Named Entity Recognition in Indonesian History Textbook Using BERT Model

Ichwanul Muslim Karo Karo*[1], Muliawan Firdaus[2], Rizki Habibi[3]
[1,2]Center of Excellence for Future Innovative Science Education; Medan
[3]Computer Science, Faculty of Mathematics and Natural Science, Universitas Negeri Medan
e-mail: *[1]**ichwanul@unimed.ac.id**, [2] muliawanfirdaus@unimed.ac.id , [3] rizki@unimed.ac.id

Abstract

The decreasing emphasis on history as an independent subject in the Indonesian education curriculum raises concerns about students' understanding of national heritage. Although historical content remains available in digital social science books, information on key entities such as people, places, and organizations is often difficult to understand and interpret, especially manually on a large scale. To address this, this research applies a Named Entity Recognition (NER) approach based on Bidirectional Encoder Representations from Transformers (BERT). It automatically identifies and classifies historical entities in the 8th-grade Social Science book published by the Ministry of Education. This research includes the stages of data extraction, text preprocessing, IOB labeling, BERT model implementation, and performance evaluation. The BERT-based NER model achieved strong results with 88.68% precision, 74.60% recall, and 81.03% F1-score, demonstrating its ability to effectively recognize entities such as PERSON, LOCATION, and ORGANIZATION. These findings suggest that BERT-based NER is a promising tool for structuring historical information, supporting historical research and educational enrichment.

Keywords— NER, BERT, IOB labeling, History

1. INTRODUCTION

Referring to the recent education curriculum, history is not explicitly recognized as a subject in some education-level institutions anymore [1], [2]. The subject is merged with other subjects like sociology, economics, and geography in social science subjects at the junior high school level[1]. Even the subject is not present in vocational schools and in some other schools. This is an irony for future generations. In fact, Indonesia has a vast history with various events, figures, places, and organizations. A great nation is a nation that does not forget its history. Learning history is not to live in the past, but to understand the present and build the future. The lack of time to study history will certainly affect the younger generation, who do not recognize the history of the nation.

Meanwhile, the technology era has changed the paradigm of presenting a history book [3]. Initially, books had to be hard copies, but now they can be presented in the form of soft files. Furthermore, sources of historical knowledge can also come from essays, journals, and articles from the web. There are many important entities about Indonesian history in the textbook. Surely, the Indonesian historical data listed in the information sources are not yet presented in a structured manner, and the various important entities are quite difficult to understand. To support historical research and cultural preservation, a method that can extract historical information automatically and structured format is needed.

Named Entity Recognition (NER) is a technique in Natural Language Processing (NLP) used to identify and classify specific entities [4] such as names of people, places, locations, organizations and etc.. This technique helps in the extraction of information from unstructured text and enables computers to understand the context and meaning of the information contained therein. Thus, historical information can be understood by students and the public through NER technology.

Many NLP algorithms or models have been developed for NER, such as BERT (Bidirectional Encoder Representations from Transformers). This method has a depth capability to understand the context in sentences [5]. It identifies entities with much higher precision even in complex languages such as Indonesian [6]. BERT is a significant breakthrough in NLP released by Google in 2018 [7]. This model architecture presents a revolutionary approach to understanding language context through a bidirectional encoding mechanism [8]. This allows the model to understand the meaning of words based on the context of the entire sentence. In its architecture, BERT uses encoder transformers as the main foundation for pre-training various NLP tasks such as sentiment analysis [9], question answering [10], recommender systems [11], text summarization [12], and NER[13]. The BERT training process involves a vast and diverse dataset of text and code. These data sources include various documents such as books, articles, and programming code. This allows the model to build a comprehensive understanding of words. BERT uses language knowledge from the training process for sentence generation or understanding. The model does not simply translate word- by-word, it considers complex factors such as grammar, semantic meaning, and the overall context of the sentence. Through a bidirectional encoding mechanism, BERT can capture the complex contextual relationships between words in a sentence, distinguishing it from previous language models that tend to be unidirectional or partial [14].

Despite the abundance of historical information in school textbooks, there is a lack of automated methods to extract key historical entities. To address this gap, this study investigates how the BERT model can be applied to extract and classify historical entities from history chapters in 8th-grade social science textbooks.

2. RESEARCH METHODS

This section presents a chronological explanation of the research process. This is depicted in the flowchart (Figure 1), the research stages start from the data extraction process, text pre-processing, IOB labeling, NER process (in the form of an algorithm and Pseudocode), and evaluation. Each process is comprehensively explained with examples.

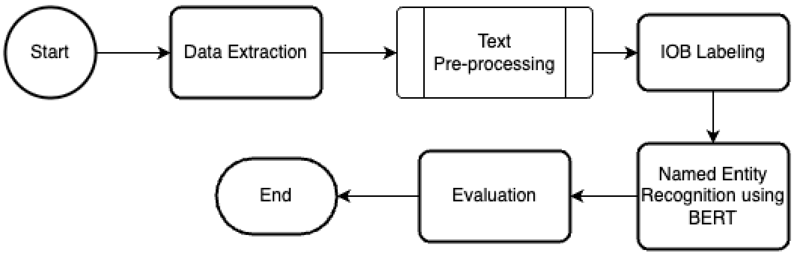

Figure 1. Research Flowchart

2.1. Data Extraction

This study extracted data from the text of the Social Science book for grade 8 (Figure 2). The book was published by the Ministry of Education in 2017[15]. There are four scientific areas in the book: Geography, History, Sociology, and Economics. In the previous curriculum, these fields of science were independent subjects[16]. However, those are now merged into a subject (social science). This research only extracts text from the chapter that discusses history, which is pages 193-258.

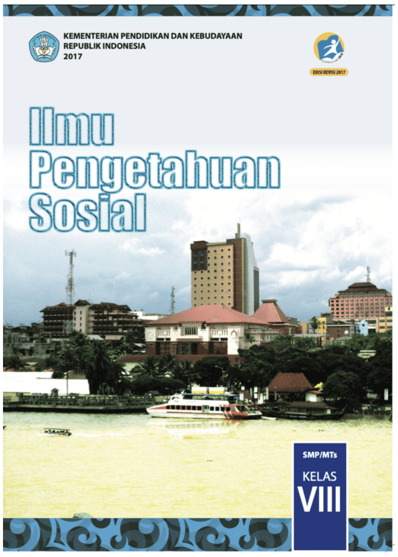

Figure 2. The Grade 8 Social Science Book Cover

2.2. Text Preprocessing

Text preprocessing is a crucial step in various NLP applications, including NER[11]. The main focus of text pre-processing is to prepare the data, then it becomes structured and analyzable. In addition, it improves the accuracy and efficiency of models [17]. There are many text pre-processing methods. This research applies some common text preprocessing methods (shown in Table 1). Tokenization is a fundamental step in text pre-processing. The text from the book is segmented into words. Segmentation allows the model to analyze and label each token individually. For example, the phrase "*Negara Indonesia*" is tokenized into ["*Negara*", "*Indonesia*"], which allows the model to recognize "*Indonesia*" as a potential entity. Without accurate tokenization, entities can be missed or misclassified. Historical texts often contain punctuation, exclamation marks, or special symbols. The symbols lack contribution to the semantic meaning of the entity. The removing punctuation and special characters step removes these elements, reducing noise and preventing the model from misinterpreting the punctuation as part of the entity. For example, in the phrase "Freedom!", the exclamation mark adds emphasis but is not relevant for entity recognition. Cleaning this up will ensure cleaner and easier-to-analyze input for the model. Stopwords are commonly used words, such as conjunctions and prepositions. Generally, it does not carry any meaningful information for entity extraction. In the context of Indonesian historical texts, words such as "*yang*", "*dan*", or "*di*" appear frequently, but provide little value in identifying names, places, or organizations. Removing these words helps to streamline the dataset, reduce computational burden, and increase the model's focus on content-rich terms that are more likely to be entities. Case normalization is the step of altering letters to lowercase. This prevents the model from treating the same word differently due to letter variations. For example, "*Indonesia*" and "*indonesia*" should be interpreted as the same entity. The use of lowercase minimizes token redundancy and ensures consistency in how the model interprets words, which is especially useful when dealing with text from different sources or formatting styles. Stemming and lemmatization transform words into their base or root form. This is especially useful in Indonesian. Verbs and nouns often appear in affixed or derived forms, for example, "*diperbudak*" is transformed into "*budak*" which represents the core meaning. This process helps disambiguate word variations and improves the model's ability to recognize entities even if they appear in different grammatical forms throughout the text.

Table 1. Text Pre-processing Methods

Methods	Description	Example in Indonesian
Tokenization	Split text into words or sentences	"*Negara Indonesia*" → ["*Negara*", "*Indonesia*"]
Removing Punctuation & Special Characters	Eliminate unnecessary symbols	"*Merdeka!*" → "*Merdeka* "
Stopword Removal	Remove common words and conjunctions	" *negara yang besar* " → "*negara besar*"
Lowercasing	Convert all text to lowercase to avoid case sensitivity	"*Merdeka atau mati*" → "*merdeka atau mati* "
Stemming and lemmatization	Reduce and convert words to their base dictionary form	" *diperbudak*" → "*budak*"

2.3. IOB Labeling

IOB (Inside-Outside-Beginning) labeling is a common tagging scheme used in NER and chunking tasks to identify and classify entities in text [18], [19], [20]. It helps in structuring unstructured text by marking different parts of an entity. IOB tagging scheme, each token in a sentence is assigned a label.

B- (Beginning) →The first token of a named entity.

I- (Inside) → Any token inside a named entity (except the first one).

O (Outside) → Tokens that are not part of any named entity.

Furthermore, the entities to be identified and labeled include PERSON, LOCATION, ORGANIZATION, and OTHER/OUTSIDE. The entity of PERSON identifies all character names in the textbook. The entity of LOCATION identifies all organizations and institutions. Table 2 presents an example implementation of IOB tagging for a sentence. B-PERSON labels a person's first name, while the last name is labeled with I-PERSON. The same is true for location or organization tagging.

Table 2. Example of IOB Tagging in a Sentence

Sentences	Entities	IOB Labeling
Portugis telah berhasil sampai di Maluku di bawah pimpinan Antonio de Abreu	PERSON = Antonio de Abreu	B-PERSON = Antonio
		I-PERSON = de Abreu
	LOCATION = Maluku	B-LOCATION = Maluku
		I-LOCATION = -
	ORGANIZATION = Portugis	B-ORGANIZATION = Portugis
		I-ORGANIZATION = -
	O = *telah, berhasil, sampai, bawah, pimpinan*	-

2.4. BERT Model

BERT (Bidirectional Encoder Representations from Transformers) is an advanced deep learning model developed for various NLP tasks [10]. It was introduced by Google in 2018. BERT is designed to fulfill the need for a model that can understand word context bi-directionally-both from left to right and right to left-unlike traditional models that are limited to one direction. This allows BERT to capture deeper semantic relationships between words, making it highly effective in tasks such as sentiment analysis, question answering, and NER. Philosophically, BERT represents a shift from rule-based or shallow statistical models towards a more context-aware approach to learning in NLP [7].

Actually, the BERT model can be trained to identify a wide variety of entities [9], [13], [21]. This research focuses on three categories: PEOPLE, ORGANIZATIONS, and LOCATIONS. These three entities are highly relevant and commonly discussed in historical texts. Names of important figures in Indonesian history, names of institutions, and geographical locations often appear in historical texts. The application of the BERT model to automatically extract these historical entities from the content of social science textbooks. This helps to structure the data to make it easier to understand and analyze.

```
BEGIN BERT Algorithm

  # STEP 1: INPUT PREPROCESSING
  INPUT: Raw Text
  TOKENIZE(text) → Convert text into WordPiece tokens
  ADD_SPECIAL_TOKENS(tokens) → Add [CLS] and [SEP] tokens
  CONVERT_TO_IDS(tokens) → Map tokens to unique vocabulary indices
  GENERATE_SEGMENT_IDS(tokens) → Assign segment embeddings
  GENERATE_POSITION_IDS(tokens) → Assign position embeddings
  EMBEDDING_LAYER(tokens) → Convert to dense vectors

  # STEP 2: TRANSFORMER ENCODER
  FOR each token in input_sequence DO:
    FOR each Transformer layer L in {1, ..., N} DO:
      SELF_ATTENTION(tokens) → Compute contextualized representations
      FEED_FORWARD_NETWORK(tokens) → Pass through a fully connected layer
      APPLY_LAYER_NORM(tokens) → Normalize activations
    END FOR
  END FOR

  # STEP 3: PRETRAINING OBJECTIVES
  IF pretraining THEN:
    SELECT_RANDOM_TOKENS(tokens) → Mask 15% of tokens as [MASK]
    MLM_PREDICTION(tokens) → Predict masked words using softmax
    NSP_CLASSIFICATION(sentence_1, sentence_2) → Predict if sentence_2 follows sentence_1
  END IF

  # STEP 4: FINE-TUNING
  IF fine_tuning THEN:
    LOAD_PRETRAINED_BERT_MODEL()
    ATTACH_TASK_SPECIFIC_LAYER(task)
    TRAIN_MODEL_ON_TASK(dataset)
  END IF

  # STEP 5: OUTPUT GENERATION
  COMPUTE_FINAL_EMBEDDINGS(tokens)
  GENERATE_PREDICTIONS(tokens)

  RETURN output_predictions

END BERT Algorithm
```

BERT requires the raw text to be converted into a specific input format so that it can process the input text effectively[10]. This involves several stages. First, the text is tokenized using the WordPiece tokenizer, which breaks words into subword units to handle rare or invisible words. Next, special tokens [CLS] and [SEP] are added to indicate the beginning and separation of sequences. These tokens are crucial in helping the model distinguish between input segments during processing. Each token is then mapped to a corresponding vocabulary index (Token ID). For example, Token ID 101 usually represents [CLS], and Token ID 102 indicates [SEP]. These IDs serve as inputs to the insertion layer.

The BERT model uses two main objectives during pre-training. Those are Masked Language Modeling (MLM) and Next Sentence Prediction (NSP)[13], [22]. In MLM, 15% of the input tokens are randomly selected and masked. However, these tokens are not always replaced with [MASK] tokens to avoid overfitting during the fine-tuning process-sometimes the word is retained or replaced with a random token. This teaches the model to infer the correct word from the context. It can thus improve its understanding of sentence-level semantics. Whereas NSP to predict whether a given sentence *B* logically follows sentence *A*, which further improves its contextual understanding.

After pre-training, BERT is fine-tuned for the specific task of entity recognition. In this phase, a task-specific classification layer is added on top of the BERT architecture[13]. The contextual embedding of each token (output from the final Transformer layer) is forwarded to this layer to predict the corresponding entity label, e.g., B-PERSON, I-LOCATION, or O for non-entity tokens. The model is trained on labeled datasets using the IOB (Inside-Outside-Beginning) tagging scheme. It helps distinguish between contiguous entities and unrelated text segments.

2.5. Evaluation Indicators

This research uses several metrics to evaluate successfully extracted entities. These are eval loss, precision, recall, and F1 score. "Eval loss" (evaluation loss) refers to the loss value calculated on a validation or test dataset during model evaluation [23]. It helps measure how well a BERT model is performing on testing data. Eval Loss describes how far the model's predictions differ from the actual labels. The lower the loss, the better the model performs in classifying entities.

Precision and recall are the two main performance metrics commonly used to evaluate NER. However, each metric captures a different aspect of model performance. Precision measures the proportion of correctly predicted positive entities out of all entities predicted as true entities (How many of the entities the model predicted were actually correct?)[24]. Recall measures the proportion of correctly predicted true entities out of all actual entities in the data [24].

Sometimes, a model may have high precision but low recall (i.e., it only predicts when it is very sure, but misses many entities), or high recall but low precision (i.e., it detects many entities but also produces many false positives)[17]. To balance these two aspects, this study uses the F1 score as in (1), which is the harmonic mean of precision and recall[24]. The best value of the F1 score is 1 or 100% (perfect precision and recall), and the worst value is 0.

$$F1\ score = \frac{2\ precision.recall}{precision + recall} \quad (1)$$

3. RESULT AND DISCUSSION

3.1. Experimental Environment

The experiments were conducted on macOS Sonoma 14.1.2 with the following hardware information: 8 GB of RAM and 245.11 GB ROM. This experiment uses the Python 3.6 programming language via Google Colab to extract text from the textbook and build NER models.

3.2. Experimental Dataset and Text Pre-processing

This research dataset is sourced from the history chapter in the Social Science textbook for junior high school students in grade VIII. The book was published by the Ministry of Education and Culture in 2017. The textbook is the main reference for history education in Indonesia, and includes important historical content such as colonialism, national liberation movements, key figures, and institutional development. A total of 1,465 sentences were extracted exclusively from pages 193 to 258. The dataset contains historical entities (PEOPLE, LOCATIONS, ORGANIZATIONS) that can be used with the BERT model to extract them.

Table 3. Example of Sentences

Id	Sentences
1	*Wilayah Indonesia sangat luas, kekayaan melimpah, dan kaya akan berbagai budaya*
2	*Kondisi inilah yang menjadi daya tarik bangsa-bangsa lain datang ke Indonesia*
3	*Hubungan dagang dan interaksi budaya merupakan contoh hubungan Indonesia dengan bangsa-bangsa asing*
4	*Hubungan dengan bangsa-bangsa asing tidak hanya berdampak positif, tetapi juga negatif*
. . .	. . .
1465	*Tugas Komisi adalah menentukan istilah-istilah modern dan menyusun suatu tata bahasa normatif serta ...*

The quality of the input data plays an important role in the performance of the NER system. Therefore, a comprehensive text pre-processing phase was implemented to prepare the dataset. Pre-processing not only improves the efficiency of the BERT model but also helps to reduce noise, minimize redundancy, and improve the model's focus on semantically relevant tokens in Indonesian.

Figure 3. Word Cloud Sentences from History Chapter before Text Preprocessing

Before pre-processing, high-frequency terms, including conjunctions and common words, appear in the raw text (Figure 3). The pre-processing pipeline applied in this study includes tokenization, punctuation and special character removal, stopword removal, lowercase letters, and stemming/lematization. After pre-processing, the number of sentences was slightly reduced to 1,450, and the number of words dropped from 16,243 to 14,751. The total number of unique terms also decreased slightly, indicating that redundant or irrelevant terms were effectively removed. This reduction not only improves computational efficiency, but also supports the model's ability to focus on linguistically meaningful components-especially important in history-related texts where entity references can be subtle and context-dependent.

Figure 4 shows the word cloud after text-preprocessing. The dominant terms reflected the historical themes of the dataset-such as "*Indonesia*", '*Belanda*', and "*Jepang*"-which confirmed the relevance of the dataset to the research topic. These dominant terms indicate that the text is rich in historical content and are suitable for evaluating the BERT model's ability to extract historical entities. From the words that appear, we argue that the history chapter of the book discusses the Dutch and Japanese colonialism in Indonesia, the resistance of the Indonesian people to colonialism, national movements and struggle organizations, the impact of colonialism on society, economy, and politics, and social and political changes during the colonial period.

Figure 4. Word Cloud Sentences from History Chapter after Text Preprocessing

Table 4 presents a comparison of the conditions before and after text processing. There is a reduction in the number of words by about 1,492 words. This could be due to processes such as stopword removal, stemming, or special character removal. The reduction of 15 sentences could be due to punctuation removal or merging of multiple sentences during preprocessing. The reduction in word count indicates that the text is cleaner and more optimized for further analysis.

Table 4. Comparison of Datasets Before and After Text Pre-processing Stages

	Before Text Preprocessing	**After Text Preprocessing**
Word	16243	14751
Unique Term	3370	3350
Sentences	1465	1450

3.3. BERT Model Performance

Figure 5 is a graph showing the training time per epoch (in minutes) in a training process with the BERT algorithm. The X-axis shows the number of epochs in the model training, from epoch 1 to epoch 15. The Y-axis shows the time taken to complete each epoch in minutes. The time per epoch fluctuates, but remains within the range of 16 - 17.2 minutes. There were a few points where the time increased significantly, especially in epochs 10 and 11, which reached the highest time (around 17.2 minutes). After epoch 11, the execution time decreased again and remained relatively stable until epoch 15. In general, the training time was quite stable in most epochs.

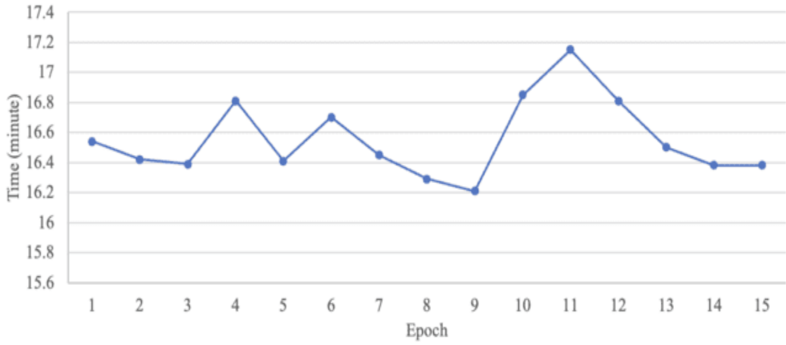

Figure 5. Training Time per Epoch

This research analyzes several factors that can cause fluctuations in training time per epoch. Variations in the number of words and entities identified per epoch are not always of the same size. Table 5 gives us an idea that if in epoch 1 the BERT algorithm processes the sentences "The Portuguese had made it to the Moluccas under Antonio de Abreu" and in epoch 2 the BERT algorithm identifies the NER in the sentence "One of the causes of resistance was the rejection of the custom of honoring the sun", then the training time of epoch 1 is higher than the training time of epoch 2. The BERT algorithm identifies four entities, while in epoch 2, only one entity. So the capacity and workload of GPU/CPU and cache memory allocation will change during training. In other words, inconsistent batch processing.

Figure 6 illustrates the evaluation loss (Eval Loss) trend over 15 epochs of the BERT model for extracting historical entities in books. The X-axis represents the number of epochs—indicating a full pass through the dataset—while the Y-axis shows the corresponding loss value on the validation set. At the beginning of training (epoch 1), the Eval Loss was relatively high (1.439). That's because the model started with random weight initialization and had not yet learned meaningful patterns. As training progressed, the Eval Loss decreased significantly, indicating that the model effectively learned to recognize historical entities such as PERSON, LOCATION, and ORGANIZATION. At the 5th epoch, the Eval Loss dropped to 0.414 and reached its lowest point at the 7th epoch (0.061), indicating that the model had achieved a high level of generalization at this stage.

This downward trend occurs due to several factors. Through backpropagation and gradient descent optimization, the model continuously updates its internal parameters to minimize prediction errors. The bidirectional BERT architecture also plays a key role, as it allows the model to learn contextual relationships between tokens. It is very important in accurately identifying entities. Additionally, the repetition of similar entity structures throughout the dataset helps the model internalize and recognize recurring patterns more effectively.

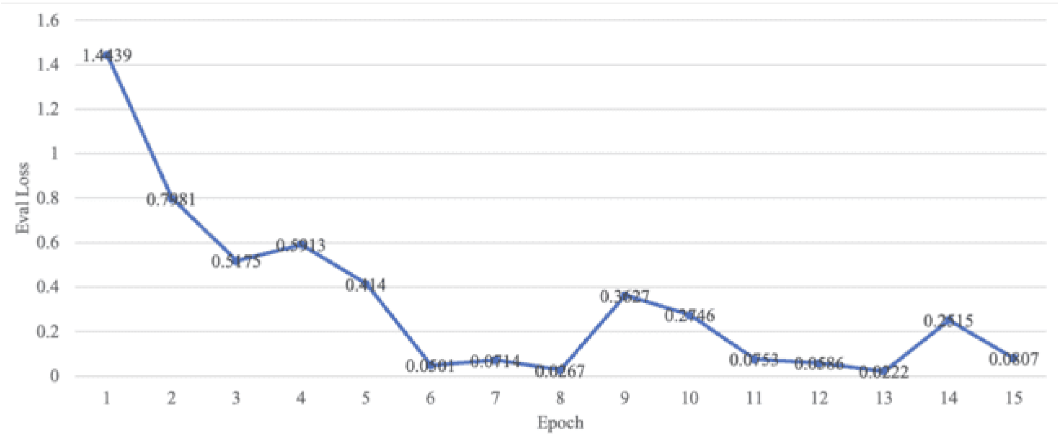

Figure 6. Eval Loss to Epoch

However, after the 7th epoch, the evaluation loss began to fluctuate. It increased to 0.3627 at the 9th epoch and varied slightly in the following epochs. These fluctuations may be caused by a tendency to overfit, where the model starts memorizing the training data instead of generalizing to unseen inputs. Other factors, such as the complexity of the evaluation batch and the optimization plateau, may also contribute to this variation. Despite the fluctuations, the evaluation loss remains relatively low, indicating that the model maintains strong performance over time. Overall, this graph shows that the BERT model converges effectively at epoch 7, and additional training beyond that point should be closely monitored to prevent overfitting. In summary, the performance of the BERT model is shown in Table 5. The Eval loss of the model is 0.306, the precision is 88.68%, the Recall is 74.6% and the F1 score is 81.03%. Surely, that is a gap and a challenge to improve the model to produce good results.

Table 5. Performance of the BERT Model

Eval Loss	Precision	Recall	F1 Score
0.306	88.68%	74.60%	81.03%

The BERT model can understand sentence context, which is very important in identifying historical figures and locations. Table 6 shows an example of the BERT model has been shown to perform well in various NER tasks. However, BERT mistakenly classified VOC as a location and Batavia as an organization. This error is likely caused by the ambiguous sentence context, where both entities often appear together in geographical and administrative contexts. This shows that although BERT understands the global context, the understanding of specific historical entities can still be incorrect if the entities have a close semantic relationship.

Table 6. Simulation of NER on Sentences for Every Epoch

Epoch	Sentences	Entities
1	*Portugis telah berhasil sampai di Maluku di bawah pimpinan Antonio de Abreu*	B- PERSON = Antonio
		B- LOCATION = Maluku
		B-ORGANIZATION = Portugis
		O = *telah, berhasil, sampai, bawah, pimpinan*
2	*Salah satu penyebab perlawanan adalah penolakan terhadap kebiasaan menghormat matahari*	PERSON = -
		LOCATION = -
		ORGANIZATION = -
		O = *Salah satu penyebab perlawanan adalah penolakan terhadap kebiasaan menghormat matahari*

Understanding the context of historical entities can help in understanding the relationship between figures and locations in the context of certain historical events. For example, the younger generation is no longer confused to say that Soekarno was a figure, not a street name, the VOC was an organization, not a vocal.

4. CONCLUSION

This research proves that the BERT (Bidirectional Encoder Representations from Transformers) algorithm can be effectively applied to perform Named Entity Recognition (NER) on Indonesian historical texts from middle school textbooks. Through a series of stages—from data extraction, text preprocessing, IOB labeling, to model training and evaluation—achieved a fairly high performance with a precision value of 88.68%, a recall of 74.60%, and an F1-score of 81.03%. These results indicate that BERT is capable of automatically and accurately identifying entities such as names of figures (PERSON), locations (LOCATION), and organizations (ORGANIZATION) from unstructured text.

The application of systematic text preprocessing techniques also contributes to improving the efficiency and accuracy of the model by reducing irrelevant words and highlighting important information in historical texts. Nevertheless, the model still shows some limitations, such as difficulty recognizing rarely appearing entities and errors in determining the boundaries of complex entities. This presents an opportunity for future improvement through the expansion of the training corpus, domain adaptation, and reinforcement of models based on local context.

Overall, this research shows that modern NLP technologies like BERT have great potential in supporting history learning, cultural preservation, and the development of digital education systems based on structured information. With this approach, the younger generation can understand the nation's history more deeply through the assistance of intelligent and adaptive technology.

REFERENCES

[1] E. Suparjan, “Perubahan Kurikulum Pendidikan Sejarah Di SMA (1994-2013),” *JISIP (Jurnal Ilmu Sosial dan Pendidikan)*, vol. 4, no. 3, 2020, doi: 10.36312/jisip.v4i3.1283.

[2] R. A. Pratama, Maskun, and N. I. Lestari, “Dinamika Pelajaran Sejarah Indonesia dalam Kurikulum 2013 pada Jenjang SMK/MAK,” *Jurnal Pendidikan Sejarah*, vol. 8, no. 2, 2019, doi: 10.21009/jps.082.02.

[3] Istianah and S. Wahyuningsih, “The hadith digitization in millennial era: A study at center for hadith studies, Indonesia,” *Qudus International Journal of Islamic Studies*, vol. 7, no. 1, 2019, doi: 10.21043/qijis.v7i1.4900.

[4] N. S. Lagutina, A. M. Vasilyev, and D. D. Zafievsky, “Name Entity Recognition Tasks: Technologies and Tools,” *Modeling and Analysis of Information Systems*, vol. 30, no. 1, 2023, doi: 10.18255/1818-1015-2023-1-64-85.

[5] A. Ushio, L. Espinosa-Anke, S. Schockaert, and J. Camacho-Collados, “BERT is to NLP what AlexNet is to CV: Can pre-trained language models identify analogies?,” in *ACL-IJCNLP 2021 - 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, 2021. doi: 10.18653/v1/2021.acl-long.280.

[6] S. Zou, Y. Xie, J. Yan, Y. Wei, and X. Luan, “Research on Image Captioning Based on Vision-language Pre-trained Models,” in *2023 9th International Conference on Big Data and Information Analytics, BigDIA 2023 - Proceedings*, 2023. doi: 10.1109/BigDIA60676.2023.10429361.

[7] S. Pichai, “Google AI updates: Bard and new AI features in Search,” Google The Keyword Blog.

[8] A. J. Keya, M. A. H. Wadud, M. F. Mridha, M. Alatiyyah, and M. A. Hamid, "AugFake-BERT: Handling Imbalance through Augmentation of Fake News Using BERT to Enhance the Performance of Fake News Classification," *Applied Sciences (Switzerland)*, vol. 12, no. 17, 2022, doi: 10.3390/app12178398.

[9] E. I. Setiawan, L. Kristianto, A. T. Hermawan, J. Santoso, K. Fujisawa, and M. H. Purnomo, "Social Media Emotion Analysis in Indonesian Using Fine-Tuning BERT Model," in 3rd 2021 *East Indonesia Conference on Computer and Information Technology, EIConCIT 2021*, 2021. doi: 10.1109/EIConCIT50028.2021.9431885.

[10] R. C. G. Ramos, H. D. Calderón-Vilca, and F. C. Cárdenas-Mariño, "A BERT-based Question Answering Architecture for Spanish Language," *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 14, 2022.

[11] D. H. Fudholi, A. Zahra, S. Rani, S. N. Huda, I. V. Paputungan, and Z. Zukhri, "BERT-based tourism named entity recognition: making use of social media for travel recommendations," *PeerJ Comput Sci*, vol. 9, 2023, doi: 10.7717/PEERJ-CS.1731.

[12] Y. Iwasaki, A. Yamashita, Y. Konno, and K. Matsubayashi, "Japanese abstractive text summarization using BERT," *Advances in Science, Technology and Engineering Systems*, vol. 5, no. 6, 2020, doi: 10.25046/AJ0506199.

[13] E. T. Luthfi, Z. I. M. Yusoh, and B. M. Aboobaider, "BERT based Named Entity Recognition for Automated Hadith Narrator Identification," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 1, 2022, doi: 10.14569/IJACSA.2022.0130173.

[14] S. Sofiana, *Konsep Bert Pada Natural Language Processing*, vol. 225. Jawa Tengah: Eureka Media Aksara, 2021.

[15] I. Kementerian Pendidikan dan Kebudayaan, *Buku Paket Ilmu Pengetahuan Sosial Kelas 8*, vol. 3, no. 4. 2017.

[16] A. P. Anar, A. Widodo, and D. Indraswati, "Menilik Jejak Historis Pendidikan IPS Di Indonesia: Konsep Dan Kedudukan Pendidikan IPS Dalam Perubahan Kurikulum Di Sekolah Dasar," *Phinisi Integration Review*, vol. 5, no. 2, 2022, doi: 10.26858/pir.v5i2.33677.

[17] I. M. Karo Karo, M. F. M. Fudzee, S. Kasim, and A. A. Ramli, "Sentiment Analysis in Karonese Tweet using Machine Learning," *Indonesian Journal of Electrical Engineering and Informatics*, vol. 10, no. 1, pp. 219–231, Mar. 2022, doi: 10.52549/ijeei.v10i1.3565.

[18] A. Tehseen, T. Ehsan, H. Bin Liaqat, X. Kong, A. Ali, and A. Al-Fuqaha, "Shahmukhi named entity recognition by using contextualized word embeddings," *Expert Syst Appl*, vol. 229, 2023, doi: 10.1016/j.eswa.2023.120489.

[19] W. Hwang, J. Yim, S. Park, S. Yang, and M. Seo, "Spatial Dependency Parsing for Semi-Structured Document Information Extraction," in *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, 2021*. doi: 10.18653/v1/2021.findings-acl.28.

[20] M. Modrzejewski, T. Le Ha, A. Waibel, M. Exel, and B. Buschbeck, "Incorporating External Annotation to improve Named Entity Translation in NMT," in *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation, EAMT 2020*, 2020.

[21] M. Mujahid, K. Kanwal, F. Rustam, W. Aljadani, and I. Ashraf, "Arabic ChatGPT Tweets Classification Using RoBERTa and BERT Ensemble Model," *ACM*

Transactions on Asian and Low-Resource Language Information Processing, vol. 22, no. 8, 2023, doi: 10.1145/3605889.

[22] L. Zhang, P. Xia, X. Ma, C. Yang, and X. Ding, "Enhanced Chinese named entity recognition with multi-granularity BERT adapter and efficient global pointer," *Complex and Intelligent Systems*, vol. 10, no. 3, 2024, doi: 10.1007/s40747-024-01383-6.

[23] A. M. A. Barhoom, B. S. Abunasser, and S. S. Abu-Naser, "Sarcasm Detection in Headline News using Machine and Deep Learning Algorithms," 2022.

[24] I. M. K. Karo, R. Ramdhani, A. W. Ramadhelza, and B. Z. Aufa, "A Hybrid Classification Based on Machine Learning Classifiers to Predict Smart Indonesia Program," in *Proceeding - 2020 3rd International Conference on Vocational Education and Electrical Engineering: Strengthening the framework of Society 5.0 through Innovations in Education, Electrical, Engineering and Informatics Engineering, ICVEE 2020*, 2020. doi: 10.1109/ICVEE50212.2020.9243195.