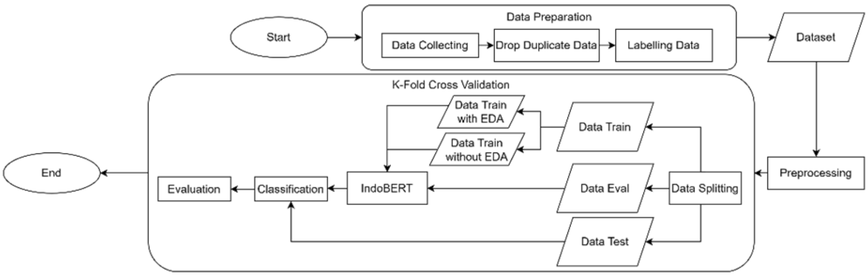

Figure 1. Research Methodology

2.1. Data Preparation

Data preparation consists of three stages: collection, drop of duplicate data, and labeling. The dataset used in this research is Indonesian SMS data collected from several cell phones that have been approved by their owners and datasets available from previous research [7] and [26]. SMS data taken from cell phones is obtained from backup results using the "SMS Backup and Restore" application downloaded from the Google PlayStore. The dataset that has been backed up has a .xml extension; then the entire dataset is converted into .csv format. Next, the process of deleting duplicate data was carried out, and then the data labeling process. Data is categorized into six classes, namely "personal SMS," "online loan SMS," "fraud SMS," "offer SMS," "operator SMS," and "gambling SMS." Two research colleagues performed annotations of the collected SMS data to avoid bias. Additionally, a guideline is created by considering the characteristics of each category to create a uniform annotation standard. Then, annotators made annotations of SMS messages until all data was successfully annotated for use in the next stage. The description and distribution of the amount of data from the dataset are presented in Table 1 and Figure 2.

Table 1. Pembagian data untuk Training dan Testing

Class	Description	Amount
personal SMS	Refers to private messages that are usually sent between individuals for non-commercial purposes.	829
online loan SMS	Represent messages related to online lending services. These messages often contain offers of quick loans.	78
fraud SMS	Denotes messages that are intended to deceive or trick the recipient of the message	138
offer SMS	Includes promotional messages from businesses or services offering different products, discounts, or special offers	823
operator SMS	Involves messages from mobile network operators, banks, and official agencies.	840
gambling SMS	Includes messages related to gambling or betting activities	42
Totals		2750

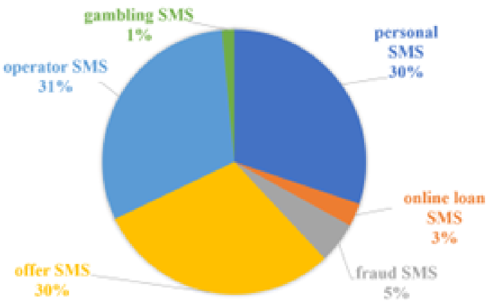

Figure 2. Data distribution

The data will be divided into train, validation, and test data, with a ratio of division for test data of 20% and 80% for training data. Then, 80% of the training data will be divided into validation data (50%) and training data (50%). The dataset still experiences an imbalance in each class's data. This can be seen in the unbalanced distribution of the amount of data for each class, where of the six classes, there are three classes that are the majority class or with more data, namely "personal SMS," "offer SMS," and "operator SMS" classes. While the other three classes are minority classes because they have a small amount of data, namely "online loan SMS," "fraud SMS," and "gambling SMS" classes (See in Table 2).

Table 2. An example of a dataset

SMS Text	Translated Messages	Class
Bangunan nya ini milik orang tua pak, tapi utk kepentingan bersama,toko nya di sewakan ,dan sewanya dibagirata kami 6 bersaudara	The building is owned by my parents, but for mutual benefit, the shop is rented out, and the rent is divided equally among us 6 siblings,	personal SMS
Kami menawarkan pinjaman rupiah online tanpa jaminan/agunan minim 5jt-500jt bunga 2% Info lkp chat wa: 0887435830844	We offer online rupiah loans without collateral / collateral minimum 5 million-500 million interest 2% More info chat wa: 0887435830844	online loan SMS
selamat . . . anda mendapatkan hadiah Rp.175jt dari program SHOPEE 0FFICIAL DAY 2021 dengan kode pin pemenang:25f4777 untuk info:klik tinyuri.com/Puncak-rejeki2021	congratulations . . . you get a prize of Rp.175 million from the SHOPEE 0FFICIAL DAY 2021 program with the winning pin code: 25f4777 for info: click tinyuri.com/Puncak-rejeki2021	fraud SMS
Dapatkan pulsa XL s/d 500Rb dan Bonus Kartu Perdana XL setiap buka rekening NYALA via ONe Mobile. Klik bit.ly/PromoXLOCBC untuk info lebih lanjut MCH87A	Get XL credit up to 500,000 and XL Starter Pack Bonus every time open NYALA account via ONe Mobile. Click bit.ly/PromoXLOCBC for more info MCH87A	offer SMS
(XL) Sisa kuota Internet GRATIS Anda sdh habis. Kelebihan kuota akan memotong kuota utama atau tarif dasar. Cek kuota di ∗123♯ atau my.xl.co.id. Info 818	(XL) Your remaining FREE Internet quota has been exhausted. Excess quota will be deducted from the main quota or basic tariff. Check your quota at ∗123♯ or my.xl.co.id. Info 818	operator SMS
Sdh bnyk yg mngg b0sq lsng kt byr brppun dft skr dis1tu5 jd onli-ne terbaik b0-la,5l0-t,casi-n0,p0-ker,t0-gel 0nli-ne dft skrng dilin-k: bit.ly/Hokibet5	There are many who are waiting for b0sq directly we pay whatever reg-ister now at the best online gambling site, ball, slot, casino, poker, online togel register now at: bit.ly/Hokibet5	gam-bling SMS

2.2. Preprocessing

Data Preprocessing is the initial stage of preparing the dataset to be used properly for further processing. Preprocessing is performed to remove unnecessary parts or text to obtain quality data to be processed. This stage aims to clean data that still has lots of noise. Generally, the processed data is data that is still unstructured and has many repetitive words or characters that are not needed in the classification process [27]. The stages that are passed in the preprocessing include Cleaning Teks and Case Folding Teks. The text cleaning process is a process of cleaning the text through several steps, which include the removal of mentions or words that begin with the character ′@,′ hashtags, or words that begin with the character ′@,′ URLs, numbers, and special characters. The next step is to remove emojis from SMS text by converting them into ASCII characters. At this stage, the entire text is converted into lowercase letters. This aims to speed up the comparison process when text processing is applied. The case folding stage is used to normalize text so that differences that occur in the writing of letters do not affect text analysis, especially when comparing words. In addition, doing case folding will make the comparison process between words more efficient because there is no impact from differences in lowercase and uppercase letters. An example of the preprocessing stage is shown in Table 3.

Table 3. Preprocessing Flow

Process	Sentence 1
Initial	Alhamdulillah, makasih yaaa udah meluangkan waktu 🥰
Cleaning	Alhamdulillah makasih yaaa udah meluangkan waktu 🥰
Removing emoji	Alhamdulillah makasih yaaa udah meluangkan waktu
Replace Three or more character	Alhamdulillah makasih ya udah meluangkan waktu
Case folding	alhamdulillah makasih ya udah meluangkan waktu

2.3. Modelling

The next stage is to build a model that divides the data into train, validation, and test data. Two schemes will be carried out on the train data, namely, train data without the EDA process and train data using the EDA process to overcome the data imbalance problem and determine its effect. The model used for this research is IndoBERT. IndoBERT method is chosen because it refers to research [28] that IndoBERT is one of the good methods because it has been trained with a large dataset. IndoBERT is a BERT model that was trained using Masked Language Modelling on Indonesian datasets [29]. The tokenization process is first carried out before a model is built using the IndoBERT pre-trained model. The tokenization process is carried out by applying a Transformer tokenizer to convert text into tokens that provide a representation of words or sub-words so that it is easily understood by the model. The results of tokenization are then ready to be implemented in fine-tuning process using the IndoBERT model. There are several architectural variants of IndoBERT, such as IndoBERT-base and IndoBERT-large. These two architectures have variations in the number of layers of transformers, attention heads, and parameters [20]. The IndoBERT base contains 12 layers of transformers, an attention head size 12, and 110 million parameters. Meanwhile, IndoBERT-large contains 24 layers of transformers, 16 attention heads, and a parameter count of 340 million. The basic architecture of these types of IndoBERT is shown in Figure 3.

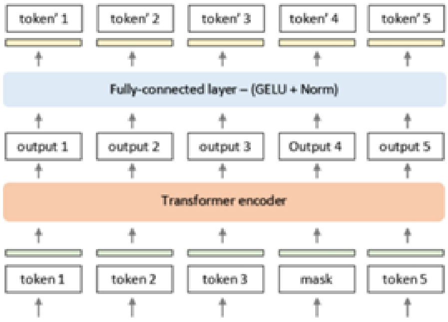

Figure 3. IndoBERT Architecture [20]

The previously trained IndoBERT model will be used for fine-tuning (taking a model previously trained on large data to be adapted to smaller data, in this case, SMS spam text classification). IndoBERT model versions are IndoBERT-base-p1 and base-p2 versions of Hugging Face. The model will be adapted to SMS text classification with six SMS categories to be classified. Training is conducted using the 'Trainer' object from Transformers library. To configure the training, the' TrainingArguments' object manages the parameters. Classification is performed using several parameters shown in Table 4. Once the training and classification are completed, the next step involves evaluating the model to assess its final performance. The outcomes of the training and validation of the model can be utilized to analyze the performance of the model and understand how well it can recognize previously unseen data.

Table 4. IndoBERT Parameter

Parameter	Value
epoch	3
batch_size	8
warmup_steps	500
weight_decay	0.01

2.4. Easy Data Augmentation (EDA)

This study has a problem of imbalanced class according to the data distribution in Figure 2. The data containing class imbalance may affect the performance of the classification method. Therefore, unbalanced data should be handled to improve the performance of the classification method [30]. To overcome the problem, this research tries to apply the EDA technique to balance the data by performing data augmentation on the minority class. EDA is a data augmentation technique to improve performance in text classification [24]. EDA is used in this research because referring to [24]. EDA can create synthesized data that is used to train the model effectively, increasing the classification accuracy. EDA has four very influential simple operations: synonym replacement, random insertion, random swap, and random deletion. Based on research [24], here are the stages that can be passed with EDA.

For the sentences obtained in the training set, we randomly select and execute one of the operations: 1. Synonym Replacement (SR): n words from the sentence that does not stop words are randomly selected. Then, one of its synonyms will be randomly selected to replace each word; 2. Random Insertion (RI): Search for a random synonym for the random word in the sentence that is not a stop word. Then, the synonym will be inserted into a random location in the sentence, which is done n times; 3. Random Swap (RS): two words in the sentence are randomly selected, and their positions are swapped by doing it n times; 4. Random Deletion (RD): Each word in the sentence is randomly deleted with probability p. Long sentences have a larger number of words when compared to short sentences; hence, they can absorb more noise while preserving original class labels. To balance this, the way to do it is to vary the number of transformed words, n, for SR, RI, and RS based on the length of the sentence l with the formula $n = \alpha$, where α is a parameter indicating the percentage of words in the sentence that are transformed (for RD we use: $p = \alpha$). Next, for every original sentence, we produced n_{aug} augmented sentence.

2.5. Evaluation

The model development is carried out using k-fold cross-validation. The k-fold cross-validation process is an evaluation method that divides the dataset into k subsets or folds of equal size. The model is trained at each iteration using k-1 folds as training and validation data. This means that from the k-1 folds that were originally training data, we divide them into two parts, becoming training data and validation data. While the remaining folds are used as test data. We use 5 folds, so this process is repeated five times, where each fold is used once as test data. The test results are then averaged to produce a more stable evaluation metric. After the model building, training, and classification processes are done, the evaluation process is performed. Measuring the classification performance of the developed model can be done using one of the mechanisms of a confusion matrix. The confusion matrix will show how far the designed model can predict the existing class correctly. In measuring performance with a confusion matrix, four terms can be used to express the results of the classification process. They are True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). TP is the positive data that is correctly detected. False FN is the reverse of True Positive, which is positive data but is detected unfavorably. The TN value is the amount of malicious data detected as true, while FP is malicious data detected as positive data [7].

Measuring the performance of the model that has been created can be done using a confusion matrix to calculate various performance metrics. Some performance metrics that are commonly used are Accuracy, Precision, Recall, and F1-Score. Accuracy shows how accurately the model can classify correctly, which means the closeness of the predicted value to the true value. Accuracy can be used to measure performance in multiclass classification. Precision shows the accuracy of the requested data with the model's predicted results, which means the ratio of correct positive predictions compared to all positive predicted results. In other words, from all of the positive classes that have been predicted correctly, how much of the data is actually positive. Recall shows the model's success in retrieving information, which is the ratio of true positive predictions compared to all true positive data. F1-Score shows the comparison of the weighted average precision and recall values.

3. RESULT AND ANALYSIS

The following are the results of what was done on the SMS spam dataset to classify SMS spam in multiclass. The dataset is preprocessed and continued with classification using the IndoBERT, which is divided into three data sets: train data, validation data, and test data. Classification is done using two IndoBERT models, using IndoBERT-base-p1 and IndoBERT-base-p2. Because there is an imbalance of data in the class, the classification process uses two types of train data, which use train data without the EDA process and train data with the EDA process so that the impact of classification results on train data with or without the EDA process can be assessed. Testing is done using 5-fold cross-validation.

3.1. The Results of Preprocessing Data

Data preprocessing is carried out with the stages of text cleaning and case folding text so that the preprocessing results are shown in Table 5 for data before preprocessing and Table 6 for data preprocessing results. The data is then preprocessed with several stages, such as cleaning the text, including removal of mentions or words starting with the character ′@,′ hashtags or words starting with the character ′♯′, URLs, numbers, and special characters. Then, text will be converted into lowercase letters at the case folding stage. The process is carried out so that the next stage can be processed properly.

Table 5. Example Data Before Preprocessing

Text	Class
Pak mau tanya kalau nilai remed lebih kecil itu diambil nilai nya yang remed atau tetap yang sebelum nya pak?	personal SMS
Yth Nasabah Bank Syariah Indonesia,segera aktivasi BSI Mobile anda.Kode aktivasi anda 78583120 berlaku hingga 18 Feb 2021 08:14:44 WIB. Mohon disimpan	operator SMS

Table 6. Pembagian data untuk Training dan Testing

Text	Class
pak mau tanya kalau nilai remed lebih kecil itu diambil nilai nya yang remed atau tetap yang sebelum nya pak	personal SMS
yth nasabah bank syariah indonesia segera aktivasi bsi mobile anda kode aktivasi anda berlaku hingga feb wib mohon disimpan	operator SMS

3.2. The Results of the Evaluation Model without EDA

The SMS spam classification process is carried out on train, validation, and testing data using the IndoBERT model, starting without using the EDA process. The process was run through a 5-fold cross-validation scheme. Model performance data at each fold is shown in Table 7. Based on the experiments conducted in Table 77, training data using IndoBERT base-p1 gives the highest accuracy of 93% on fold 5 and 92% when using IndoBERT base-p2 on folds 2, 3, and 5. After conducting a 5-fold cross-validation on the classification model, an average calculation is performed to produce a classification report reflecting overall model performance. This average is done to get a more stable and consistent evaluation in measuring model performance against different datasets at each iteration. The average model performance results without using the EDA technique are shown in Table 8.

Table 7. Classification Results without EDA Technique

Fold	Class	IndoBERT base-p1				IndoBERT base-p2			
		Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy
Fold 1	personal SMS	0.99	0.95	0.97	0.91	0.99	0.95	0.97	0.88
	online loan SMS	0.94	1.00	0.97		0.88	0.94	0.91	
	fraud SMS	0.64	0.93	0.76		0.95	0.78	0.86	
	offer SMS	0.97	0.87	0.91		0.95	0.75	0.83	
	operator SMS	0.86	0.89	0.87		0.76	0.96	0.85	
	gambling SMS	0.57	1.00	0.73		0.78	0.88	0.82	
Fold 2	personal SMS	0.95	0.97	0.96	0.91	0.96	0.95	0.96	0.92
	online loan SMS	0.93	0.88	0.90		1.00	0.94	0.97	
	fraud SMS	0.83	0.89	0.86		0.92	0.85	0.88	
	offer SMS	0.97	0.83	0.90		0.95	0.88	0.91	
	operator SMS	0.85	0.95	0.90		0.85	0.96	0.90	
	gambling SMS	0.83	0.62	0.71		1.00	0.62	0.77	
Fold 3	personal SMS	0.84	0.98	0.90	0.89	0.95	0.96	0.95	0.92
	online loan SMS	0.72	0.87	0.79		0.67	0.93	0.78	
	fraud SMS	0.87	0.71	0.78		0.72	0.75	0.74	
	offer SMS	0.94	0.96	0.95		0.99	0.88	0.93	
	operator SMS	0.93	0.81	0.87		0.88	0.93	0.91	
	gambling SMS	1.00	0.38	0.55		1.00	0.88	0.93	
Fold 4	personal SMS	0.98	0.95	0.97	0.90	0.99	0.92	0.95	0.90
	online loan SMS	1.00	0.73	0.85		1.00	0.87	0.93	
	fraud SMS	0.62	0.93	0.74		0.93	0.96	0.95	
	offer SMS	0.94	0.90	0.92		0.92	0.87	0.90	
	operator SMS	0.87	0.89	0.88		0.81	0.94	0.87	
	gambling SMS	0.67	0.44	0.53		1.00	0.33	0.50	
Fold 5	personal SMS	0.93	0.99	0.96	0.93	0.97	0.97	0.97	0.92
	online loan SMS	0.89	1.00	0.94		0.89	1.00	0.94	
	fraud SMS	0.92	0.82	0.87		0.96	0.79	0.86	
	offer SMS	0.91	0.98	0.94		0.90	0.96	0.93	
	operator SMS	0.97	0.86	0.91		0.92	0.88	0.90	
	gambling SMS	0.88	0.78	0.82		0.67	0.67	0.67	

Table 8. Average Classification Result without EDA Technique

Class	IndoBERT base-p1				IndoBERT base-p2			
	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy
personal SMS	0.94	0.97	0.95	0.91	0.97	0.95	0.96	0.91
online loan SMS	0.90	0.90	0.89		0.89	0.94	0.91	
fraud SMS	0.78	0.86	0.80		0.90	0.83	0.86	
offer SMS	0.95	0.91	0.92		0.94	0.87	0.90	
operator SMS	0.90	0.88	0.89		0.84	0.93	0.89	
gambling SMS	0.79	0.64	0.67		0.89	0.68	0.74	

Based on the average classification results shown in Table 8, the model obtained an accuracy of 91% using IndoBERT base-p1 and IndoBERT base-p2. Although it obtained a good accuracy value, it was not followed by good predictions in minority classes, such as in the "gambling SMS" class when using IndoBERT base-p1, which gave results that have a 79% precision value, 64% recall and 67% f1-score and when using IndoBERT base-p2 which gave results that have 89% precision value, 68% recall, and 74% f1-score. This shows that there is an imbalance in recall and f1-score. With recall values of 64% and 68%, the model still seems to have difficulty identifying most gambling SMS. In addition, the f1-score, which is a metric for combining precision and recall, gives values of 67% and 74%, which shows that the F1-score results are not optimal. While the model is very sensitive in predicting the majority class, it tends to provide better performance in identifying the majority class that dominates in the dataset. This is due to class imbalance in the dataset used, where there are several classes that have a much higher frequency when compared to other classes, so the model pays less attention to minority classes. To overcome the class imbalance problem, the research is continued by augmenting data in minority classes to make the data amount balanced in all classes.

3.3. The Results of the Evaluation Model with EDA

The SMS spam classification process is continued with another experiment to overcome the problem of imbalance in class data by performing data augmentation on minority classes. The minority class data that will be augmented are ”online loan SMS,” ”fraud SMS,” and ”gambling SMS” classes. The data augmentation process uses the EDA technique by increasing the number of samples in the minority class so that the amount will be close to the majority class so that the model can learn patterns in the class evenly. The EDA process focuses on the minority class of train data so that there is a change in the amount of data in train data. This is in line with [24] which performs augmentation on the training set. The classification process is carried out using the IndoBERT through a 5-fold cross-validation scheme. Model performance data at each fold is shown in Table 9.

Based on the experiments conducted in Table 9, data training using IndoBERT base-p1 gives the highest accuracy of 94% on fold 4 and 94% when using IndoBERT base-p2 on folds 1 and 4. This shows an increase in model performance after adding data to the minority class using the EDA technique. This is in line with [24] which results in an increase after EDA is applied. The average classification results with the EDA technique are calculated to produce a classification report that reflects the overall performance of the model. This average is done to get a more stable and consistent evaluation in measuring model performance against different datasets at each iteration. The average model performance results from the classification process after using the EDA technique are shown in Table 10.

Table 9. Classification Results with EDA Technique

Fold	Class	IndoBERT base-p1				IndoBERT base-p2			
		Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy
Fold 1	personal SMS	0.99	0.93	0.96	0.93	0.99	0.93	0.96	0.94
	online loan SMS	1.00	1.00	1.00		0.94	1.00	0.97	
	fraud SMS	0.90	0.96	0.93		0.93	0.93	0.93	
	offer SMS	0.92	0.93	0.92		0.92	0.95	0.93	
	operator SMS	0.89	0.92	0.90		0.90	0.93	0.91	
	gambling SMS	0.89	1.00	0.94		1.00	1.00	1.00	
Fold 2	personal SMS	0.99	0.93	0.96	0.93	0.99	0.93	0.96	0.93
	online loan SMS	1.00	0.94	0.97		0.94	0.94	0.94	
	fraud SMS	0.96	0.85	0.90		1.00	0.85	0.92	
	offer SMS	0.96	0.93	0.94		0.94	0.93	0.93	
	operator SMS	0.86	0.96	0.91		0.88	0.97	0.92	
	gambling SMS	1.00	0.75	0.86		1.00	0.62	0.77	
Fold 3	personal SMS	0.92	0.94	0.93	0.92	0.94	0.94	0.94	0.92
	online loan SMS	0.79	1.00	0.88		0.75	1.00	0.86	
	fraud SMS	0.80	0.86	0.83		0.74	0.82	0.78	
	offer SMS	0.96	0.94	0.95		0.96	0.93	0.94	
	operator SMS	0.90	0.87	0.88		0.91	0.89	0.90	
	gambling SMS	0.89	1.00	0.94		0.88	0.88	0.88	
Fold 4	personal SMS	0.96	0.96	0.96	0.94	0.98	0.96	0.97	0.94
	online loan SMS	0.88	1.00	0.94		1.00	1.00	1.00	
	fraud SMS	0.90	0.93	0.91		1.00	1.00	1.00	
	offer SMS	0.93	0.96	0.94		0.93	0.93	0.93	
	operator SMS	0.93	0.90	0.92		0.89	0.93	0.91	
	gambling SMS	1.00	0.78	0.88		1.00	0.56	0.71	
Fold 5	personal SMS	0.96	0.96	0.96	0.93	0.92	0.98	0.95	0.93
	online loan SMS	0.80	1.00	0.89		0.84	1.00	0.91	
	fraud SMS	0.95	0.71	0.82		0.96	0.82	0.88	
	offer SMS	0.93	0.94	0.93		0.96	0.91	0.94	
	operator SMS	0.91	0.92	0.92		0.92	0.90	0.91	
	gambling SMS	1.00	0.89	0.94		0.90	1.00	0.95	

Table 10. Average classification result with EDA technique

Class	IndoBERT base-p1				IndoBERT base-p2			
	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy
personal SMS	0.96	0.94	0.95	0.93	0.96	0.95	0.96	0.93
online loan SMS	0.89	0.99	0.94		0.89	0.99	0.94	
fraud SMS	0.90	0.86	0.88		0.93	0.88	0.90	
offer SMS	0.94	0.94	0.94		0.94	0.93	0.93	
operator SMS	0.90	0.91	0.91		0.90	0.92	0.91	
gambling SMS	0.96	0.88	0.91		0.96	0.81	0.86	

Based on the classification results shown in Table 10, after the EDA process, it can be seen that the performance of the classification model using both IndoBERT base-p1 and IndoBERT base-p2 has improved significantly, especially in classes that were previously in the minority, such as ”online loan SMS,” ”fraud SMS,” and ”gambling SMS” classes. The increase occurred in the precision, recall, and f1-score values. The improvement indicates that the model has recognized the classes previously as a minority class to achieve good performance. In addition, the accuracy owned by the model also shows an increase, which was originally 91%, increased by 2% to 93% when using IndoBERT base-p1 and IndoBERT base-p2. Therefore, this research shows that the EDA technique not only solves the problem of imbalanced classes in the data but also can improve the model's performance.

3.4. Performance Patterns: A Visual Summary Results

In this subsection, a few charts that visualize the performance of the multiclass SMS spam classification model will be presented to make it easier to compare the performance of the model and the improvements that occur. Visualization is performed by presenting f1-score comparison graphs when using IndoBERT base-p1 and IndoBERT base-p2, as shown in Figure 4 and Figure 5. F1-score was chosen to see model performance comparison because this metric provides a balanced display between precision and recall. The comparison graph of the f1-score value before and after EDA when using IndoBERT base-p1 is shown in Figure 4.

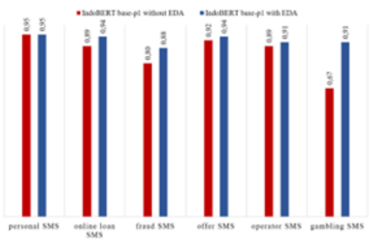

Figure 4. F1-Score Comparison Before and After EDA using IndoBERT Base-p1

Figure 5 shows a comparison graph of f1-score values using IndoBERT base-p1 without applying EDA techniques and by applying EDA techniques. From the graph, it can be seen that there is an increase in each class after applying the EDA technique. Then, when viewed from the minority class, which is the object for the EDA technique, it shows the f1-score results, with an average increase in f1-score from the three minority classes of 12%. When observed from each class, it was found that the highest increase occurred in the ”gambling SMS” class, where before applying the EDA technique, the f1-score value in that class was 67%, and after the application of the EDA technique, the f1-score value increased by 24% to 91%. The comparison graph of F1-Score values before and after EDA when using IndoBERT base-p2 is shown in Figure 5.

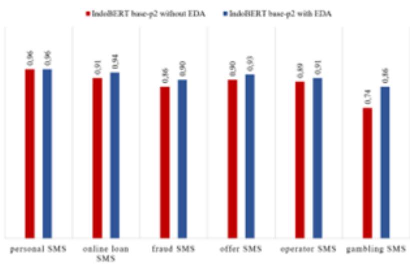

Figure 5. F1-Score Comparison Before and After EDA using IndoBERT Base-p2

Figure 5 shows a comparison graph of f1-score values using IndoBERT base-p2 with and without EDA. From the graph, it can be seen that there is an increase in each class after applying EDA. Then, when viewed from the minority class, which is the object for EDA, the f1-score results show an average increase in the f1-score from the three minority classes of 6%. When observed from each class, it was found that the highest increase occurred in the ”gambling SMS” class by 12%, where before EDA was applied, the f1-score value in that class was 74%, and after EDA implementation, the f1-score value became 86%. Furthermore, based on Table 8 and Table 10, the model performance from precision, recall, and f1-score comparisons in the minority class, both using IndoBERT base-p1 and IndoBERT base-p2 without using EDA and by using EDA is evaluated to understand the impact of data augmentation techniques with EDA on model performance. The table shows a general improvement in performance across various performance metrics, although components are experiencing stagnation and decline. It can be seen that there are variations in the precision value changes in each class after implementing the EDA technique. There are class that has no change in precision value but still have a good precision number, there are classes that experience a decrease that is not too far from the previous value, and there are several classes that experience a significant increase in precision value such as in the ”fraud SMS” class which previously had a precision of 78% increased to 90% and in the ”gambling SMS” class which previously had a precision value of 79% increased to 96% when using IndoBERT base-p1. Likewise, when using IndoBERT base-p2, there was an increase in several classes, such as in the class of ”fraud SMS” increasing from 90% to 93%, “operator SMS” increasing from 84% to 90%, and ”gambling SMS” increasing from 89% to 96%, so that after the EDA technique was applied, it gave almost evenly distributed precision values in all classes.

Meanwhile, when it is observed from a recall perspective, the results show a variation in the change of recall value in each class after applying EDA. In the IndoBERT base-p1 variation, most classes experienced an increase of recall value, such as in the ”online loan SMS” class, which increased from 90% to 99%, the ”offer SMS” which increased from 91% to 94%, the ”operator SMS” class which increased from 88% to 91%, and the ”gambling SMS” class increased from 64% to 88%. Likewise, the IndoBERT base-p2 variation shows variations in recall value changes where, after applying EDA, there is also an increase in recall value in most classes. The increase of recall value occurs in the ”online loan SMS” class from 94% to 99%, the ”fraud SMS” class increases from 83% to 88%, the ”offer SMS” class increases from 87% to 93%, and the ”gambling SMS” class from 68% increases to 81%. Although there is a decrease and no change in the recall value in some other classes, the results after EDA were applied show that the pattern of recall values in all classes is almost evenly distributed with good values.

The findings of this research are that IndoBERT provides success in performing SMS spam classification to solve the SMS spam problem, and EDA effectively overcomes imbalanced data, improving text classification performance. Generally, before EDA was applied, the accuracy when using IndoBERT base-p1 or IndoBERT base-p2 was 91%. The good accuracy results show that IndoBERT can be used well in text classification in SMS spam detection. This is in line with several studies [17, 20] by looking at the use of IndoBERT in text classification and giving good accuracy. Then, after performing EDA, there was a significant improvement, where the accuracy increased to 93% on both IndoBERT base-p1 and base-p2 models. This reflects that EDA positively improves the model’s performance in classifying SMS spam more accurately. Although the difference between the IndoBERT base-p1 and the IndoBERT base-p2 is not very significant, they both show similar improvements after the data augmentation process. The consistent increase in accuracy in both model variants shows that EDA effectively improves model accuracy in classifying different categories of SMS spam. The problem of data imbalance is successfully overcome using EDA and can even improve model accuracy, which in this case is **in line with research** [24].

Based on the literature search, this research obtained better model accuracy than previous research [31] which tried with machine learning approaches Naïve Bayes and KNN with 67% accuracy for Naïve Bayes and 70% for KNN. Furthermore, there is a difference in the dataset used, which in this research uses a larger dataset than in [31] with a dataset of 569 and only divided into three categories. The difference shows that this research has higher complexity regarding the number of diverse categories and data volume but achieved good and improved accuracy after combining with EDA. However, the classification model built using IndoBERT faces challenges in recognizing specific messages with a minor dataset, such as in the ”gambling SMS” class; before applying EDA, the model had a low recall value; this shows that the model has difficulty in identifying most of the messages that are categorized in that category. Thus, to overcome this, future research can try to apply other augmentation techniques, such as SMOTE, Back Translation, and An Easier Data Augmentation (AEDA) technique, to help improve model performance against imbalanced data in multiclass SMS spam classification. This research contributes to developing a model that can identify SMS spam. This research implies that it can provide references to the implementation of IndoBERT in other text classification tasks, including combining it with EDA techniques for cases with imbalanced data problems.

4. CONCLUSION

This research evaluated the IndoBERT method's performance in addressing the SMS spam problem in six categories. It also addressed the problem of imbalanced data by using EDA to evaluating its effect on IndoBERT classification. By using IndoBERT and EDA techniques, this research successfully achieved its objective with a significant level of accuracy, which resulted in IndoBERT achieving 91% accuracy in both model variants. Although imbalanced data was also a limitation, oversampling with EDA proved effective, significantly improving precision, recall, and f1-score on minority classes. On average, the f1-score improvement was 12% from the minority classes, with the highest improvement in the ”gambling SMS” class reaching 24% when using IndoBERT base-p1. The same trend was seen in IndoBERT base-p2, with an average increase of 6% in the f1-score of the minority class and the highest increase of 12% in the ”gambling SMS” class. These performance improvements contributed to an overall model accuracy improvement of 93% after applying EDA. The novelty of this research is obtaining the ability to combine IndoBERT and EDA, which can effectively help classify spam SMS with good performance. However, there are limitations to the study, such as the IndoBERT model's inability to handle some types of messages well, which results in low performance in certain classes. Therefore, future research can compare the performance of IndoBERT using other deep learning methods in terms of SMS spam classification to gain deeper insight into effective methods to overcome the challenges of SMS spam classification. In the augmentation process, other techniques can be tried in addition to the augmentation techniques used in this study to further contribute to improving the reliability of the model against diverse and imbalanced data and improve the overall performance of multiclass SMS spam classification. Thus, this research contributes significantly to understanding and addressing the problem of SMS spam and paves the way for further research in this topic.

5. ACKNOWLEDGEMENTS

Thank you to the volunteers willing to provide the SMS on their personal cell phones to be used as a dataset. We also thank the annotators and all those who helped in this research.

6. DECLARATIONS

AUTHOR CONTIBUTION

The first author is the researcher who contributed to the preparation of the manuscript, starting from the idea, building model, results, and conclusions. The second and third authors are informatics engineering lecturers who became supervisors who contributed by providing innovative and constructive suggestions and valuable insights for this manuscript.

FUNDING STATEMENT

This research did not receive any specific grants from any other funding agency.

COMPETING INTEREST

All authors declare that there are no financial interests or personal conflicts related to the research in this manuscript.

REFERENCES

[1] F. D. Pramakrisna, F. D. Adhinata, and N. A. F. Tanjung, “Aplikasi Klasifikasi SMS Berbasis Web Menggunakan Algoritma Logistic Regression,” *Teknika*, vol. 11, no. 2, pp. 90–97, 2022, https://doi.org/10.34148/teknika.v11i2.466.

[2] A. A. N. A. Surya Utama and A. A. Sri Indrawati, “Perlindungan Terhadap Pengguna Layanan Seluler yang Terganggu dengan Adanya Short Message Service (SMS) Spam,” *Kertha Semaya : Journal Ilmu Hukum*, vol. 10, no. 9, pp. 1–10, jul 2022, https://doi.org/10.24843/ks.2022.v10.i09.p09.

[3] F. R. Suprihati, “Analisis Klasifikasi SMS Spam Menggunakan Logistic Regression,” *Jurnal Sistem Cerdas*, vol. 4, no. 3, pp. 155–160, 2021, https://doi.org/10.22219/repositor.v3i4.32080.

[4] H. Baaqeel and R. Zagrouba, “Hybrid SMS Spam Filtering System Using Machine Learning Techniques,” in *2020 21st International Arab Conference on Information Technology (ACIT)*. IEEE, nov 2020, pp. 1–8, https://doi.org/10.1109/ACIT50332.2020.9300071.

[5] E. Sankar, Y. Y. S. S. Babu, and M. Tridev, "SMS Spam Detection Using Machine Learning," *International Journal of Scientific Research in Engineering and Management (IJSREM) International Journal of Scientific Research in Engineering and Management*, vol. 7, no. 3, pp. 1–11, 2023, https://doi.org/10.55041/IJSREM18832.

[6] I. Indriyani and P. Dewanti, "Truecaller's Spam Call and SMS Blocking Solution for Surveillance on Social Media," *Jurnal Mekintek : Jurnal Mekanikal, Energi, Industri, Dan Teknologi*, vol. 13, no. 1, pp. 19–29, apr 2022, https://doi.org/10.35335/mekintek.v13i1.121.

[7] M. S. Ghofany, R. Dwiyansaputra, F. Bimantoro, and Khairunnas, "Indonesian SMS Spam Detection Using TF-RF Feature Weighting Method and Support Vector Machine Classifier," in *Proceedings of the First Mandalika International Multi-Conference on Science and Engineering 2022, MIMSE 2022 (Informatics and Computer Science) (MIMSE-I-C-2022)*, vol. 35, no. 1. Atlantis Press, 2022, pp. 117–129, https://doi.org/10.2991/978-94-6463-084-8_12.

[8] A. Theodorus, T. K. Prasetyo, R. Hartono, and D. Suhartono, "Short Message Service (SMS) Spam Filtering using Machine Learning in Bahasa Indonesia," in *2021 3rd East Indonesia Conference on Computer and Information Technology (EIConCIT)*. IEEE, apr 2021, pp. 199–203, https://doi.org/10.1109/EIConCIT50028.2021.9431859.

[9] P. A. Raharja, M. F. Sidiq, and D. C. Fransisca, "Comparative Analysis of Multinomial Naïve Bayes and Logistic Regression Models for Prediction of SMS Spam," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 6, no. 3, pp. 1–7, jul 2022, https://doi.org/10.30865/mib.v6i3.4019.

[10] M. H. S. Ajat, "Klasifikasi SMS Spam Dengan Komparasi Metode SVM dan Naïve Bayes," *METHODIKA: Jurnal Teknik Informatika dan Sistem Informasi*, vol. 9, no. 1, pp. 31–34, mar 2023, https://doi.org/10.46880/mtk.v9i1.1694.

[11] S. A. Sireesha, S. B. Karthik, K. Srena, S. N. Gopal, and S. K. Reddy, "SMS Spam Detection Using Machine Learning," *Scandinavian Journal of Information Systems*, vol. 35, no. 1, pp. 749–754, 2023, https://doi.org/10.5281/zenodo.7807440.

[12] A. N. R. Hasanah, R. A. Krestianti, and S. Wati, "Implementasi Algoritma Regresi Logistik untuk Binary Classification dalam Spam SMS dan WhatsApp," in *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, vol. 7, no. 1, 2023, pp. 80–93, https://doi.org/10.29407/inotek.v7i1.3413.

[13] A. Kurniasih and L. P. Manik, "On the Role of Text Preprocessing in BERT Embedding-based DNNs for Classifying Informal Texts," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 6, pp. 927–934, 2022, https://doi.org/10.14569/IJACSA.2022.01306109.

[14] H. Jayadianti, W. Kaswidjanti, A. T. Utomo, S. Saifullah, F. A. Dwiyanto, and R. Drezewski, "Sentiment analysis of Indonesian reviews using fine-tuning IndoBERT and R-CNN," *ILKOM Jurnal Ilmiah*, vol. 14, no. 3, pp. 348–354, dec 2022, https://doi.org/10.33096/ilkom.v14i3.1505.348-354.

[15] M. V. Koroteev, "BERT: A Review of Applications in Natural Language Processing and Understanding," *arXiv preprint arXiv:2103.11943*, pp. 1–18, mar 2021, https://doi.org/10.48550/arXiv.2103.11943.

[16] H. M. Lee and Y. Sibaroni, "Comparison of IndoBERTweet and Support Vector Machine on Sentiment Analysis of Racing Circuit Construction in Indonesia," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 7, no. 1, pp. 99–106, 2023, https://doi.org/10.30865/mib.v7i1.5380.

[17] S. M. Isa, G. Nico, and M. Permana, "Indobert for Indonesian fake news detection," *ICIC Express Lett*, vol. 16, no. 3, pp. 289–297, 2022, https://doi.org/10.24507/icicel.16.03.289.

[18] N. N. Qomariyah, T. Sun, and D. Kazakov, "NLP Analysis of COVID-19 Radiology Reports in Indonesian using IndoBERT," in *2022 4th International Conference on Biomedical Engineering (IBIOMED)*. IEEE, oct 2022, pp. 65–70, https://doi.org/10.1109/IBIOMED56408.2022.9988223.

[19] B. Juarto, "Indonesian News Classification Using IndoBert," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 11, no. 2, pp. 454–460, 2023, https://doi.org/10.1109/IBIOMED56408.2022.9988223.

[20] L. Geni, E. Yulianti, and D. I. Sensuse, "Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using IndoBERT Language Models," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI)*, vol. 9, no. 3, pp. 746–757, 2023, https://doi.org/10.26555/jiteki.v9i3.26490.

[21] M. I. Amal, E. S. Rahmasita, E. Suryaputra, and N. A. Rakhmawati, “Analisis Klasifikasi Sentimen Terhadap Isu Kebocoran Data Kartu Identitas Ponsel di Twitter,” *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 8, no. 3, pp. 645–660, dec 2022, https://doi.org/10.28932/jutisi.v8i3.5483.

[22] D. A. Oyeyemi and A. K. Ojo, “SMS Spam Detection and Classification to Combat Abuse in Telephone Networks Using Natural Language Processing,” *Journal of Advances in Mathematics and Computer Science*, vol. 38, no. 10, pp. 144–156, oct 2023, https://doi.org/10.9734/jamcs/2023/v38i101832.

[23] D.-C. Li, S.-C. Chen, Y.-S. Lin, and W.-Y. Hsu, “A Novel Classification Method Based on a Two-Phase Technique for Learning Imbalanced Text Data,” *Symmetry*, vol. 14, no. 3, pp. 1–23, mar 2022, https://doi.org/10.3390/sym14030567.

[24] A. Wirawan, H. D. Cahyono, and Winarno, “Easy Data Augmentation in Sentiment Analysis of Cyberbullying,” in *2023 6th International Conference on Information and Communications Technology (ICOIACT)*. IEEE, nov 2023, pp. 443–447, https://doi.org/10.1109/ICOIACT59844.2023.10455817.

[25] H. R. Nafiisah and F. Z. Ruskanda, “Content-based Multiclass Classification on Indonesian SMS Messages,” in *2022 International Symposium on Electronics and Smart Devices (ISESD)*. IEEE, nov 2022, pp. 1–6, https://doi.org/10.1109/ISESD56103.2022.9980769.

[26] R. Dwiyansaputra, G. S. Nugraha, F. Bimantoro, and A. Aranta, “Deteksi SMS Spam Berbahasa Indonesia menggunakan TF-IDF dan Stochastic Gradient Descent Classifier,” *Jurnal Teknologi Informasi, Komputer, dan Aplikasinya (JTIKA)*, vol. 3, no. 2, pp. 200–207, 2021, https://doi.org/10.29303/jtika.v3i2.145.

[27] G. Z. Nabiilah, I. N. Alam, E. S. Purwanto, and M. F. Hidayat, “Indonesian multilabel classification using IndoBERT embedding and MBERT classification,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 1, p. 1071, feb 2024, https://doi.org/10.11591/ijece.v14i1.pp1071-1078.

[28] B. Wilie, K. Vincentio, G. I. Winata, S. Cahyawijaya, X. Li, Z. Y. Lim, S. Soleman, R. Mahendra, P. Fung, S. Bahar, and A. Purwarianti, “IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding,” *arXiv preprint arXiv:2009.05387*, sep 2020, https://doi.org/10.48550/arXiv.2009.05387.

[29] S. Saadah, K. M. Auditama, A. A. Fattahila, F. I. Amorokhman, A. Aditsania, and A. A. Rohmawati, “Implementation of BERT, IndoBERT, and CNN-LSTM in Classifying Public Opinion about COVID-19 Vaccine in Indonesia,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 4, pp. 648–655, aug 2022, https://doi.org/10.29207/resti.v6i4.4215.

[30] P. T. Putra, A. Anggrawan, and H. Hairani, “Comparison of Machine Learning Methods for Classifying User Satisfaction Opinions of the PeduliLindungi Application,” *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 22, no. 3, pp. 431–442, jun 2023, https://doi.org/10.30812/matrik.v22i3.2860.

[31] E. D. Pratama, “Implementasi Model Long-Short Term Memory (LSTM) pada Klasifikasi Teks Data SMS Spam Berbahasa Indonesia,” *The Journal on Machine Learning and Computational Intelligence (JMLCI)*, vol. 1, no. 2, 2022, https://doi.org/10.26740/vol1iss2y2022id12.