Emotion Detection in Indonesian Music Comments Using BERT, Traditional Machine Learning and Their Hybrid Model

Kevin
Computer Science Department
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia 11480
kevin094@binus.ac.id

David Malcom Maximillian
Computer Science Department
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia 11480
david.maximillian@binus.ac.id

Edrick Louis
Computer Science Department
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia 11480
edrick.louis@binus.ac.id

Kartika Purwandari
Computer Science Department
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia 11480
kartika.purwandari@binus.ac.id

Abstract— **In the era of digital communication, social media comes as a platform to express honest feelings and opinions, especially in Indonesia where user engagement continues to increase. Music, often related to emotions, gives various reactions in comments on social platforms. However, detecting emotions from Indonesian-language comments is still a challenge due to the use of informal language, emoticons, etc. This research proposes and evaluates three approaches: traditional machine learning employs TF-IDF and SVD for feature extraction, Bidirectional Encoder Representations from Transformers, and a hybrid model that combines both. The results show that BERT achieves the highest accuracy (79%). The hybrid model achieves an accuracy of 77% with the training time (0.97 s). Meanwhile, logistic Regression alone offers the fastest performance (0.35 s) with the lowest memory usage but lower accuracy (64%). This study highlights the potential of hybrid approaches to balance accuracy and computational efficiency, contributing to a better understanding of music-related emotional expression in Indonesian online communities.**

Keywords—Emotion Classification, BERT, Traditional Machine Learning, Hybrid Model, Indonesian Language.

I. Introduction

In current times, social media has been an essential part of human life. People can express their opinions and views freely. That makes social media a compelling text data source, as the opinions are diverse from people with completely different views and backgrounds. Social media usage rate especially in Indonesia has increased significantly [1]. This growth presents an opportunity to understand their opinion on various matters. Several topics are popular in Indonesia such as sports, music, and entertainment.

Music and emotion are closely linked. Humans can convey their feelings such as anger, joy, and sadness through song. Music experience made it possible for humans to feel their emotions deeper [2]. However, how it is being perceived may differ as human emotions are complex. It explains why humans have different reactions to the same song as they have different situations and emotions when listening to the song. They show their opinion by reacting and commenting on songs on social media. This sentiment data is essential to understanding the impact of music and its public views.

However, there are several limitations when detecting emotion in the Indonesian language. Informal grammar, emoticon usage, and multilingual mixing pose a problem in detecting the emotion accurately. Also detecting emotions only from text suffers significantly due to the absence of expression, body language, etc. However, previous research [3] has concluded by leveraging deep learning and transformers, machines are capable of classifying complex human emotions. Bidirectional Encoder Representations from Transformers (BERT) is a transformer architecture that has been consistently tested and used for such complex emotion recognition. While previous studies have successfully applied BERT, limited studies focused on specifically Indonesian Language Comment and hybridize BERT and traditional machine learning.

To enhance classification accuracy and efficiency, this paper proposes a hybrid approach of deep learning and ensemble learning [4]. Utilizing BERT as feature extractor and traditional machine learning as classifier. BERT is suitable for embedding informal and complex words and traditional is robust and low resource cost. The purpose is to know how effective this hybrid combination is to classify emotions such as happy, sadness, fear, love, and anger [5] in Indonesian comments on music related in social media. Through this paper, we want to contribute to creating an alternative model that is accurate yet efficient in terms of time and computer cost to a specific language while giving insight into how music is perceived in Indonesia. This paper is structured as follows: Section II reviews prior work, Section III shows our method approach, Section IV shows the experimental results, and Section V summary with future directions.

II. Literature Review

A. Emotion Classification

Emotion Classification is based on the emotion model, which defines how they can represent emotion [6]. The model assumed that emotions can manifest in various situations that must be differentiated for each emotion. Many studies have shown that the important thing for this analysis like:

1. Discrete emotion models (DEMs): This model concept involves placing basic emotions that can be recognized widely. Paul Ekman's model widely used for emotion classification shows that emotion is divided into 6 fundamental emotions, which are happiness, sadness, anger, disgust, surprise, and fear [7]. This is just one of several DEMs outside that can be used to design the emotion classification, depending on how the DEMs fit into the research.
2. Dimensional emotion models (DiEMs): This model presupposes that there exists a relation between one emotion with another emotion. Thus, the dimensional emotion specifies the feelings based on a few dimensions that are determined by specific factors (Valence, Arousal, and Power) [7]. These factors are useful for analyzing emotional expressions.

Building on this research, emotion classification is used to determine the Indonesian language on music comments, where music experience made it possible for humans to express their honest feelings. Therefore, this model is needed to complete this project.

B. BERT Model for Emotion Class

BERT is a revolutionary approach model in Natural Language Processing NLP. BERT uses the encoder architecture from transformers as a main component in the pre-training process for various NLP tasks such as Sentiment Analysis (SA), Question Answering (QA), and Text Summarization (TS). BERT can understand language by training on the Masked Language Modeling (MLM) and the Next Sentence Prediction (NSP) mechanisms [3]. In MLM, some words in a sentence will be masked, and the model will learn to guess the missing words from the context around it, so that the model will be able to understand bi-directional context.

BERT has shown the best performance on various NLP tasks without requiring any significant architectural modifications for specific tasks. This demonstrates BERT's flexibility and power in understanding complex emotional contexts from text [8]. It can be concluded that BERT has great potential in emotion classification tasks, especially because of BERT ability that can capture context in a sentence bidirectionally.

C. Traditional Machine Learning

Traditional machine learning refers to algorithms that rely on manually engineered features rather than automatic feature extraction from raw data. The examples for the traditional model are Random Forest, Logistic Regression, Support Vector Machines (SVM), and XGBoost. These models were commonly used for text classification in the past. The reason these models are gradually being left behind is their lack of ability to capture a complex human emotion and their dependency on feature engineering [9].

Despite traditional machine learning struggling to understand the word's true meaning, especially in complex sentences and languages, as in the Indonesian comment. Their main point lies in their efficient resource usage and are interpretable. Resulting in a possibility where traditional machine learning outperforms BERT in an environment where speed and computational efficiency are prioritised over achieving the highest possible accuracy. These possibilities motivate the usage of BERT as a feature extractor, which can automatically learn contextual representations and nullify the main weakness of traditional machine learning.

D. Hybrid Model for Emotion Classification

In recent years, hybrid models have gained interest, with the ability to combine two or more models. Combining the advantages of each model to fill the shortcomings of each model. Allowing researchers to increase the efficiency of the results of one model [12]. Hybrid models produce many results in a lot of different ways, depending on certain situations. Rather a single model, reliable only on one certain condition.

This paper uses those approaches, which are a hybrid approach of deep learning and ensemble learning. Utilizing BERT as a contextual feature extractor, which can capture context in a sentence bidirectionally also comprehend the semantic and syntactic aspects of Indonesian-language music comments. Traditional machine learning serves as an effective classifier due to it is computationally efficient, interpretable and capable of providing stable prediction when the features are well engineered [11].

Previous studies have also demonstrated this hybrid model. A hybrid of BERT was used for feature extraction and integrated with traditional machine learning as a classifier. The study showed that hybrid models had well-performing results compared to single models, showing the advantage of combining deep learning and ensemble learning [11].

E. Evaluation for Emotion Classification

Studies have proved that emotions are a complex phenomenon that is difficult to represent in a simple way. There are many discrete emotion schemes that need to be evaluated from multiple perspectives [13]. It depends on the right approach and selection of the model that is contextual and flexible. To achieve a good result it requires quality and efficiency. Evaluating emotion classification models not only involves performance metrics like accuracy, precision, recall, and F1-score but also suitability for a specific domain and the amount of resources available. As shown in the previous studies [13], even though Ekman's scheme is widely used, it lacks some emotion nuances where emotions can be subtle and varied.

In addition, the interaction between feature representation and model architecture must be considered in order to get the right approach and selection [14]. Attention-based transformer models such as RoBERTa-MA and DistilBERT-MA [14] show improved performance but still face some challenges in identifying more abstract emotions [15]. Therefore, pre-processing and feature enrichment strategies may offer future improvements. In this study, we compare BERT and traditional machine learning models both individually and as a hybrid, to evaluate how contextual understanding and robust classification can complement each other in detecting emotions from Indonesian language music comments with the most effective resource usage.

III. METHODOLOGY

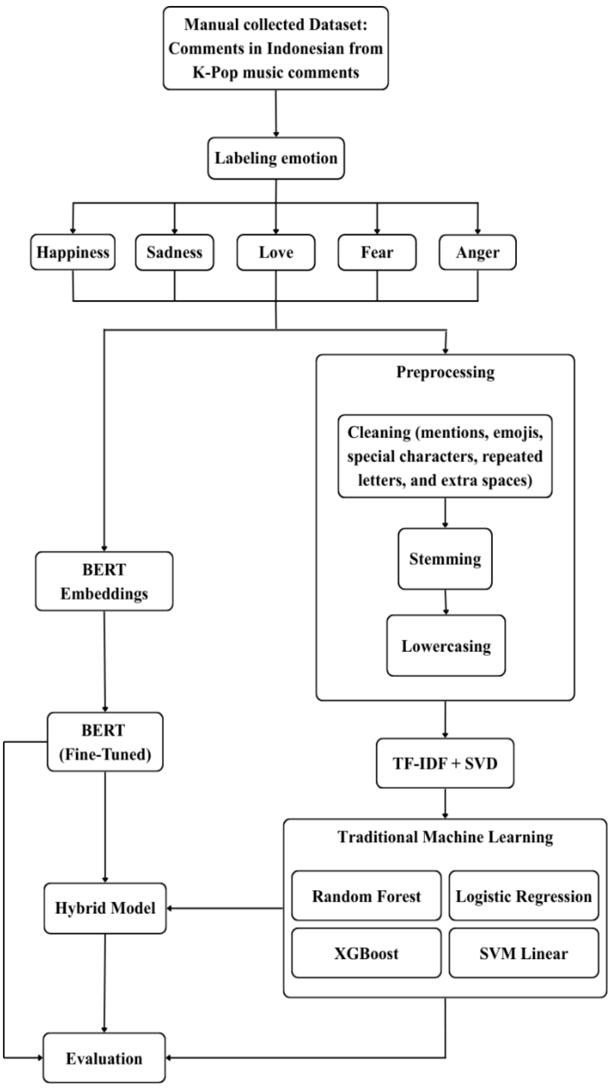

Fig. 1. Methodology Diagram

This study uses several stages to classify emotions from Indonesian comments, especially those related to Korean Music content on social media. The overall steps are summarized in the methodology diagram, as shown in Fig. 1. The methodology stages used in this study can be explained as follows:

A. Data Collection

Data was obtained from 1,000 comments on the Youtube and TikTok platforms written in Indonesian and taken manually. The comments taken are user reactions to Korean music content which is the main source of data in this study. All collected comments were stored in a structured format using Github works as the dataset for this research. This dataset is accessible through the following link https://github.com/kevinsutrisna/Emotion-Dataset/tree/main.

B. Data Labelling

After the data is collected, the process of labeling emotions on each comment will be classified manually. Each comment is classified into one of five emotion categories, namely happiness, sadness, fear, love, and anger.

C. Preprocessing

The preprocessing stage is carried out to improve the quality of data text before being used in the model training process. This process is only necessary to the traditional models since BERT performs better without removing words from the original sentence. In this stage, all comments are converted into lowercase and cleaned from irrelevant elements such as URLs, mentions, emojis, special characters, repeated letters, and extra spaces. Non-standard or slang words will be adjusted into their standard using the normalization dictionary in CSV format. Afterwards, common words which do not have a major contribution to the emotional meaning will be removed using the stopword list from *Sastrawi* library and each word is reduced to its basic form through stemming with the *Sastrawi Stemmer* library. As a result, to simplify word variations into one basic form as shown on Table I.

TABLE I EXAMPLE OF COMMENTS BEFORE AND AFTER PREPROCESSING

Before Preprocessing	After Preprocessing	Label
Lagunya enak didengar di telinga saya, seperti suasana fantasi 😊	lagu enak dengar telinga suasana fantasi	Happiness
tpi aku ttep sedih sihh part nya jin Kya gk kerasa ada gtu 🥺 🥺	aku tetap sedih part jin seperti tidak rasa begitu	Sadness
Suaranya nakutin bgt, asli takut 😱	suara takut baget asli takut	Fear
MENYALA PARA SAYANG KUU 🔥 🔥 ❤️ 🖤	nyala para sayang ku	Love
Saya sebagai orang katolik sangat tidak menyukai MV itu	saya orang katolik tidak suka mv	Anger

D. Feature Extraction

At this stage, feature extraction is performed to convert the comment text data into numeric form that can be processed by the classification algorithm model. Two approaches that are utilized, namely the Term Frequency–Inverse Document Frequency (TF-IDF) and BERT.

First, for the traditional machine learning model, the TF-IDF method works by calculating the importance weights for each word with the entire corpus. TF-IDF is applied to pre-processed comment text data that has gone through Text Cleaning, Slang Normalization, Stopword Removal, and Stemming. The resulting TF-IDF vectors were then simplified by reducing their dimensionality through Singular Value Decomposition (SVD). These simplified data were used as input for the traditional machine learning model during the training process.

Second, for the transformer-based approach, the IndoBERT model as a feature extraction tool. The preprocessed text data is used as input to the BERT tokenizer, then a token-specific representation [CLS] vector is extracted that reflects the overall meaning of the sentence. This vector is used in two ways: as direct input to the BERT fine-tuning model for classification, and second as a feature concatenated with the TF-IDF reduction to form the hybrid features. This combination provides comprehensive representation for the classification process.

E. Modeling

The modeling stage is carried out to classify emotions from each comment text data that has gone through the preprocessing and feature extraction stages. In this study, we used three different modeling approaches: a traditional machine learning model, a fine-tuned BERT model, and a hybrid model that combines features from BERT and traditional machine learning. All models were run on Google Colaboratory using the T4 GPU runtime to ensure a consistent computational environment. All models were trained using an 80% training and 20% testing stratified split.

1. Traditional Machine Learning Model

In this approach, several machine learning algorithms were tested to find out the most effective model for emotion classification using features extracted from TF-IDF. The models that were tested are Random Forest, Linear SVM, Logistic Regression, and XGBoost. Each model was trained with TF-IDF vectors from preprocessed text data and SVD to reduce dimensionality.

2. BERT Model Fine-Tuning

In this second approach, we use the IndoBERT model which is a variant of BERT that is specially trained using Indonesian. This model has been refined using a dataset that has been tokenized and converted into tensor form. Tensors themselves are multi-dimensional data structures that represent data in numeric form, tensors are similar to arrays or matrices but tensors can have more than two dimensions. The training process is carried out using Adam Optimizer and Sparse Categorical Crossentropy Loss Function, for 10 epochs. This model aims to classify emotions directly from representation in text using an End-to-end deep learning technique.

3. Hybrid Model: BERT + Traditional Machine Learning

In this approach, we combine the feature representations obtained from the BERT model with the traditional machine learning. The feature that is being used is the TF-IDF reduction results concatenated with the BERT embeddings. These hybrid features were then used as inputs for all the traditional machine learning that were being tested (Random Forest, Linear SVM, Logistic Regression, and XGBoost). From this approach, the hybrid model is expected to have good balance between how well it can predict things and how quickly it can run, sitting between simpler traditional models and the more complex, resource-heavy BERT model that has been fine-tuned.

F. Model Evaluation

An evaluation of all classification models is conducted after the modeling is completed. This evaluation approach is carried out in order to measure performance and effectiveness in detecting emotions from Indonesian-language music comments. Several metrics were used in this classification task, which are:

1. Accuracy

Accuracy measures how well the model correctly predicts emotions from the data text of comments. This metric shows how well the model recognizes the emotions from the text comment.

2. Classification Report

The Classification report provides detailed information such as precision, recall, and F1-score for each emotion label. From the classification report result, it tells how well the model is able to distinguish between each emotion.

3. Computation Time Used

Computation time is used to find out how fast the model is trained and used for prediction tasks. Time is calculated from the start of the training process to the end of training.

4. Memory Usage Tracking

Memory usage tracking is used to find out how much memory is consumed during the training process. This is used to evaluate the efficiency of the three model approaches.

5. Predict Time

Predict time is the amount of time for a trained model to make predictions on new data. This measurement shows how long it takes from the moment the model gets an input text until it gives the correct emotion label.

With these five metrics, it is possible to compare traditional machine learning with TF-IDF, BERT with fine-tuning, and the hybrid BERT + traditional machine learning approach. The evaluation results are used to assess the performance and balance between accuracy and efficiency of all approaches.

IV. Results and Discussions

This section presents the experimental results of the singular model and the hybrid model that have been tested for the task of detecting emotions in Indonesian music commentary.

TABLE II COMPARISON TABLE OF MODELS EVALUATION

Model	**Metric**			
	Accuracy	***Training Time***	***Memory Used (MiB)***	***Predicting Time***
Random Forest	0.59	0.67 s	1612.79	0.046 s
Logistic Regression	0.64	0.35 s	1531.93	0.011 s
SVM (Linear)	**0.66**	**0.29 s**	**1504.68**	**0.034 s**
XGBoost	0.56	1.16 s	1705.34	0.055 s
BERT	**0.79**	**335.95 s**	**5868.97**	**1.12 s**
Random Forest + BERT	0.67	4.43 s	4680.54	0.658 s
Logistic Regression + BERT	**0.77**	**4.10 s**	**4682.55**	**0.455 s**
SVM (Linear) + BERT	0.73	0.57 s	4682.81	0.461 s
XGBoost + BERT	0.59	23.98 s	4873.65	0.541 s

Based on the Comparison Table of Models Evaluation (Table II), it was found that the Singular BERT Model achieved the highest overall accuracy of 0.79. This confirms the Transformer-based Deep Learning model's dominance in complex natural language processing tasks such as Indonesian music commentary. Meanwhile, the best hybrid model was Logistic Regression + BERT, with an accuracy of 0.77, it demonstrates great potential by combining BERT's rich features with a simple classification layer. Meanwhile, the best singular machine learning model was the SVM (Linear) with an accuracy of 0.66.

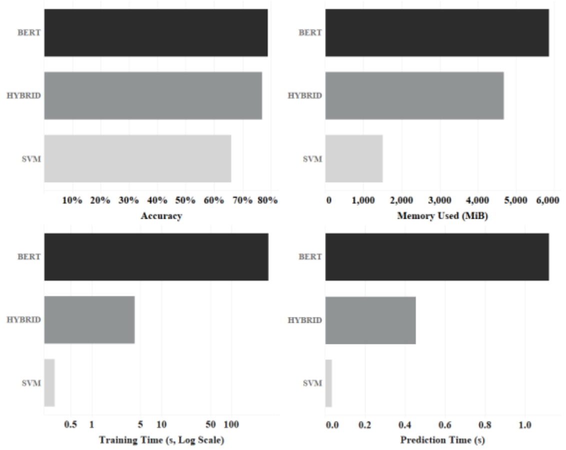

Fig. 2. Bar chart comparison of the three models.

It is worth noting from Fig. 2. that there is a significant trade-off between model performance and computational cost. While BERT provides the highest accuracy, its training time is significantly longer than other models, at 335.95 seconds, this is significantly longer and requires 5868.97 MiB of memory, which is significantly larger than a single ML model like SVM, with its training time of only 0.29 seconds and memory of only 1504.68 MiB. A hybrid model like Logistic Regression + BERT offers a compromise, achieving high accuracy (0.77) and a significantly shorter training time (4.10 seconds) than the pure BERT model, as only the BERT feature extraction part needs to be run during training. Here are the three models with the highest accuracy from each category which will be discussed further in more detail: BERT (Deep Learning), Logistic Regression + BERT (Hybrid), and Linear SVM (traditional machine learning).

TABLE III BERT CLASSIFICATION

	Precision	Recall	F1-score	Support
Happiness	0.88	0.70	0.78	40
Sadness	0.85	0.55	0.67	40
Fear	1.00	0.95	0.97	40
Love	0.58	0.90	0.71	40
Anger	0.81	0.85	0.83	40
Accuracy			0.79	200
Macro AVG	0.82	0.79	0.79	200
Weighted AVG	0.82	0.79	0.79	200

TABLE IV LOGISTIC REGRESSION + BERT CLASSIFICATION

	Precision	Recall	F1-score	Support
Happiness	0.70	0.70	0.70	40
Sadness	0.79	0.55	0.65	40
Fear	0.95	0.90	0.92	40
Love	0.64	0.80	0.71	40
Anger	0.82	0.90	0.86	40
Accuracy			0.77	200
Macro AVG	0.78	0.77	0.77	200
Weighted AVG	0.78	0.77	0.77	200

TABLE V SVM (LINEAR) CLASSIFICATION

	Precision	Recall	F1-score	Support
Happiness	0.91	0.50	0.65	40
Sadness	0.58	0.55	0.56	40
Fear	1.00	0.70	0.82	40
Love	0.57	0.60	0.59	40
Anger	0.54	0.95	0.69	40
Accuracy			0.66	200
Macro AVG	0.72	0.66	0.66	200
Weighted AVG	0.72	0.66	0.66	200

The following is a discussion of the model performance per class based on the classification report results contained in Table III (BERT), Table IV (Logistic Regression + BERT), and Table V (SVM (Linear)):

1. BERT Model: Table III shows that the BERT model shows superior performance, especially in the 'Fear' Class, supported by perfect precision (1.00) and very high Recall (0.95), resulting in an F1 score of 0.97. This demonstrates the BERT model's outstanding ability to identify emotional patterns in this specific class. However, the model's main weakness is its low precision in the 'Love' Class, which is 0.58, indicating its inability to classify other emotions into the 'Love' Class.
2. Hybrid Logistic Regression + BERT Model: Table IV shows that this hybrid model successfully overcomes some of BERT's vulnerabilities. It achieves very high F1 scores on both the 'Fear' Class (0.92) and the 'Anger' Class (0.86). The accuracy of this hybrid model is only slightly lower than that of the single BERT model (Δ = 0.02), indicating that the features generated by BERT are highly representative, and the simple classification performed by the Logistic Regression model can suffice for this task. Therefore, this combined model can result

in significantly lower computational costs compared to running the entire fine-tuned single BERT model.

3. SVM Model (Linear): From Table V, it shows that the best single machine learning model, namely SVM, can achieve the highest F1 score on the 'Fear' Class (0.82), with perfect Precision (1.00). This shows that the TF-IDF feature is able to create a clear linear separation for these specific emotions. Unfortunately, this model has a poor Recall on the 'Happiness' Class which is only 0.50 and the lowest F1 score on the 'Love' Class which is only 0.59. This proves that the TF-IDF representation is not adequate to effectively distinguish these emotions.

V. CONCLUSION

Indonesia's comments usually consisted of multiple languages, trendy slang and sarcasm. A model that is cost-efficient and time-efficient is essential for a real-time capture of people's emotions. This paper compares three possible approaches models to classify emotion in the Indonesian language. The first approach uses several traditional machine learning techniques to extract features and classify, the result was the lowest accuracy with decent training and cost time. The second way uses BERT as its feature extractor and classifier, resulting in the highest accuracy but comes with a heavy time cost and computational cost. The third uses a hybrid approach, combining several traditional machine learning as a feature extractor and BERT as a classifier, keeping the strength of SVM with its low cost and combining it with BERT's strength of the deep understanding and robustness of the language. Among all hybrids, the Logistic Regression + BERT model achieved the best performance overall.

There are cases where one approach was better than the other one. The hybrid model exchanges its accuracy for cheap and fast but reliable prediction. Capturing a huge amount of data and processing it efficiently, creating a model that is able to monitor high streaming data and fast changing trends by being trained daily. Whereas BERT excels in accuracy and is suitable for high stake emotion classification when the slight accuracy difference matters. The hybrid approach is ideal for capturing large amounts of streaming data, monitoring fast changing trends, and providing robust emotion classification in a cost effective way. Logistic Regression with hybrid is therefore suitable for practical applications where both speed and accuracy are important.

In future works, implementing models on real time streaming data (Kafka or similar frameworks) is essential to assess latency, throughput, and maintenance cost under live and high-volume conditions. At the same time, expanding the dataset beyond the current 1,000 comment samples will help to avoid overfitting, improve model generalization and better capture the complexity of Indonesian social media comments. In addition, a new continuous learning strategy needs to be developed to keep up with the fast changing trends, slang and memes.

ACKNOWLEDGMENT

The authors would like to thank Bina Nusantara University for funding the publication of this paper.

AUTHOR CONTRIBUTIONS

Kevin (First Author): Focused on the implementation of the BERT, traditional machine learning, and hybrid model, conducted the comparative analysis, and wrote the introduction and conclusion.

Edrick Louis (Second Author): Contributed to the methodology, developed the discussion section by interpreting and analyzing the results.

David Malcom Maximillian (Third Author): Responsible for data collection and dictionary, conducted the literature review at the early stage of study and wrote the abstract, contributed to the critical review and final editing of the paper.

Kartika Purwandari (Supervisor / Fourth Author): Provided supervision, conceptual guidance, and validation throughout the research, ensured academic quality, and gave final approval of the manuscript.

REFERENCES

[1] D. Susilo and Teguh Dwi Putranto, “Indonesian youth on social media: Study on content analysis,” Atlantis Press, pp. 94–97. doi: https://doi.org/10.2991/sshr-17.2018.15.

[2] Juslin, Patrik N, Musical emotions explained: Unlocking the secrets of musical affect. Oxford University Press, 2019.

[3] F. A. Acheampong, H. Nunoo-Mensah, and W. Chen, “Transformer models for text-based emotion detection: a review of BERT-based approaches,” Artificial Intelligence Review, vol. 54, Art. no. 8, 2021.

[4] E. M. Younis, S. M. Zaki, E. Kanjo, and Houssein, Essam H, “Evaluating ensemble learning methods for multi-modal emotion recognition using sensor data fusion,” Sensors, vol. 22, Art. no. 15, 2022.

[5] D. Purwitasari, A. Adi, A. P. Kurniawan, and K. A. Nur, “A hybrid method on emotion detection for indonesian tweets of COVID-19,” Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), vol. 7, Art. no. 2, 2023.

[6] F. A. Acheampong, C. Wenyu, and H. Nunoo-Mensah, “Text-based emotion detection: Advances, challenges, and opportunities,” Engineering Reports, vol. 2, Art. no. 7, 2020.

[7] A. Maruf, F. Khanam, M. M. Haque, J. Z. Masud, Mridha, Muhammad Firoz, and Z. Aung, “Challenges and opportunities of text-based emotion detection: a survey,” IEEE access, vol. 12, pp. 18416–18450, 2024.

[8] A. Chiorrini, C. Diamantini, A. Mircoli, and D. Potena, “Emotion and sentiment analysis of tweets using BERT.,” 2021, pp. 1–7.[2]

[9] S. González-Carvajal and E. C. Garrido-Merchán, “Comparing BERT against traditional machine learning text classification,” 2020.

[10] A. S. Talaat, “Sentiment analysis classification system using hybrid BERT models,” Journal of Big Data, vol. 10, Art. no. 1, 2023.

[11] C. N. Dang, Moreno-García, María N, and Fernando, “Hybrid deep learning models for sentiment analysis,” Complexity, vol. 2021, Art. no. 1, 2021.

[12] F. Zhang, J. Chen, Q. Tang, and Y. Tian, “Evaluation of emotion classification schemes in social media text: an annotation-based approach,” BMC psychology, vol. 12, Art. no. 1, 2024.

[13] G. Pinto, J. M. Carvalho, F. Barros, S. C. Soares, A. J. Pinho, and S. Brás, “Multimodal emotion evaluation: A physiological model for cost-effective emotion classification,” Sensors, vol. 20, Art. no. 12, 2020.

[14] I. Ameer, N. Bölücü, Siddiqui, B. Can, G. Sidorov, and A. Gelbukh, “Multi-label emotion classification in texts using transfer learning,” Expert Systems with Applications, vol. 213, p. 118534, 2023.

[15] Y. Du, R. G. Crespo, and O. S. Martínez, “Human emotion recognition for enhanced performance evaluation in e-learning,” Progress in Artificial Intelligence, vol. 12, Art. no. 2, 2023