Bulletin of Electrical Engineering and Informatics
Vol. 14, No. 5, October 2025, pp. 3632~3644
ISSN: 2302-9285, DOI: 10.11591/eei.v14i5.9731

MAS-TENER: a modified attention score transformer encoder for Indonesian skill entity recognition

Meilany Nonsi Tentua[1], Suprapto[2], Afiahayati[2]
[1]Department of Informatics, Faculty of Science and Technology, Universitas PGRI Yogyakarta, Yogyakarta, Indonesia
[2]Department of Computer Science and Electronics, Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada, Yogyakarta, Indonesia

Article Info

Article history:

Received Dec 16, 2024
Revised Jul 14, 2025
Accepted Jul 29, 2025

Keywords:

Information retrieval
Named entity recognition
Natural language processing
Pretrained model language
Skill recognition

ABSTRACT

Skill entity recognition is a crucial task for aligning educational curricula with the evolving needs of the industry, particularly in multilingual job markets. This study introduces modified attention score transformer encoder (MAS-TENER), a novel transformer-based model designed to enhance the recognition of skill entities from Indonesian job descriptions. The proposed model modifies the attention mechanism by integrating relative positional embeddings and removing the scaling factor in self-attention. These improvements enhance the context of tokens, allowing for the accurate establishment of hard skills, soft skills, and technology skills. The MAS-TENER model was pre-trained and fine-tuned using a combinF.ation of job description datasets and additional corpora, achieving an F1-score of 90.46% at the entity level. The experimental results demonstrate the model's ability to handle unstructured, mixed-language job descriptions, with significant potential for curriculum reform and the development of new workforce capabilities. The study demonstrates the efficacy of the MAS-TENER model as an effective response for any natural language processing (NLP) task in low-resource languages. Moreover, the scope of long-term job market analytics in action research has been a key skill set in the education policy arena, demonstrating collaborative workforce capabilities.

Corresponding Author:

Suprapto
Department of Computer Science and Electronics, Faculty of Mathematics and Natural Sciences
Universitas Gadjah Mada
Sekip Utara, Bulak Sumur, Yogyakarta 55281, Indonesia
Email: sprapto@ugm.ac.id

1. INTRODUCTION

The higher education sector plays a vital role as a primary supplier of the workforce to engineering and economic industries [1]. However, rapid industrial developments have created challenges for educational institutions, as they are tasked with producing graduates equipped with up-to-date skills [2], [3]. The alignment of university curricula with industry needs is crucial; higher education institutions should engage more closely with stakeholders, as there is a significant gap between education and the labour market in Indonesia. This gap contributes to "intellectual unemployment," where graduates cannot find jobs that match their qualifications [4].

To bridge this gap, higher education institutions must focus on producing human resources with competencies and skills that meet industry demands [5]. One possible approach is to extract relevant job skills from job postings to inform curriculum changes. Online job platforms in Indonesia that provide information on industry-required skills include Indeed [6] and JobStreet.id [7]. These job postings, often containing unstructured job descriptions, present an opportunity for natural language processing (NLP) techniques to extract and analyze essential skill requirements [8]. To illustrate this process, Figure 1 provides an example of skill recognition in an Indonesian job description, highlighting how specific skills (hard skills, soft skills, technology) are identified within the text, which underscores the need for effective entity recognition models.

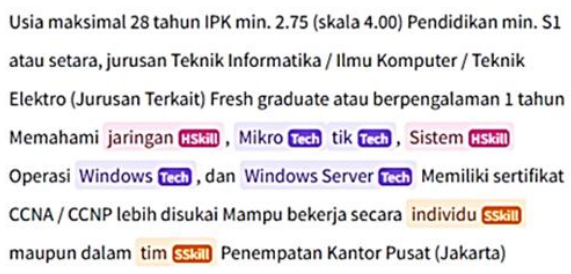


Figure 1. Sample of skill recognition in Indonesian job description

Information extraction, a subfield of NLP, can identify and recognize entities such as job skills within these unstructured job descriptions. Named entity recognition (NER) is a fundamental task in this process, responsible for identifying entities within text [9]. In recent years, transformer-based models have demonstrated significant success in NER tasks, primarily due to their self-attention mechanism, which effectively captures complex word dependencies within a sentence [10].

Transformer-based models, such as pretrained language models (PLMs) and bidirectional encoder representations from transformers (BERT) [11], have become the standard for NER tasks due to their superior contextual understanding. Fine-tuning PML BERT offers distinct advantages, achieving high accuracy with smaller training datasets, thereby reducing the time and costs required to develop specialized models for job skill extraction. The identification of job skill requirements on job portals using fine-tuned, language-specific BERT-based PML has been conducted by various researchers. This PML includes English [12], [13], as well as Vietnamese [14], Chinese [15], German [16], and Finnish [17].

Existing pre-trained models for Indonesian, such as Indonesian bidirectional encoder representations from transformers (IndoBERT) [18] and Indonesian BERT for tweets (IndoBERTweet) [19], have demonstrated adaptability to low-resource languages; however, they lack specific training on job description datasets, resulting in suboptimal performance in identifying skill-related entities. Moreover, the reliance of standard transformer architectures on absolute positional embeddings limits their ability to capture directional and distance-based contextual relationships, which are crucial for recognizing skills [20]-[22]. Additionally, the implementation of relative position embedding in attention can also be applied to BERT [23].

To overcome these challenges, this study introduces the modified attention score transformer encoder (MAS-TENER), a novel transformer-based model designed for skill entity recognition in Indonesian job descriptions. MAS-TENER addresses the limitations of conventional models by incorporating relative positional embeddings and removing the scaling factor in the self-attention mechanism. These innovations enhance the model's ability to capture contextual dependencies between tokens, improving the precision and recall of skill recognition.

The MAS-TENER model was pre-trained and fine-tuned on specialized datasets, including annotated Indonesian job descriptions and supplementary corpora from Wikipedia, enabling it to handle mixed-language content and unstructured data effectively. Experimental results demonstrate that MAS-TENER achieves a state-of-the-art F1-score of 90.46% at the entity level, outperforming existing models such as BERT and IndoBERT in precision and recall.

This research makes a significant contribution to the field in three key ways. First, it provides a modified attention mechanism specifically designed for technologies that recognize skills in low-resource, multilingual contexts. Second, it establishes a new benchmark for skill entity recognition, contributing to both new developments in workforce development and curriculum revision. Third, MAS-TENER's robust performance in extracting skills from job descriptions positions it as a valuable tool for bridging the gap between academia and industry, particularly in Indonesia and similar contexts.

2. METHOD

The MAS-TENER model was developed to enhance skill entity recognition from Indonesian job descriptions by modifying the attention mechanism in transformer encoders. This section outlines the data collection, model architecture, pre-training, and fine-tuning processes involved in building MAS-TENER. To provide an overview of our approach, Figure 2 illustrates the proposed methodology, depicting the workflow from data collection and preprocessing to model training and evaluation, highlighting the integration of relative position embeddings in the transformer architecture.

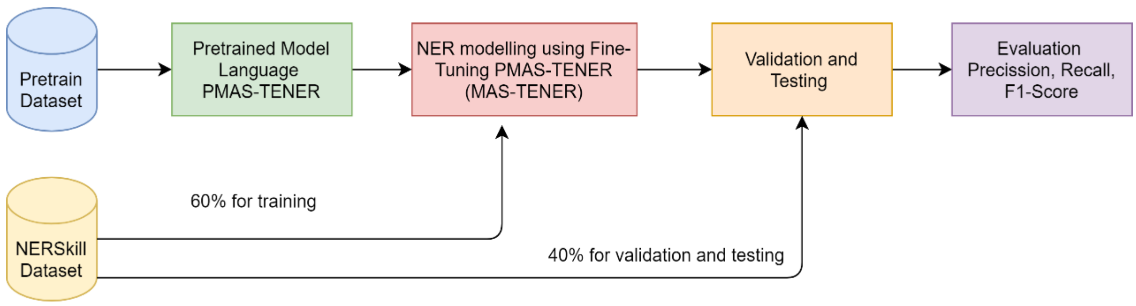


Figure 2. The proposed methods

2.1. Data collection and preparation

Job description data were collected from several Indonesian job portals, including Loker.id [24], JobStreet [7], Indeed [6], and Job.id [25] to build the skill recognition model. These job postings include information such as company name, job location, job field, and job descriptions, along with the required skills. To illustrate the data preparation process, Figure 3 outlines the steps involved in data collection and preprocessing for building MAS-TENER, including filtering job descriptions, removing special characters, and splitting the dataset into pre-training and fine-tuning subsets.

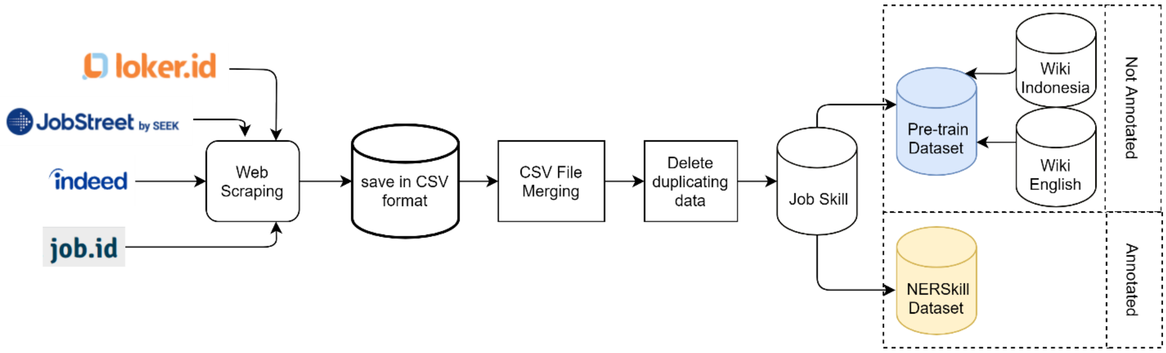


Figure 3. Data collection and preparation

The data was filtered to include job descriptions containing at least five words. Special characters and HTML tags were removed to simplify the text. A total of 34,966 job descriptions were collected and stored in CSV format. The dataset was divided into two parts: i) pre-train dataset, consisting of 30,354 job descriptions for language model pre-training and ii) NER-Skill dataset: consisting of 4,394 job descriptions for fine-tuning the NER model [26].

Due to the limited size of the pre-training dataset, additional data from Indonesian Wikipedia [18] and English Wikipedia [11] were added to enhance the model's ability to handle both Indonesian and English vocabulary, as job postings in Indonesia often include mixed-language content.

2.2. Dataset annotation

The NER-Skill [26] dataset was annotated manually following the BIO tagging format (beginning, inside, and other) [27]. The following labels were used for annotation, as seen in Table 1.

Table 1. The labels in the NER-Skill dataset

Label	Description
B-HSkill and I-HSkill	Hard skills, such as programming and data analysis.
B-SSkill and I-SSkill	Soft skills, such as teamwork and communication.
B-Tech and I-Tech	Technology-specific skills, such as Python and MySQL.
O	Other tokens that do not belong to any skill category.

The NER-Skill dataset annotation format follows the CoNLL 2003 format, and each job description was tokenized prior to the annotation process. The dataset's imbalance was managed during training by appropriately weighting underrepresented labels. To illustrate the annotation structure, Figure 4 presents a sample of the NER-Skill dataset, showing tokenized job description text with corresponding BIO tags to demonstrate how skill entities are labelled for model training.

```
PHP         B-Tech
,           O
JavaScript  B-Tech
,           O
Jquery      B-Tech
,           O
HTML/XHTML  B-Tech
,           O
CSS         B-Tech
,           O
MySQL       B-Tech
Mampu       O
bekerja     O
secara      O
individu    B-SSkill
dan         O
Tim         B-SSkill
Mampu       O
bekerja     O
dibawah     B-SSkill
tekanan     I-SSkill
```

Figure 4. A sample of the NER-Skill dataset with BIO tagging

2.3. Modified attention score transformer encoder

The core of MAS-TENER is a modified transformer encoder designed to enhance skill entity recognition by addressing limitations in the conventional self-attention mechanism. Specifically, the model incorporates relative position embedding and eliminates the scaling factor to improve the attention mechanism.

The model utilizes sinusoidal functions to represent positional embeddings, thereby extracting the positional relationships among tokens within a sentence. This function allows the model to comprehend the direction and distance among skill entities, which is essential for accurate recognition. As in (1) and (2) are used to position embedding.

$$p_{(pos,2i)} = sin\left(\frac{pos}{10000^{\frac{2i}{d}}}\right) \tag{1}$$

$$p_{(pos,2i+1)} = cos\left(\frac{pos}{10000^{\frac{2i}{d}}}\right) \tag{2}$$

where: p is the position encoding, pos is the word position, i is the feature embedding i, and d is the word embedding size.

Instead of the standard attention score calculation, MAS-TENER computes attention scores using relative positional embeddings without the scaling factor. This approach enables the model to focus more on the relationships between tokens than their distances. To illustrate this modification, Figure 5 depicts the architecture of the modified attention score mechanism in the transformer encoder, showing how relative positional embeddings are integrated into the query, key, and value computations to enhance skill entity recognition.

This modified attention mechanism, as illustrated in Figure 5, enables MAS-TENER to more effectively capture token relationships, resulting in enhanced skill recognition performance. To extract feature representations, namely Q (query), K (key), and V (value), we project the input sequence using three matrices: $W^Q \in R^{m\times d_q}, W^K \in R^{m\times d_k}$, and $W^V \in R^{m\times d_v}$. Here, d_k and d_q are set to be equal. These matrices are used in (3) to calculate the Q, K, and V matrices.

$$Q = x_i W^Q, K = x_i W^K, V = x_i W^V \tag{3}$$

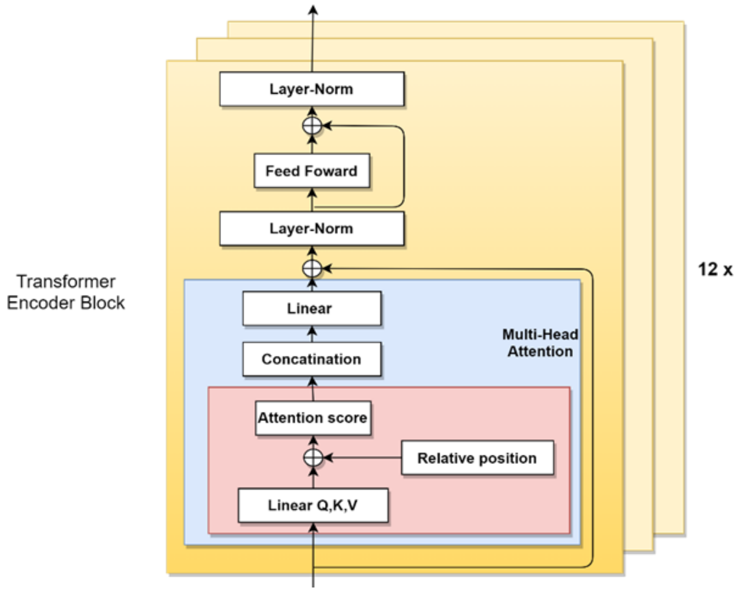


Figure 5. The modified attention score in encoder transformer

The input sequence will produce an attention score $z = (z_1, z_2, ..., z_n)$ with the same length as the input $z_i \in R^{d_z}$. The attention score is computed by considering both the dot product of the query and key vectors, as well as the relative position information. The relative position information is encoded using relative position embeddings, which capture the positional relationships between tokens in the sequence. Relative positional embedding is a vector $a_{ij}^Q, a_{ij}^K, a_{ij}^V \in R^{d_a}$, where a_{ij}^Q, a_{ij}^K, and a_{ij}^V are the input elements x_i and x_j direction information, respectively. The attention scores are then typically passed through a softmax function to obtain attention weights, which determine the importance or relevance of each token in relation to the query. The attention score between two tokens is calculated using (4) and (5).

$$\alpha_{ij} = \frac{exp\ e_{ij}}{\sum_{k=1}^{n} exp\ e_{ik}} \tag{4}$$

$$z_i = \sum_{j=1}^{n} \alpha_{IJ} (V + a_{ij}^V) \tag{5}$$

where: $e_{ij} = (Q)(K)^T + (Q)\ a_{ij} + (K)^T a_{ij}^T$, x_i and x_j are input vectors, a_{ij} is the relative position embedding, and z_i is the attention score.

By incorporating relative position information into the self-attention mechanism, MAS-TENER can better capture dependencies between tokens, improving the accuracy of skill recognition.

2.4. Pre-training of language model

The implementation of embedding the relative position in the attention score is applied to the BERT architecture. We named PMAS-TENER for the pre-trained model language, which was pre-trained using the combined pre-train dataset (job descriptions and Wikipedia corpus). The PMAS-TENER model is trained using the masked language model (MLM) technique, a powerful approach for pre-training language models. In MLM, specific tokens in a sentence are randomly masked, and the model predicts these masked tokens based on the surrounding context [28].

This pre-training process involved the following settings: epochs=300, learning rate=1e-4, batch size=64, gradient accumulation steps=4, and weight decay=10. The pre-training was conducted using the NVIDIA A100-SXM machine with stochastic gradient descent (SGD) as the optimization method. This process allowed the model to learn language patterns and semantic relationships within the job descriptions. To demonstrate the pre-training approach, Figure 6 illustrates the MLM technique used in PMAS-TENER, showing how tokens are randomly masked and predicted based on the surrounding context to enhance the model's understanding of job description text.

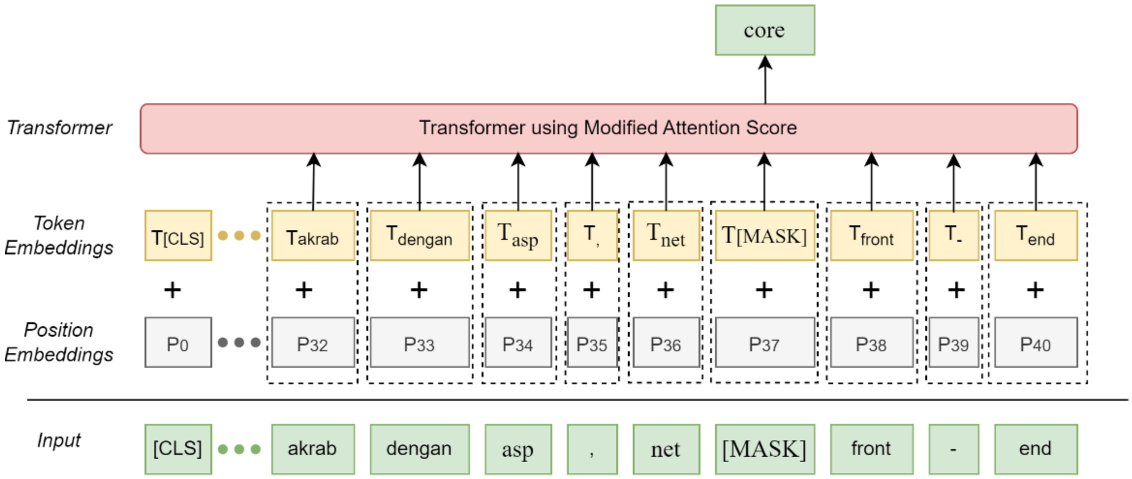


Figure 6. MLM technique in PMAS-TENER

2.5. Fine-tuning for named entity recognition-skill recognition

After pre-training, the model was fine-tuned on the NER-Skill dataset, which contains manually labelled job descriptions. The fine-tuning process adjusted the pre-trained model to recognize skill entities specific to job descriptions using the following hyperparameters: epochs=6, learning rate=3e-5, batch size=32, dropout=0.2, optimizer=Adam, and sequence length=256 tokens. To illustrate this process, Figure 7 depicts the fine-tuning architecture for NER-skill recognition using the pre-trained PMAS-TENER model, showing how the model processes input embeddings and predicts BIO tags for skill entities in job descriptions.

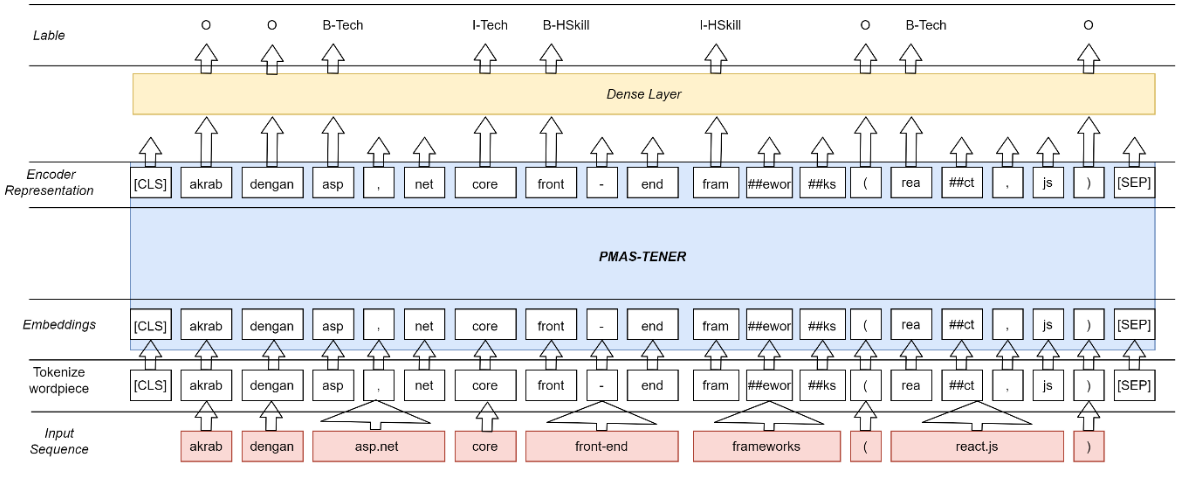


Figure 7. Fine-tuning for NER-skill recognition using pre-trained PMAS-TENER

In Figure 7, *E*'s and *R*'s represent the input embedding and the contextual representation of sub-token s, and [CLS] is the special symbol used to get the full input representation. These words are then tokenized into subword units using a WordPiece tokenizer to ensure compatibility with the transformer model. The embedding layer takes these tokens and translates them into dense vector embeddings that include both positional information and token-specific information. For instance, within the encoder of PMAS-TENER, the input embeddings are processed using a variant of self-attention with relative positional embeddings, which empowers the model to understand the contextual relationships between tokens. The encoder representation outputs refined, contextualized features for each token, which are then passed to the dense layer. This layer performs a final classification, assigning labels to each token.

2.6. Model evaluation

The performance of the MAS-TENER model was evaluated using the NER-Skill test dataset. The model's performance is measured using the precision, recall, and F1-score. These values are calculated using a confusion matrix, which identifies the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) predictions. The precision, recall, and F1-score can be calculated using the formulas in (7) to (9) [29]:

$$precision = \frac{TP}{TP+FP} \tag{7}$$

$$recall = \frac{TP}{TP+FN} \tag{8}$$

$$F1\ score = \frac{2 \times precision \times recall}{precision + recall} \tag{9}$$

These metrics were calculated at both the token and entity levels [30], allowing a detailed assessment of the model's ability to recognize skill entities within job descriptions.

A baseline comparison was conducted using the standard attention mechanism without incorporating relative position information (Model 1), the attention mechanism incorporating relative position information on matrix *K* with scaling (EBERT-RP) [31], and the attention mechanism incorporating relative position information on matrix *K* and *Q* with scaling (Model 2).

3. RESULTS

The training procedure for the four PMLs: i) Model 1 (standard BERT without relative position embeddings), ii) EBERT-RP (BERT with relative position embeddings on matrix *K* and scaling), iii) Model 2 (BERT with relative position embeddings on matrix *K* and *Q* with scaling), and iv) MAS-TENER, consists of a 12-layer, 12-head encoder transformer and the maximum token length at 256 tokens. To further analyze the training behavior of MAS-TENER and the baseline models, we examined the loss curves over iterations during the fine-tuning phase. Figure 8 plots the loss function for four different models: Model 1 (blue), EBERT-RP (red), Model 2 (yellow), and MAS-TENER (green).

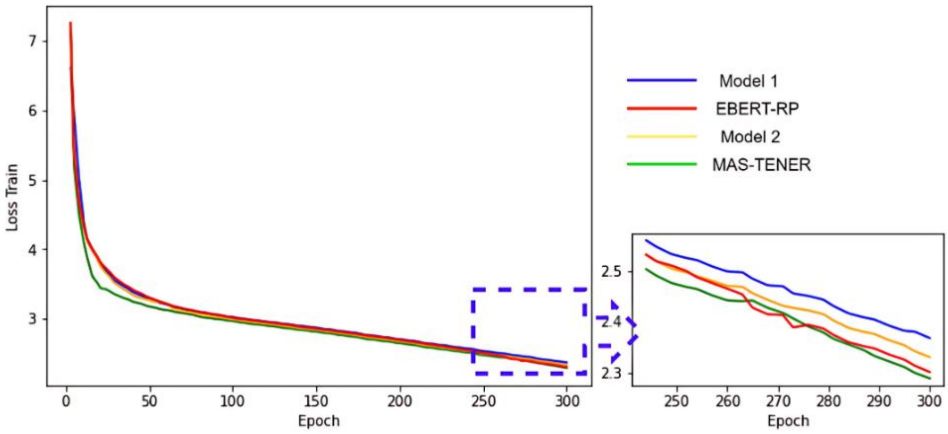


Figure 8. Loss function over iterations for Model 1 (blue), EBERT-RP (red), Model 2 (yellow), and MAS-TENER (green), the inset provides a zoomed-in view of the later iterations

As shown in Figure 8, the overall convergence trend, while the inset provides a zoomed-in view of the loss values in the later iterations, highlighting the stability and final convergence of each model. This comparison demonstrates the effectiveness of MAS-TENER's modified attention mechanism in achieving faster and more stable convergence compared to the baselines.

The pre-training configurations of the evaluated language models demonstrate substantial complexity, with Model 1 comprising 110,591,667 parameters, while EBERT-RP, Model 2, and PMAS-TENER each incorporate 111,377,331 parameters. These extensive parameter sets were trained over computational durations of 109 hours for Model 1, 125 hours and 46 minutes for EBERT-RP, 126 hours and 57 minutes for Model 2, and 120 hours and 32 minutes for PMAS-TENER.

After developing the PMAS-TENER model through pre-training, a fine-tuning process was conducted to adapt the model for skill entity recognition specific to job descriptions. The process involved fine-tuning the pre-trained model to classify tokens accurately, which included hard skills, soft skills, and technology-specific competencies. Fine-tuning was performed using hyperparameters carefully selected to optimize performance. The model was trained for six epochs with a learning rate of 3e-5 and a batch size of 32, aiming to balance the computational component with the need for generalization. We applied a dropout of 0.2 to help prevent overfitting and used the Adam optimizer, as it is generally more efficient with sparse gradients and has effective dynamic learning rates. Additionally, a sequence length of 256 tokens ensured the model could process job descriptions of varying lengths without truncating important contextual information. This structured and methodical fine-tuning process enhanced PMAS-TENER's ability to recognize skill entities with high precision and recall.

To evaluate the training stability and convergence of the MAS-TENER model compared to the baseline models, we analyzed the loss function over epochs during both pre-training and fine-tuning phases. Figure 9 illustrates the training and validation loss curves for four different configurations: Figure 9(a) Model 1, Figure 9(b) EBERT-RP, Figure 9(c) Model 2, and Figure 9(d) MAS-TENER. In each subfigure, the green line represents the training loss, while the red line represents the validation loss. These plots demonstrate how the modified attention mechanism in MAS-TENER affects convergence compared to baseline models.

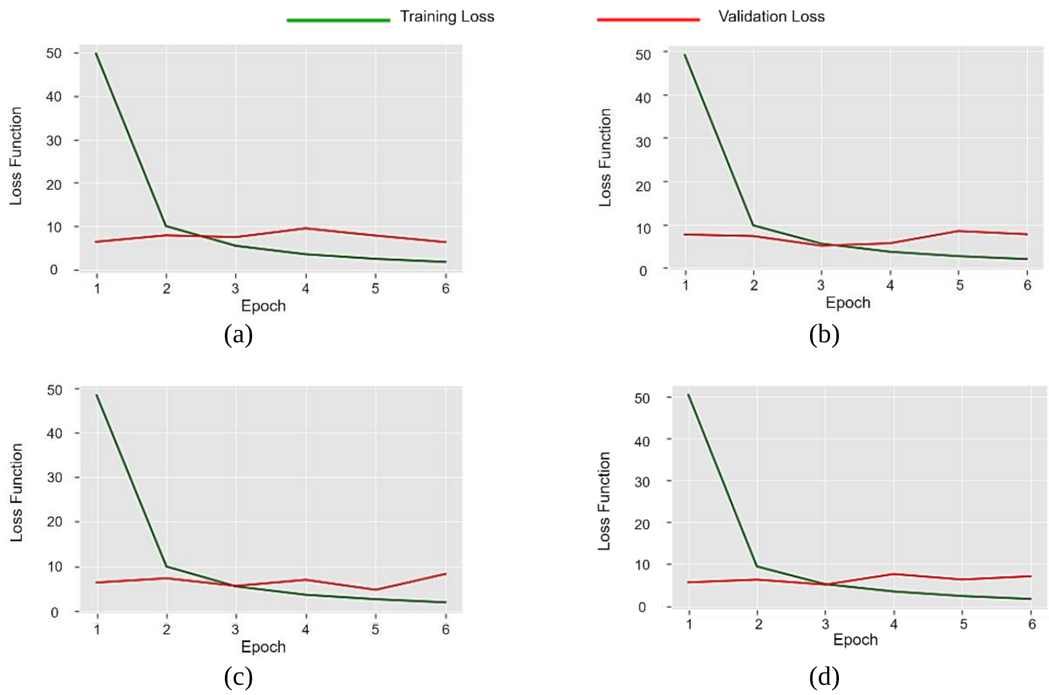


Figure 9. Comparison of NER modelling loss functions; (a) fine-tuning Model 1, (b) fine-tuning EBERT-RP, (c) fine-tuning Model 2, and (d) MAS-TENER

As shown in Figure 9, the MAS-TENER model was evaluated on the NER-Skill test dataset to assess its effectiveness in recognizing skill entities from Indonesian job descriptions. This section presents the experimental results, focusing on key performance metrics such as precision, recall, and F1-score. We also compare MAS-TENER's performance with baseline models and discuss the implications of the findings for skill entity recognition and curriculum alignment with industry needs.

3.1. Performance at the token level

To evaluate MAS-TENER's performance at the token level, we compared it with Model 1 (standard BERT without relative position embeddings), EBERT-RP, and Model 2 (BERT with relative position embeddings and scaling). Table 2 presents the precision, recall, and F1-scores for each label, highlighting MAS-TENER's improvements in recognizing hard skills, soft skills, and technological competencies. The precision and Recall for the B-Tech label achieved 89.07% and 81.50%, respectively, indicating that the model performs well in identifying technological skills from job descriptions. Similarly, the F1-score of MAS-TENER for soft skills (B-SSkill) was 81.48%, outperforming both baseline models.

Table 2. The evaluation skill recognition model at the token level

Label	Model 1			EBERT-RP [31]			Model 2			MAS-TENER		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
B-HSkill	59.26	65.81	62.36	69.46	53.24	60.28	71.29	56.76	63.20	63.69	65.56	**64.62**
B-SSkill	71.96	87.14	78.83	76.29	86.27	80.97	76.13	83.95	79.85	79.16	83.95	**81.48**
B-Tech	84.29	84.30	84.29	88.27	81.45	84.72	85.10	85.14	85.12	89.07	81.50	85.12
I-HSkill	65.21	52.28	58.03	71.98	51.22	59.85	66.52	59.62	62.88	63.18	66.88	**64.98**
I-SSkill	68.76	54.96	61.09	79.61	53.64	64.09	61.49	62.73	62.10	78.40	52.42	**62.83**
I-Tech	73.36	59.97	65.99	82.97	59.61	69.37	72.83	70.43	71.61	79.50	67.53	**73.03**
O	95.54	95.37	95.46	94.45	97.36	95.89	95.40	96.44	95.92	95.72	96.28	**96.00**
Average	74.05	71.40	72.29	80.43	68.97	73.60	75.54	**73.58**	74.38	**78.39**	73.45	**75.44**

3.2. Performance at the entity level

We also evaluated MAS-TENER's performance at the entity level to assess its ability to accurately recognize complete skill entities. Table 3 reports the precision, recall, and F1-scores for MAS-TENER compared to Model 1, EBERT-RP, and Model 2, demonstrating MAS-TENER's superior performance in entity-level skill recognition. The entity-level evaluation further demonstrates that the MAS-TENER model recognizes B-SSkill and I-HSkill entities, achieving a 97.12% F1-score for B-SSkill and a 91.20% recall for I-HSkill. Including relative position embeddings significantly enhances the model's ability to capture relationships between tokens, especially for skill-related entities, where distance and direction are essential for accurate entity classification.

Table 3. The evaluation skill recognition model at the entity level

Label	Model 1			EBERT-RP [31]			Model 2			MAS-TENER		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
B-HSkill	85.12	90.77	87..85	87.53	91.67	89.56	90.63	88.96	89.79	88.58	91.16	89.85
B-SSkill	94,72	97.34	96.01	95.41	98.27	96.82	96.54	96.81	96.67	96.42	97.82	97.12
B-Tech	92.01	92.97	92.48	93.97	94.63	94.30	93.81	93.41	93.61	94.78	91.65	93.19
I-HSkill	85.65	80.42	82.95	89.32	87.19	88.24	85.76	90.78	88.20	85.98	91.20	88.51
I-SSkill	94.38	83.00	88..32	93.49	86.48	89.85	91.32	91.19	91.26	95.23	88.87	91.94
I-Tech	80.76	69.65	74.80	88.47	76.38	81.98	82.69	78.98	80.79	84.91	79.57	82.16
Average	88.77	85.69	87.07	91.37	89.10	90.12	90.12	90.02	90.05	**90.98**	**90.05**	**90.46**

3.3. Comparison with other pretrained models language

To further assess MAS-TENER's performance, we compared it with existing pre-trained models, including BERT [11], TENER [21], and IndoBERT [18], as well as the baseline models (Model 1, EBERT-RP, and Model 2). Table 4 presents the evaluation results at both token and entity levels, reporting precision, Recall, and F1-scores for each model. This comparison highlights MAS-TENER's improvements over standard transformer models in recognizing skills from Indonesian job descriptions, particularly at the entity level.

Table 4. Comparison of MAS-TENER with pretrained and baseline models at token and entity levels

Model NER		Level token			Entity level		
		Precision	Recall	F1-score	Precision	Recall	F1-score
Baseline	From pretrained BERT [11]	**81.68**	75.03	77.70	87.90	85.90	86.30
	From pretrained IndoBERT [18]	81.65	**76.47**	**78.59**	87.40	85.70	86.50
	TENER [21]	73.24	66.18	68.94	85.55	82.52	83.86
	Model 1	74.05	71.4	72.29	88.77	85.69	87.07
	EBERT-RP [31]	80.43	68.97	73.60	**91.37**	89.1	90.12
	Model 2	75,54	73,58	74,38	90,12	90,02	90,05
Propose	MAS-TENER	78.39	73.45	75.44	90.98	**90.05**	**90.46**

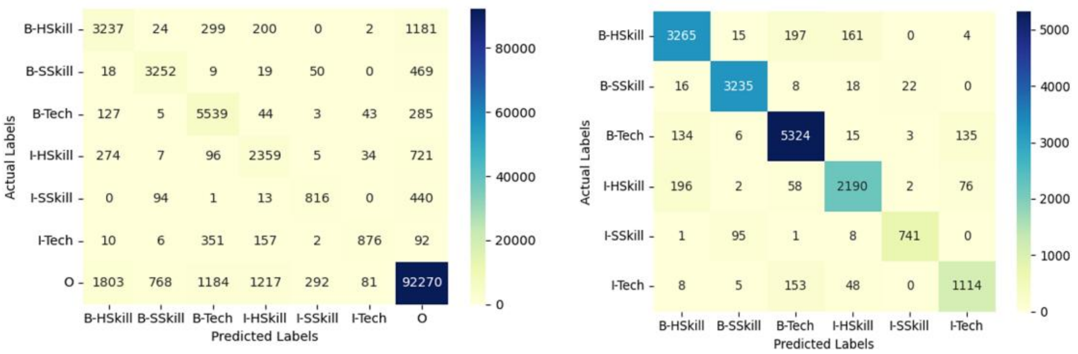
As shown in Table 4, MAS-TENER's F1-score of 90.46% at the entity level surpasses all baselines, underscoring the effectiveness of the modified attention mechanism. This performance is attributed to the modified attention score mechanism, which effectively captures the contextual relationships in job descriptions.

4. DISCUSSION

The experimental results demonstrate that the MAS-TENER model significantly improves skill entity recognition from unstructured job descriptions in Indonesian. By incorporating relative positional embeddings and eliminating the scaling factor in the attention mechanism, MAS-TENER achieves better precision, recall, and F1-score compared to both baseline and existing pre-trained models.

The improved ability to recognize hard skills and technological competencies is valuable for educational institutions and industries, as it enables better alignment of curricula with real-world job requirements. Additionally, the capability to work with mixed-language content (Indonesian and English) in job postings enhances the applicability of MAS-TENER in multilingual job markets.

However, there are still challenges in recognizing specific skills, particularly I-HSkill and I-SSkill, which often overlap with other categories such as B-Tech. This issue stems from ambiguities in job descriptions, where terms like "programming" may refer to hard skills and technologies. Figure 10 shows MAS-TENER confusion matrices Figure 10(a) token level and Figure 10(b) entity level.



(a) (b)

Figure 10. MAS-TENER confusion matrices; (a) token level and (b) entity level

The confusion matrices for MAS-TENER at the token level a) and entity level b) offer valuable insights into its classification performance across skill categories (B-HSkill, B-SSkill, B-Tech, H-HSkill, I-SSkill, I-Tech, and O). At the token level a), the model demonstrates strong performance for B-HSkill and B-SSkill, but notable errors emerge, including significant misclassifications of B-Tech as B-HSkill and I-Tech as I-HSkill or B-HSkill. At the entity level b), MAS-TENER continues to face challenges with specific skill entities, particularly misidentifying B-HSkill as I-HSkill or I-Tech, while also exhibiting confusion between B-Tech and I-Tech. Future work should address these ambiguities by improving entity disambiguation and contextual modelling.

MAS-TENER performs state-of-the-art skill entity recognition for Indonesian job descriptions, outperforming standard transformer models like BERT [11], TENER [21], and IndoBERT [18]. The model's success underscores the importance of relative positional information in NER tasks and highlights its potential for real-world applications in workforce development and curriculum reform.

5. CONCLUSION

This study introduced MAS-TENER, a modified transformer-based model designed to enhance skill entity recognition from Indonesian job descriptions by addressing limitations in conventional attention mechanisms. The primary objective was to enhance the accuracy of recognizing job-related skills, including hard skills, soft skills, and technological competencies, by integrating relative position embeddings and eliminating the scaling factor in the self-attention mechanism. The experimental results confirmed that MAS-TENER achieved a state-of-the-art F1-score of 90.46% at the entity level, outperforming existing models such as BERT, TENER and IndoBERT in precision and recall for skill recognition tasks.

These findings have significant implications for workforce development and educational alignment with industry needs. Through the precise extraction of skills from job descriptions, MAS-TENER can help educational institutions gain a better understanding of the demand for skills, enabling them to update their curricula to align with identified skills gaps in the workforce compared to what they are currently providing to students. Beyond that, MAS-TENER is useful in multilingual and mixed-language settings and is particularly applicable to diverse job markets, providing a novel service to job domains that otherwise relied on unstructured job description data.

Future work can expand on MAS-TENER's methodology for improving entity disambiguation and contextualization, such as the overlapping skill categories discussed above. This model lays the groundwork for future advancements in NLP for low-resource languages, enabling broader applications in job market analysis and educational reform. The research validates MAS-TENER's approach to skill recognition, paving the way for future developments that enhance the utility and scalability of NLP models in addressing real-world skill assessment needs.

ACKNOWLEDGMENTS

We are grateful to the Department of Computer Science and Electronics, Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada, for providing the NVIDIA A100-SXM machine.

FUNDING INFORMATION

The authors state that no funding is involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Meilany Nonsi Tentua	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
Suprapto		✓		✓		✓		✓		✓	✓	✓		
Afiahayati	✓		✓	✓			✓			✓	✓	✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors state no conflict of interest.

DATA AVAILABILITY

The dataset used in this study is described in [26] and is openly available at Mendeley Data (http://doi.org/10.17632/5s8r9ndfvc.3).

REFERENCES

[1] H. Lauder and K. Mayhew, “Higher education and the labour market: an introduction,” *Oxford Review of Education*, vol. 46, no. 1, pp. 1–9, Jan. 2020, doi: 10.1080/03054985.2019.1699714.
[2] H. Tushar and N. Sooraksa, “Global employability skills in the 21st century workplace: A semi-systematic literature review,” *Heliyon*, vol. 9, no. 11, pp. 1–14, Nov. 2023, doi: 10.1016/j.heliyon.2023.e21023.
[3] ILO, *World employment and social outlook*. 2023, doi: 10.54394/xmeg0270.
[4] W. Bank, *The Promise of Education in Indonesia*. World Bank, Washington, DC, 2020, doi: 10.1596/34807.
[5] J. Muzam, “The Challenges of Modern Economy on the Competencies of Knowledge Workers,” *Journal of the Knowledge Economy*, vol. 14, no. 2, pp. 1635–1671, Jun. 2023, doi: 10.1007/s13132-022-00979-y.
[6] Indeed, “Job Search _ Jobs by Indeed,” id.indeed.com. [Online]. Available: https://id.indeed.com/. (Date accessed: Aug. 12, 2020).
[7] Jobstreet, “Lowongan kerja di Indonesia - Karir,” id.jobstreet.com. [Online]. Available: https://id.jobstreet.com/. (Date accessed: Sep. 03, 2020).

[8] R. Boselli, M. Cesarini, F. Mercorio, and M. Mezzanzanica, "Classifying online Job Advertisements through Machine Learning," *Future Generation Computer Systems*, vol. 86, pp. 319–328, Sep. 2018, doi: 10.1016/j.future.2018.03.035.
[9] A. Debbarma, P. Bhattacharya, and B. S. Purkayastha, "Named entity recognition for a low resource language," *International Journal of Recent Technology and Engineering*, vol. 8, no. 3, pp. 587–590, Sep. 2019, doi: 10.35940/ijrte.B2085.098319.
[10] A. Vaswani *et al.*, "Attention Is All You Need," in *Advances in neural Information Processing Systems*, 2017, pp. 1–11.
[11] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for computational Linguistics: Human Language Technologies*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2019, pp. 4171–4186, doi: 10.18653/v1/N19-1423.
[12] M. Zhang, K. N. Jensen, R. van der Goot, and B. Plank, "Skill Extraction from Job Postings using Weak Supervision," in *CEUR Workshop Proceedings*, 2022, vol. 3218.
[13] J.-J. Decorte, J. Van Hautte, T. Demeester, and C. Develder, "JobBERT: Understanding Job Titles through Skills," *The International Workshop on Fair, Effective And Sustainable Talent management using data science*, pp. 1–9, 2021.
[14] H. T. Tran, H. H. P. Vo, and S. T. Luu, "Predicting Job Titles from Job Descriptions with Multi-label Text Classification," in *2021 8th NAFOSTED Conference on Information and Computer Science (NICS)*, IEEE, Dec. 2021, pp. 513–518, doi: 10.1109/NICS54270.2021.9701541.
[15] Q. Chen *et al.*, "CareerMiner: Automatic extraction of professional network from large Chinese resume data," *Franklin Open*, vol. 6, pp. 1–14, Mar. 2024, doi: 10.1016/j.fraope.2023.100065.
[16] J. Käser, T. Hanne, and R. Dornberger, "Large Language Models for Named Entity Recognition (NER) of Skills in Job Postings in German," in *Lecture Notes in Networks and Systems*, vol. 1047, pp. 21–31, 2024, doi: 10.1007/978-3-031-64836-6_3.
[17] M. Gavrilescu, F. Leon, and A. A. Minea, "Techniques for Transversal Skill Classification and Relevant Keyword Extraction from Job Advertisements," *Information (Switzerland)*, vol. 16, no. 3, pp. 1–24, 2025, doi: 10.3390/info16030167.
[18] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," in *Proceedings ofthe 28th International Conference on Computational Linguistics*, Stroudsburg, PA, USA: International Committee on Computational Linguistics, 2020, pp. 757–770, doi: 10.18653/v1/2020.coling-main.66.
[19] F. Koto, J. H. Lau, and T. Baldwin, "INDOBERTWEET: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization," in *Empirical Methods in Natural Language Processing (EMNLP 2021), Dominican Republic (virtual)*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2021, pp. 10660–10668, doi: 10.18653/v1/2021.emnlp-main.833.
[20] P. Dufter, M. Schmitt, and H. Schütze, "Position Information in Transformers: An Overview," *Computational Linguistics*, vol. 48, no. 3, pp. 733–763, 2022, doi: 10.1162/coli_a_00445.
[21] H. Yan, B. Deng, X. Li, and X. Qiu, "TENER: Adapting Transformer Encoder for Named Entity Recognition," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence (Italy): Association for Computational Linguistics, 2019, pp. 1–10.
[22] X. Zhou, Z. Ren, S. Zhou, Z. Jiang, T. Z. Yu, and H. Luo, "Rethinking Position Embedding Methods in the Transformer Architecture," *Neural Processing Letters*, vol. 56, no. 2, 2024, doi: 10.1007/s11063-024-11539-7.
[23] Z. Huang, D. Liang, P. Xu, and B. Xiang, "Improve transformer models with better relative position embeddings," in *Findings of the Association for Computational Linguistics Findings of ACL: EMNLP 2020*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2020, pp. 3327–3335, doi: 10.18653/v1/2020.findings-emnlp.298.
[24] Loker.id, "Situs Lowongan Kerja Indonesia," loker.id. [Online]. Available: https://www.loker.id/. (Date accessed: Jun. 03, 2021)
[25] Job.id, "Bursa Lowongan Kerja Terbaru," job.id. [Online]. Available: https://job.id/. (Date accessed: Nov. 23, 2020).
[26] M. N. Tentua, Suprapto, and Afiahayati, "NERSkill.Id: Annotated dataset of Indonesian's skill entity recognition," *Data in Brief*, vol. 53, pp. 1–5, Apr. 2024, doi: 10.1016/j.dib.2024.110192.
[27] N. Alshammari and S. Alanazi, "An Arabic dataset for disease named entity recognition with multi-annotation schemes," *Data*, vol. 5, no. 3, pp. 1–8, Jul. 2020, doi: 10.3390/data5030060.
[28] A. Deshmukh and A. Raut, "Applying BERT-Based NLP for Automated Resume Screening and Candidate Ranking," *Annals of Data Science*, vol. 12, no. 2, pp. 591–603, Apr. 2025, doi: 10.1007/s40745-024-00524-5.
[29] H. Kim and H. Kim, "Fine-grained named entity recognition using a multi-stacked feature fusion and dual-stacked output in Korean," *Applied Sciences (Switzerland)*, vol. 11, no. 22, pp. 1–10, Nov. 2021, doi: 10.3390/app112210795.
[30] K. Mendbayar and M. Aono, "Token Level Evaluation and Feature Enhancer for Transformer-based Models," in *2021 8th International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA)*, IEEE, Sep. 2021, pp. 1–6, doi: 10.1109/ICAICTA53211.2021.9640269.
[31] M. N. Tentua, Suprapto, and Afiahayati, "an Enhanced Bidirectional Encoder Transformers With Relative Position for Indonesian Skill Recognition," *ICIC Express Letters - An International Journal of Research and Survey*, vol. 18, no. 4, pp. 325–332, 2024, doi: 10.24507/icicel.18.04.325.

BIOGRAPHIES OF AUTHORS

Meilany Nonsi Tentua received a B.S. degree in Mathematics from Universitas Gadjah Mada, Indonesia, in 1996. She received a Master's degree in Electrical Engineering and a Doctorate in Computer Science from Universitas Gadjah Mada, Indonesia, in 2023. She has been a lecturer at Informatics, Universitas PGRI Yogyakarta, Indonesia, since 2006. She has been an associate Professor at Informatics, Universitas PGRI Yogyakarta, since 2024. Her research interests include machine learning, data science, data mining, natural language processing, and artificial intelligence. She can be contacted at email: meilany@upy.ac.id.

Suprapto received a B.S. degree in Mathematics from Universitas Gadjah Mada, Indonesia, in 1987. He received his M.S. degree in computer science from the Universitas Indonesia, Indonesia, and a Doctorate in Computer Science from Universitas Gadjah Mada, Indonesia, in 2015. From 2001, he was a lecturer at the Department of Computer Science and Electronics, Universitas Gadjah Mada, Indonesia. Since 2024, he has been a Professor at the Department of Computer Science and Electronics, Gadjah Mada University, Indonesia. His research interests include simulation modelling and computational modelling. He was a recipient of a sandwich program at UI-UMUC Maryland, the USA, in 1990. Since 2021, he has served as chairman of the Certification Department in the IndoCEISS. Since 2015, he has been a registered member of Aptikom and IAENG. He can be contacted at email: sprapto@ugm.ac.id.

Afiahayati received a B.S. degree in Computer Science from Universitas Gadjah Mada, Indonesia, in 2008. She received her M.S. degree in computer science from Universitas Gadjah Mada, Indonesia, and her Ph.D. in Biosciences and Informatics from Keio University, Japan, in 2015. She is an Assistant Professor at the Department of Computer Science and Electronics, FMIPA, Universitas Gadjah Mada. Currently, she is chair of the Master of Artificial Intelligence program at the Department of Computer Science and Electronics, FMIPA, Universitas Gadjah Mada. Her research interests include metagenomic assembly, genome assembly, comparative genomics, machine learning, probabilistic methods, bioinformatics, data mining, and artificial intelligence. She can be contacted at email: afia@ugm.ac.id.