Classifying Fraud in Indonesia Short Message Service (SMS) Using Machine Learning and Deep Learning

1st Jefry Christiawan
Computer Science Department
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
jefry.christiawan@binus.ac.id

2nd Collin Kliveson
Computer Science Department
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
collin.kliveson@binus.ac.id

4th Derwin Suhartono
Computer Science Department
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
dsuhartono@binus.edu

3rd Henry Lucky
Computer Science Department
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
henry.lucky@binus.ac.id

***Abstract*—In the era of internet-based communication, Short Message Service (SMS) remains a critical platform for businesses and personal communication, particularly in Indonesia where smartphone penetration is high. However, the increase in SMS usage has led to a rise in fraudulent activities, posing significant risks to users. This research aims to enhance SMS classification in Indonesia by utilizing advanced machine learning and deep learning techniques. This research compares and investigates the effectiveness of various models, including Hybrid CNN-LSTM, Naive Bayes, Support Vector Machine (SVM), and advanced models like IndoBERT. Our approach involves classifying SMS into three categories: Fraud, Advertisement, and Normal messages, using a dataset of 1143 Indonesian SMS. The results demonstrate that combining several traditional machine learning methods into ensemble learning can significantly improve classification accuracy. Our hybrid model approach, especially the stacking ensemble learning combining CNN-LSTM, Support Vector Machine (SVM), and Naive Bayes, shows the best performance with over 94% score for accuracy, precision, recall, and F1-score. This research shows the potential of hybrid models in enhancing the accuracy and reliability of SMS classification, therefore contributing to a more secure and user-friendly SMS communication in Indonesia.**

Keywords—SMS Classification, BERT, Machine Learning, Deep Learning, Support Vector Machine, Naïve Bayes, CNN-LSTM, Ensemble Learning

I. INTRODUCTION

In the era of internet-based instant messaging applications, Short Message Service (SMS) still has become a highly popular communication platform. With just a Valid phone number, SMS become a powerful tool for businesses to reach their audience and communicate. Moreover, SMS apps are built in every mobile phone which means that receiving SMS messages is integrated into all mobile devices worldwide. As of 2022, around 67% of the population in Indonesia owns smartphones which means they have capability to receive messages from Short Message Services (SMS) [1]. The trend is expected to continue increasing until 2028 [2].

Recently, due to a large number of users, there has been a significant increase in fraudulent activity messages. During the period from August to mid-November 2023, the Ministry of Communication and Information Technology received reports of 958 cases of telephone and SMS misuse for online fraud [3]. The methods used by fraudsters are diverse, with one common approach being smishing text messages, typically sent by fraudsters via SMS in various formats. However, their ultimate goal remains consistent with stealing our personal data. Usually the content of smishing text messages include a link that may request your personal information or sometimes download files or applications automatically [4].

Numerous studies have explored the implementation of various machine learning techniques to classify spam messages effectively. Among the methodologies explored are traditional approaches such as the Naive Bayes Classifier (NB) [5], Support Vector Machine (SVM) [6], and Random Forest Algorithm [7]. Moreover, advancements in deep learning have led to the exploration of sophisticated models such as Bidirectional Encoder Representations from Transformers (BERT) [8] and the cutting-edge Generative Pre-trained Transformer 3 (GPT-3) model [9]. These models have demonstrated impressive capabilities in tasks ranging from natural language understanding to text generation, offering promising avenues for spam detection and classification. Additionally, there are models designed for specific languages, such as IndoBERT for Indonesian language processing [8], which means we have more tools used for classifying SMS.

However, most studies mainly concentrate on finding and blocking spam messages. They often miss a crucial point that "not all unwanted messages are the same". Some people might be okay with getting ads or marketing stuff, but they really don't like fraudulent messages trying to trick or harm them [10]. Understanding this difference is important for making better SMS Classification that suit users' needs.

The primary goal of our research is to develop a more effective SMS classifier using machine learning and deep learning from the previous research for classifying Indonesia SMS into 3 categories (Fraud, Advertisement, and Normal) messages. Although deep learning gives a high accuracy on classifying SMS, we also see that some hybrid models of machine learning that combined with deep learning manage to give a higher accuracy [9]. Because of that, in this research,

we try to use IndoBERT for word embedding in Indonesian SMS and hybrid machine learning model as a SMS Classifier.

II. Literature Review

Several machine learning based classification methods have been proposed over the past five years. In the field of SMS Classification, some of these approaches are Naive Bayes Classifier, Support Vector Machine, Random Forest Algorithm, Natural Language Processing, Convolutional Neural Network, Long Short-Term Memory, etc.

A work [5] shows that Naive Bayes Classifier for SMS Spam classification, reaches an accuracy of 99% when evaluating a dataset comprising 5574 messages. This high accuracy is noteworthy, especially considering that it was achieved without utilizing deep learning techniques. As time has passed, the way we classify SMS messages has changed. Now, researchers have begun to explore hybrid methodologies to get better results. They're mixing traditional ways of classifying messages with fancy new computer techniques. This mix, or hybrid approach, seems to be promising because it takes the best parts of both worlds. As comparison, another studies [11] that combine Naïve Bayes with Neural Networks manage to increase the accuracy of using only Naïve-Bayes Classifier. However the accuracy still lower with 97% even with the same dataset. Another works [7] with same dataset using random forest algorithm achieve over 97% in accuracy, precision, and F-measure. This research gives the highest accuracy compare to other works if using TF-IDF in processing data.

Following this trend, some researchers tried something new by mixing the Naive Bayes Algorithm with BERT [12]. This mix gave good results, with performance of 97-99% in accuracy, precision, and recall. It also did well in remembering things correctly and not missing important stuff. This success shows that combining old-fashioned methods with super-smart new ones can make classification SMS messages more accurate. Another experiment [13] in filtering SMS also done in China with over 90% in accuracy using Naïve Bayes in 4 times experiment. Another research [14] combined four BERT models (BERT, DistilBERT, RoBERT, and SpanBERT) with Random Forest, Decision Tree, and SVM. And the result show that a combination of DistilBERT and SVM achieved the best scores of all evaluation metrics with approximately 97% in accuracy, precision, recall, and F1-Score and 94% in Area Under the Curve Score.

Additionally, when researchers looked into Support Vector Machines (SVM) [6], they found out it's really good at classifying spam messages. On average, it gets it right about 98.9% of the time. SVMs have become like the strong and reliable guardians of SMS classification, working well with all classification of different message collections. A research [15] use One Class-Support Vector Machine (OC-SVM) and it achieved an accuracy and f1 score percentage in 98%, this is the best result compared to other models used in this research including normal SVM itself. Another Work with discrete HMM [16] shows 95.9% accuracy but bad in precision and recall which is around 80%.

Another interesting research [17] used Convolutional Neural Network (CNN) and it obtained an accuracy score of 98.4%, AUC of 95.5%, and F1 score of 98.3%. At the same time, when scientists tried combining CNN and LSTM [18], they found it worked pretty well too. They manage to achieve above 90% in some evaluation metrics, such as accuracy, precision, F1-Score, and AUC. However, the percentage of recall quite low compared with others which is 87.87% in performance. Another work also done to Swahili dataset [19] with hybrid CNN-LSTM and manage to outperform other models with the highest accuracy of 99.98%.

With the previous result makes deep learning become more utilized than machine learning in classification. Research that comparing LSTM and Gated Recurrent Unit (GRU) [20] showed over 98% accuracy and outperform machine learning. Another innovation work with their proposed transformer model [21] by encoding SMS manage to get accuracy of 98.52%, precision in 97.81%, recall in 94.51% and F1-Score of 96.13%.

Other researchers also try to solve this problem with new approach. They think that feature extraction is the most essential part to increase the performance of spam classification. Their work [22] combined three techniques (CNN, LSTM, and TF-IDF) to create feature extraction model. This method manage to get over 99% in all evaluation metrics.

A research [23] try an approach of making a hybrid system by combining K-means with Naïve Bayes, Logistic Regression, and SVM. The result shows that the hybrid system of K-means and SVM achieved the highest accuracy of 98.8%. But the highest precision is achieved by SVM individually. Another research [24] chose to compare several machine learning models individually instead of making a hybrid model, this research compared K-Nearest Neighbor (KNN), Multilayer Perceptron (MLP), and Linear Support Vector Machine (LSVM). The result shows that MLP achieved an accuracy of 98.18%, KNN achieved an accuracy of 98.15%, and LSVM achieved an accuracy of 98.13%. It means MLP is performed slightly superior to KNN and LSVM in terms of accuracy but the difference is really small.

Furthermore, ensemble learning also popular in researchers to build a new model independently by combining machine learning and deep learning models. This is different with hybrid model because ensemble work independently, while hybrid model work together to predict something. Research done with GPT-3 as embedding and ensemble model [9] which made from SVM, KNN, LightGBM, and CNN manage to reach over 99% in accuracy, precision, recall, and F1-Score.

Looking at what's been happening lately in Indonesia, researchers have explored IndoBERT [8], a variant of the BERT model fine-tuned on Indonesian language. Studies employing IndoBERT have showcased its superior performance, particularly when coupled with data augmentation techniques such as EDA (Easy Data Augmentation). IndoBERT achieved accuracy of 99.35% and 99.28% respectively when evaluated using the Average 10-fold cross validation accuracy technique and Train test split 80 : 20 method. This shows that IndoBERT is very good at classifying SMS messages, And as scientists keep working on this stuff, combining different methods will probably help make SMS classification even better. Another research [25] with BERT using 4 dataset with total approximately 30,000 thousand manage to perform over 97% in every dataset. This shows that BERT is the most powerful tools now in handling text classification.

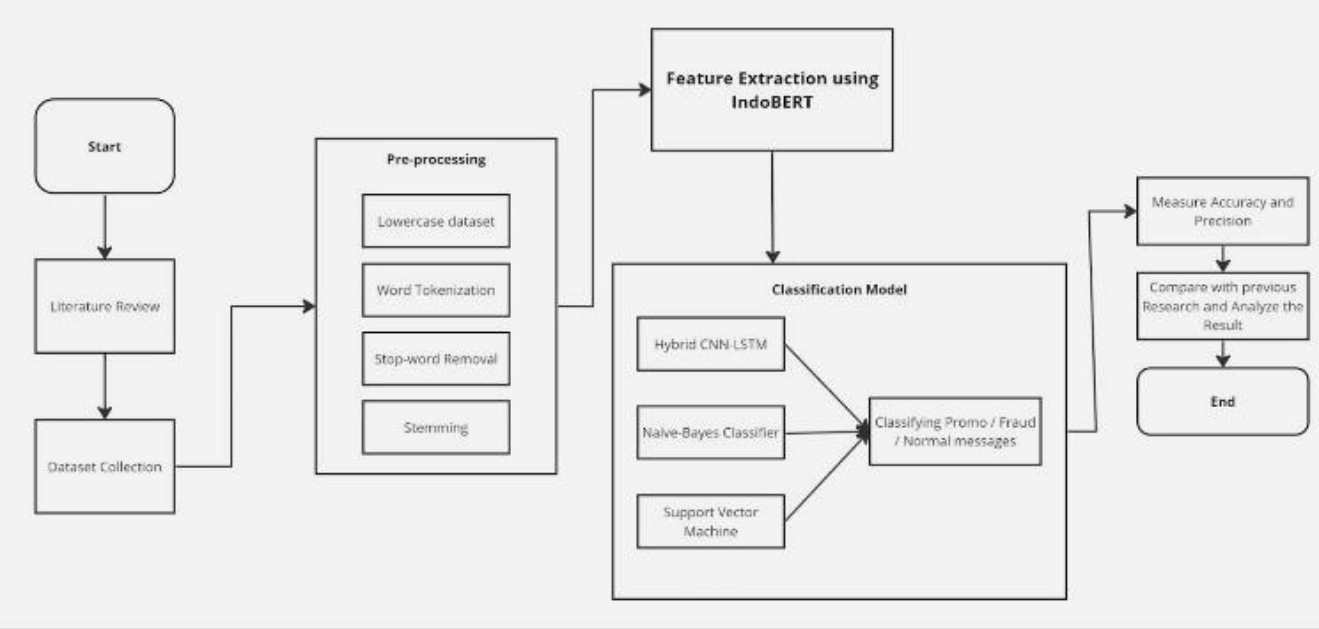

Fig. 1. Methodology Process

III. METHODOLOGY

A. Dataset

The dataset used in this paper is obtained from github and consists of a collection of Indonesian SMS. The dataset contains 1,143 Indonesian SMS text with three classes: normal, Fraud, and advertisement. The dataset consists of 569 normal messages, 335 Fraud messages, and 239 advertisement messages.

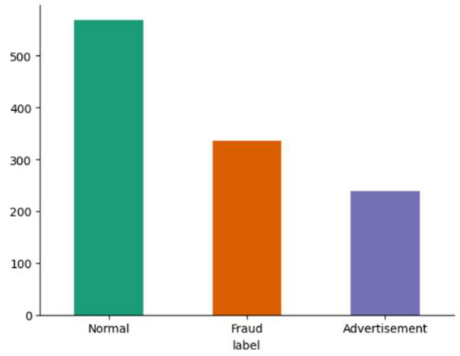

Fig. 2. Dataset

B. Pre-processing

Before utilizing the dataset for feature extraction and training, our initial step involves preprocessing it, which includes converting the dataset to lowercase, then do word tokenization to break down the text into individual tokens, then implementing stop-word removal to eliminate common word from the text, and then lastly, we do stemming in order to remove similar words.

1) *Lowercase Dataset:* In Lowercase Dataset, every words in the dataset are transformed into lowercase. This ensures that the text is uniform and consistent for further processing.

2) Word Tokenization: In Word Tokenization, the provided data or information are segmented into tokens. Basically it split sentences into words. For example a word "Saya adalah Binusian", it divides into tokens of words "Saya", "adalah", "Binusian".

C. Feature Extraction

In order utilizing the Indonesian SMS data for the model we need to transform it, from text data to numerical features. There are multiple techniques available for transforming text into a numerical feature, but we decided to use BERT because of its powerful language understanding capabilities. Here, we use the IndoBERT because our dataset is in Indonesian language.

BERT transforms text data into numerical features through a deep learning approach. The process involves several steps:

1) *Tokenization:* convert sentences into words that the model can understand. This is done using a tokenizer that splits the input text into tokens and then maps those tokens to their corresponding indexes in BERT's vocabulary.

2) *Embedding:* convert each token into a numerical vector through an embedding process. BERT utilizes multiple embeddings such as token embeddings, segment embeddings, and position embeddings.

D. Classification

For ensuring the reliability of our classification model, we split the dataset into two portions: one for training and the others for testing, with a split of 8:2 because our dataset size is not very big. Our method involves combining different techniques from Hybrid CNN-LSTM, Naive-Bayes Classifier, and Support Vector Machine algorithms to form a ensemble learning. By using ensemble learning, We aim to enhance the model's capability to understand the SMSmessages. This way, we can capture small details in the text data, which should help us classify Indonesian SMS messages more accurately.

1) *Hybrid CNN-LSTM:* Hybrid CNN-LSTM (Convolutional Neural Network - Long Short Term Memory) is a deep learning method that combine the advantages of using CNN and LSTM. CNN can be used for extract features, such as sequences of words or characters in this case, While LSTM are good at understanding sequences over long periods, which makes them great for tasks like translating languages, understanding speech, and predicting trends in time-based data.

Application in our case, starting from the CNN Layer that extracting features from SMS with convolutional layers. Convolutional filters will detect local patterns and features within word embedding. Next, continue to Max Pooling Layer with a pool size of 4 will reduce the dimentionality of feature maps in order to get most leading feature and make the computation become efficient. After that step is done, we continue with using LSTM for capturing contextual information within the sequences. also LSTM can remember important information over long sequences and discarding irrelevant information. Finally, flatten the output and Connect the flattened output to a dense layer with 3 neurons, corresponding to 3 classes (Fraud, Advertisement, normal) by using a softmax activation to produce probabilities for each class.

2) *Naïve-Bayes Classifier:* Naive-Bayes is a well suited for classification task such as SMS Classification due to its efficiency and simplicity. In this case, it involves categorizing text messages into three classes: fraud, advertisement and normal messages. Naive-Bayes is particularly effective for this task because it works well with small datasets like SMS data, and it can also handle high-dimensional feature spaces efficiently. At its core, Naive Bayes operates based on Bayes Theorem. which calculates the posterior likelihood of a class given an input (P(A|B)), using prior probabilities (P(A)) and likelihood probabilities (P(B|A)).

In the application of classification SMS, we calculate posterior probability of each class (Fraud, advertisement, normal) given word message as an input. The formula will be considering multiple words in every messages and also based on strong independent word assumptions.

Parameters used in naïve Bayes Classifier can be seen in Table I.

TABLE I. NAÏVE BAYES PARAMETERS

Parameters	Value
Distribution assumptions	Gaussian
Var smoothing	1e-6

3) *Support Vector Machine (SVM)*

Support Vector Machine (SVM) belongs to a category of machine learning algorithms that utilized to understand data as well as identifying patterns within it. While it's originally designed for numbers, SVM can also handle categories by converting them into numbers. Before using the data, SVM makes sure it's in a standard format. The primary objective of SVM is to discover an optimal hyperplane for the separation of distinct groups in the data, like distinguishing between fraud, advertisement and normal messages in our case. This hyperplane, identified by maximizing the margin between various classes, serves as a dividing line. SVM is capable of managing data that can be separated linearly and non-linearly by utilizing diverse kernel functions for feature space transformation. By introducing a margin around the decision boundary, SVM accommodates potential errors while aiming for superior classification accuracy. SVM is good at both simple and complex classification tasks and works by minimizing potential risks. It's considered one of the best methods for sorting text into different categories.

The list of parameters that we will use for SVM can be seen in Table II.

TABLE II. SVM PARAMETERS

Parameters	Value
Kernel Type	Radial Basis Function
Regularization Parameters	1.0
Gamma	Scale
Class Weights	- Normal: 0.67 - Fraud: 1.14 - Normal: 1.59

E. Evaluation Metrics

For the evaluation metrics, we will compare our model performance with the model that exists from previous research. The result of the model's performance will be measured by some measurements, such as:

1) *Accuracy.* Accuracy is a common used evaluation metrics in evaluating the correctness on our classification model. This metrics is very common and straightforward, however they have a limitation in measuring correctness model if the data is imbalanced. We still can achieve high accuracy even the model is mostly wrong classified the minority classes.

2) *Precision.* In terms of imbalanced dataset, precision is very good in understanding the model performaces in identifying minority classes without falsely labeling instances from majority class.

3) *Recall.* Recall focuses on measuring model's ability in capturing positive instances. Like Precision, recall is also good for imbalanced dataset because it can measure the accuracy of minority class without being affected by the majority.

4) *F1-Score.* Combining precision and recall makes F1-Score become a powerful evaluation metrics in measuring model's performance in imbalanced dataset.

TABLE III. MODEL COMPARISON

Models	Evaluation Metrics			
	Accuracy	Precision	Recall	F1-Score
Voting Ensemble Learning	93.01%	93.77%	93.01%	93.20%
Stacking Ensemble Learning	**94.32%**	**94.64%**	**94.32%**	**94.42%**
Bagging Ensemble Learning	93.89%	94.52%	93.89%	94.05%
Support Vector Machine	**94.32%**	**94.66%**	**94.32%**	**94.42%**
Naïve Bayes	93.89%	94.51%	93.89%	94.05%
Hybrid CNN-LSTM	86.46%	88.07%	86.46%	86.70%

IV. EXPERIMENTS AND RESULT

A. Model Comparison

There is no significant difference between ensemble learning and single machine learning model. We have already done experiments on 6 different models: Voting Ensemble Learning, Stacking Ensemble Learning, Bagging Machine Learning, Support Vector Machine, Naïve-Bayes Classifier, and Hybrid CNN-LSTM. All of which have achieved almost the same performance with around 93-95% in all four evaluation metrics, with exception of hybrid CNN-LSTM in classifying Indonesian SMS into 3 labels: Fraud, Advertisement, and normal messages.

From Table III, we can see that stacking ensemble learning that made from hybrid CNN-LSTM, SVM and Naïve-Bayes then use logistic regression for classification achieved the best performance. It managed to achieve 94% in all evaluation metrics. Let's take comparison with single model SVM. It achieves accuracy and F1-Score by the same with stacking. The other evaluation metrics, such as precision and recall are still dominated by SVM by only 0.01% differences. This conclude that ensemble learning does not give significant impact to the improvement of the model
performance in classifying Indonesian SMS.

B. The effect of Ensemble Learning

Ensemble itself does not have a significant impact if we compared with the performance of single machine learning model. However, ensemble learning manages to improve the performance of weak learners.

As we can see, from table III, the performance by evaluation metrics for hybrid CNN-LSTM is around 88-89%. The weakness of this model is because we have a small dataset size. Using deep learning usually requires a bigger size of dataset because deep learning captures more detailed features. Because of that, the smaller size of dataset will cause decrease in the performance.

In our case, ensemble learning with stacking method manages to outperform other ensemble learning. Nevertheless, ensemble learning does not have an impact on increasing performance if we compare it with a single machine learning model, such as Support Vector Machine and Naïve Bayes Classifier.

We can see that ensemble learning with stacking method has the same performance as Support Vector Machine, while Ensemble learning with bagging method has the same performance with Naïve-Bayes Classifier. This inferred that single machine learning model performs good at smaller dataset size rather which mean they does not need ensemble learning to increase the performance.

C. Confusion Matrix & Discussion

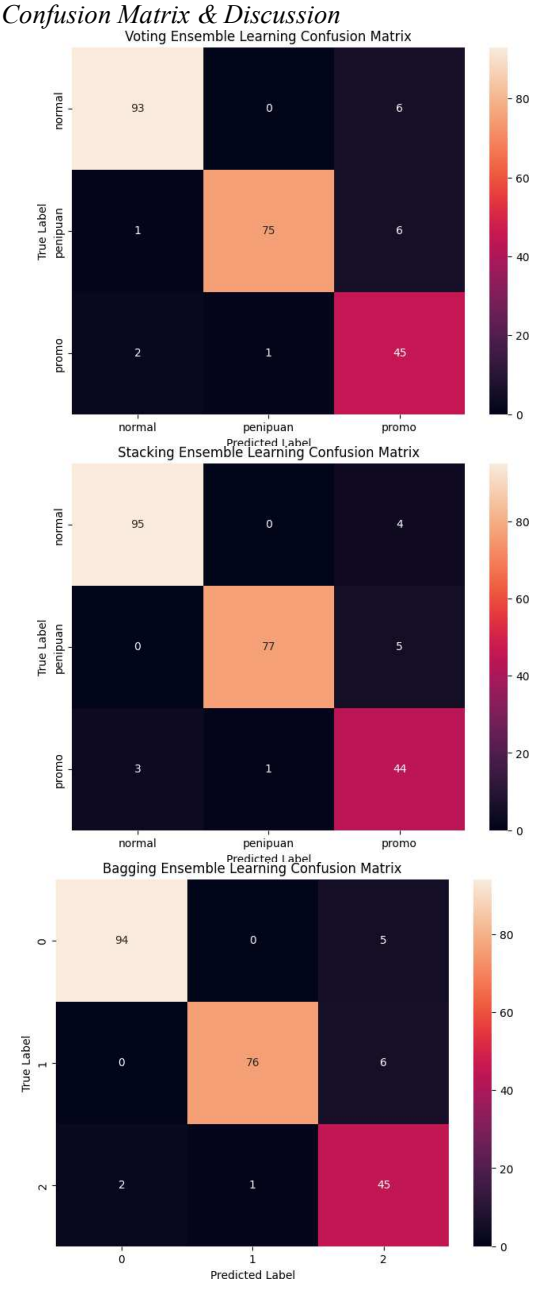

Fig. 3. Confusion Matrix of Ensemble Learning

From the confusion matrix, the dataset looks small after splitting for train & test. It's only about 200 data of SMS. Because of that, we believe that if we can get more data about SMS, we can manage to get a higher performance model.

Based on confusion matrix shown in Fig 3, we can infer that ensemble learning with stacking method and using logistic regression as meta model achieved highest accuracy.

V. CONCLUSION

To conclude, Ensemble Learning with Stacking method managed to achieve the best performance among other ensemble learning models. Also, Ensemble learning with stacking did enhance weaker model which is Hybrid CNN-LSTM although they still cannot outnumber single learning model in performance. The Ensemble with stacking managed to achieve 94% in accuracy, precision, recall, and F1-Score.

Additionally, our model manages to get consistent in performance percentage. This is because BERT embeddings itself that are really good at embedding the text. So even with the model that is not in good performance which is hybrid CNN-LSTM. They still manage to handle imbalance dataset and show only 1-2% differences in their performance.

REFERENCES

[1] A. 'Ahdiat, "67% Penduduk Indonesia Punya Handphone pada 2022, Ini Sebarannya.".

[2] "Number of smartphone users in Indonesia from 2018 to 2028," Statista. Accessed: Apr. 01, 2024. [Online]. Available: https://www.statista.com/forecasts/266729/smartphone-users-in-indonesia

[3] "Tekan Kasus Penipuan Online, Kominfo Buka AduanNomor.id," Kominfo. Accessed: Apr. 01, 2024. [Online]. Available: https://www.kominfo.go.id/content/detail/52935/siaran-pers-no-466hmkominfo112023-tentang-tekan-kasus-penipuan-online- kominfo-buka-aduannomorid/0/siaran_pers

[4] P. E. Wicaksono, "Penipuan Lewat SMS Masih Mengintai Mangsa, Kenali Ciri-Cirinya," Liputan6. Accessed: Apr. 01, 2024. [Online]. Available: https://www.liputan6.com/cek-fakta/read/5188420/penipuan-lewat-sms-masih-mengintai-mangsa-kenali-ciri-cirinya?page=2

[5] C. Bukola Asaju, J. Ekorabon, and R. Ojochegbe, "Short Message Service (Sms) Spam Detection and Classification Using Naïve Bayes," *International Journal of Mechatronics, Electrical and Computer Technology (IJMEC) (Print)*, vol. 11, no. 40, pp. 4931– 4936, 2021, [Online]. Available: https://aeuso.org

[6] N. N. A. Sjarif, Y. Yahya, S. Chuprat, and N. H. F. M. Azmi, "Support vector machine algorithm for SMS spam classification in the telecommunication industry," *Int J Adv Sci Eng Inf Technol*, vol.10, no. 2, 2020, doi: 10.18517/ijaseit.10.2.10175.N. N. Amir Sjarif, N. F. Mohd Azmi, S. Chuprat, H. M. Sarkan, Y. Yahya, and S. M. Sam, "SMS spam message detection using term frequency-inverse document frequency and random forest algorithm," in *Procedia Computer Science*, 2019. doi: 10.1016/j.procs.2019.11.150.

[7] D. Ramdhan, A. P. Kemala, and A. Chowanda, "SHORT MESSAGE SERVICE (SMS) SPAM FILTERING USING DEEP LEARNING IN BAHASA INDONESIA," *ICIC Express Letters, Part B: Applications*, vol. 13, no. 10, 2022, doi:10.24507/icicelb.13.10.1093.

[8] A. Ghourabi and M. Alohaly, "Enhancing Spam Message Classification and Detection Using Transformer-Based Embedding and Ensemble Learning," *Sensors*, vol. 23, no. 8, 2023, doi: 10.3390/s23083861.

[9] A. Soenaidy and Y. A. Sutrisno, "KAJIAN SIKAP KONSUMEN TERHADAP SMS ADVERTISING."

[10] H. H. Mansoor and S. H. Shaker, "Using classification techniques to SMS spam filter," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 12, 2019, doi: 10.35940/ijitee.L3206.1081219.

[11] 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, doi: 10.9734/jamcs/2023/v38i101832.

[12] M. Taufiq Nuruzzaman, C. Lee, M. F. A. Bin Abdullah, and D. Choi, "Simple SMS spam filtering on independent mobile phone," *Security and Communication Networks*, vol. 5, no. 10, 2012, doi: 10.1002/sec.577.

[13] C. Oswald, S. E. Simon, and A. Bhattacharya, "SpotSpam: Intention Analysis-driven SMS Spam Detection Using BERT Embeddings," *ACM Transactions on the Web*, vol. 16, no. 3, Sep. 2022, doi: 10.1145/3538491.

[14] S. Y. Yerima and A. Bashar, "Semi-supervised novelty detection with one class SVM for SMS spam detection," in *International Conference on Systems, Signals, and Image Processing*, IEEE Computer Society, 2022. doi: 10.1109/IWSSIP55020.2022.9854496.

[15] T. Xia and X. Chen, "A discrete hidden Markov model for SMS spam detection," *Applied Sciences (Switzerland)*, vol. 10, no. 14, 2020, doi: 10.3390/app10145011.

[16] M. Popovac, M. Karanovic, S. Sladojevic, M. Arsenovic, and A. Anderla, "Convolutional Neural Network Based SMS Spam Detection," in *2018 26th Telecommunications Forum, TELFOR 2018 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2018. doi: 10.1109/TELFOR.2018.8611916.

[17] A. Ghourabi, M. A. Mahmood, and Q. M. Alzubi, "A hybrid CNN-LSTM model for SMS spam detection in arabic and english messages," *Future Internet*, vol. 12, no. 9, Sep. 2020, doi: 10.3390/FI12090156.

[18] I. S. Mambina, J. D. Ndibwile, D. Uwimpuhwe, and K. F. Michael, "Uncovering SMS Spam in Swahili Text Using Deep Learning Approaches," *IEEE Access*, vol. 12, 2024, doi: 10.1109/ACCESS.2024.3365193.

[19] P. Poomka, W. Pongsena, N. Kerdprasop, and K. Kerdprasop, "SMS Spam Detection Based on Long Short-Term Memory and Gated Recurrent Unit," *International Journal of Future Computer and Communication*, vol. 8, no. 1, pp. 11–15, Mar. 2019, doi: 10.18178/ijfcc.2019.8.1.532.

[20] X. Liu, H. Lu, and A. Nayak, "A Spam Transformer Model for SMS Spam Detection," *IEEE Access*, vol. 9, 2021, doi: 10.1109/ACCESS.2021.3081479.

[21] H. A. Al-Kabbi, M. R. Feizi-Derakhshi, and S. Pashazadeh, "Multi-Type Feature Extraction and Early Fusion Framework for SMS Spam Detection," *IEEE Access*, vol. 11, 2023, doi: 10.1109/ACCESS.2023.3327897.

[22] H. Baaqeel and R. Zagrouba, "Hybrid SMS spam filtering system using machine learning techniques," in *Proceedings - 2020 21st International Arab Conference on Information Technology, ACIT 2020*, Institute of Electrical and Electronics Engineers Inc., Nov. 2020. doi: 10.1109/ACIT50332.2020.9300071.

[23] G. José De Sousa, D. Carlos, G. Pedronette, P. Papa, and I. Rizzo Guilherme, "SMS Spam Detection Through Skip-gram Embeddingsand Shallow Networks."

[24] T. Sahmoud and M. Mikki, "Spam Detection Using BERT."