Journal of Applied Engineering and Technological Science
Vol 6(1) 2024: 60-68

# *PRE-TRAINED BERT ARCHITECTURE ANALYSIS FOR INDONESIAN QUESTION ANSWER MODEL*

**Sudianto***[1]
Department of Informatics, Institut Teknologi Telkom Purwokerto, Indonesia[1]
sudianto@ittelkom-pwt.ac.id[1]

Received: 27 March 2024, Revised: 22 June 2024, Accepted: 25 July 2024
**Corresponding Author*

***ABSTRACT***
*Developing a question-and-answer system in Natural Language Processing (NLP) has become a significant concern in the Indonesian language context. One of the main challenges in developing a question-and-answer system is the limited dataset, which can cause instability in system performance. The limitations of the dataset make it difficult for the question-and-answer model to understand and answer questions well. The proposed solution uses Transfer Learning with pre-trained models such as BERT. This research aims to analyze the performance of the BERT model, which has been adapted for question-and-answer tasks in Indonesian. The BERT model uses an Indonesian language dataset adapted specifically for question-and-answer tasks. A customization approach tunes BERT parameters according to the given training data. The results show that the model is improved by minimizing the loss function. Evaluation of the trained model shows that the best validation loss is 0.00057 after 150 epochs. In addition, through in-depth evaluation of the similarity of question texts, the BERT model can answer questions measurably, according to existing knowledge in the dataset.*
***Keywords :*** *BERT, Chatbot, Indonesian, NLP, Transfer Learning, Question Answer*

## 1. Introduction

Automated questioning can change how we interact with digital information, from information seeking to deployment in virtual assistants or chatbots. Automated Question and Answer (Q&A) makes information more accessible to access, increasing productivity and providing more precise and accurate responses. In an era where Artificial Intelligence (AI) is increasingly involved in many aspects of life as (Sudianto et al., 2023a, 2023c, 2023b), tourism (Alotaibi et al., 2020; Naufal et al., 2023), education (Jhaerol & Sudianto, 2023; Nguyen et al., 2021) and e-commerce (Arnumukti et al., 2023) automated question-and-answer has the potential to streamline human-machine communication and provide intelligent solutions to natural language processing tasks, which are more relevant, adaptive, and efficient in the changing world. Changes in the dynamics of information and human interaction in the digital era.

The problems that exist in building the compatibility of the automatic question-and-answer language and the accuracy of interpreting the meaning of the language are challenging to understand, especially the automatic question-and-answer in Indonesian. According to data from TyDi QA, Indonesian has many dialect variations and uses different slang in different regions. The challenge in handling this variation is that the model must understand and interpret various forms of colloquialism that only sometimes correspond to formal Indonesian (Clark et al., 2020) . On the other hand, the data resources needed to train the automatic Q&A model in Indonesia still need to be improved. A systematic review shows that most datasets for automated question-and-answer are available in English, while Indonesian has less than 5% of the total available data (Cortes et al., 2022). In addition, the automated Q&A system must recognize and process these morphologies appropriately, which is a significant challenge because it requires an in-depth understanding of Indonesian grammar and morphology (Clark et al., 2020). Automatic question-and-answer through chatbots can be a solution for users who interact in Indonesian and computers who can understand sentence structures, word meanings, and local contexts in one Indonesian language. Indonesian is a complex and diverse language, and understanding the questions asked in the language requires providing appropriate and relevant answers. In this context, the development of automated Q&A using technologies such as BERT (Bidirectional Encoder Representations from Transformers), optimized for Indonesian, allows Indonesian-speaking users. With automatic debriefing that focuses on Indonesian, how can the potential of BERT work in automatic debriefing in Indonesian? Therefore, automated questions and answers need to be seriously reviewed, especially in Indonesian, for Indonesian-speaking users to be more inclusive and beneficial for users.

Previous research on automated question-and-answer services has developed. One of them is a question-and-answer for student admission with Deep Learning on the Sense framework (Nguyen et al., 2021) chatbot for Tourism Recommendations in Saudi Arabia with LSTM (Alotaibi et al., 2020), a chatbot for Tracking Student Educational Program Interests on social media with K-Nearest Neighbor's and Neural Network (Nasa-Ngium et al., 2023), and Chatbot for Indonesian for facial beauty products with IndoBERT (Indah Rahajeng & Purwarianti, 2021). Unlike the previous studies, the automated question-and-answer approach was constructed using a pre-trained model such as BERT in this study.

BERT was chosen to achieve a better level of language comprehension, overcome the constraints of limited training data, and save time and resources in development. BERT can improve user experience and increase the efficiency of Artificial Intelligence (AI) development (Devlin et al., 2019; Galassi et al., 2021). The main advantage of BERT is its ability to read text in both directions, namely from left to right and right to left (Devlin et al., 2019). BERT allows the model to capture nuances and complexities in sentences that one-way models might miss. For example, the words "can" in the sentences "The snake can bite" and "He can come tomorrow" have very different meanings. BERT can more accurately understand these differences by reading the context from both directions. Other studies show that self-attention in BERT models often captures patterns relevant to language's syntactic and semantic structure. BERT can capture critical information, such as relationships between words and sentences, crucial for deeper language understanding (Kovaleva et al., 2019). In addition, using pre-trained helps achieve Green AI by optimizing efficiency and contributing to environmental sustainability without burdening ecosystems and natural resources (Schwartz et al., 2020). Therefore, this study aims to perform a pre-training BERT architecture analysis for the Indonesian language question-answer model by showing the challenges in handling language variations in the automated question-and-answer model.

The flow of the discussion in this study is meticulously structured to provide a comprehensive understanding of the topic. It begins with a thorough exploration of related work, followed by a detailed discussion of relevant studies, materials, and methods. The data and data acquisition process used are then described, leading to the presentation of the results and subsequent discussion. Finally, the study concludes with insightful suggestions for further research, ensuring that the audience is well-informed and guided through the research process.

## 2. Related Work

The BERT (Bidirectional Encoder Representations from Transformers) architecture has been a milestone in natural language processing (NLP) by combining the power of Transformer networks to create rich and contextual language representations (Devlin et al., 2019). BERT development allows the model to understand and generate in-depth text by utilizing context from both directions in a sentence. BERT's key advantage lies in its ability to perform transfer learning at scale, resulting in pre-trained language models that can be adapted to various NLP tasks without requiring complex additional architectures.

BERT has found extensive application in research across a spectrum of NLP languages and tasks, even in languages with rich morphology and those less commonly studied (Koto et al., 2020; Pratama & Rjito, 2021). The successful adaptation of BERT in languages akin to Indonesian and those with similar linguistic features has been well-documented. However, the road to applying BERT to agglutinative languages or those with complex morphological structures is fraught with challenges. The evolution of automated question-and-answer systems, incorporating various deep learning-based methodologies such as LSTM, K-Nearest Neighbor, and of course, BERT (Alotaibi et al., 2020; Nasa-Ngium et al., 2023; Naufal et al., 2023)and education (Jhaerol & Sudianto, 2023; Nguyen et al., 2021), has been rapid. Yet, significant hurdles persist in adapting the question-and-answer system for multilingual use, particularly for languages like Indonesian with unique linguistic characteristics.

Specific research in Indonesian processing highlights the challenges caused by this language's agglutinative nature and rich morphology. Previous efforts have been made to develop appropriate NLP tools, resources, and models for the Indonesian language, but there is still room for further improvement. Studies on question-and-answer systems for Indonesians have included a variety of approaches, including the use of deep learning methods, but consistent evaluation of metrics and techniques needs more attention. Performance comparisons between different models and methodologies using standard evaluation metrics such as accuracy, precision, recall, and F1 scores are crucial to understanding the progress and weaknesses of different approaches (Cahyawijaya et al., 2021; Wilie et al., 2020).

While significant strides have been made, there are gaps in the literature that demand our attention. These gaps include the need for enhanced integration of BERT in the Indonesian context, advancements in multilingual question-and-answer systems, and deeper exploration of pre-trained models for specific languages. This area presents a fertile ground for further research, offering the potential to tackle the remaining challenges in NLP for languages that are less common or possess linguistic features as intricate as Indonesian. The urgency and importance of this research cannot be overstated.

## 3. Material and Method

### 3.1 Dataset

In this study, the Indonesian automatic question-and-answer test uses datasets on existing Independent Learning Independent Campus (MBKM) programs. The data was collected through the MBKM (Dikti, 2020) activities guidelines and the Frequently Asked Questions (FAQ) website. Figure 1 shows the distribution of questions from various possible answers to questions and auto-answers in Indonesian of 600 unique tags and 60 tags. Then, it is made into JSON (JavaScript Object Notation) format.

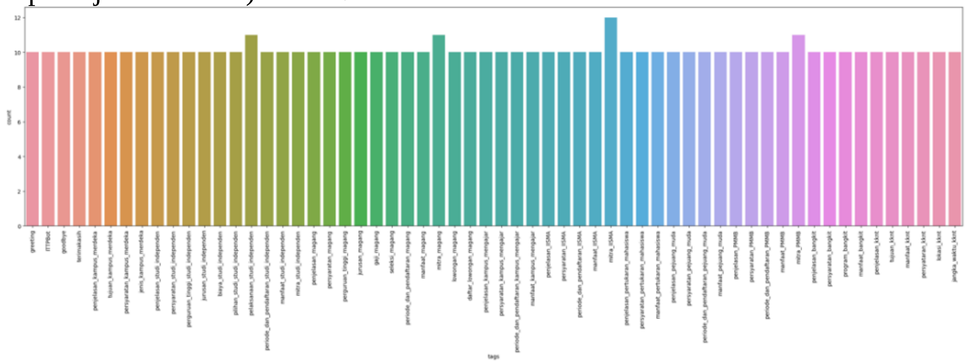

Fig. 1. Distribution of data from label various questions.

### 3.2 Pre-Trained BERT

The pre-training model is learning that has been directed at certain tasks using big data before. The model has learned from a lot of text or data that existed before. The pre-training process involves teaching the model to predict words or parts of sentences based on the surrounding context, meaning it uses the information in the text as a clue for learning (Imamura & Sumita, 2019; Lee & Hsiang, 2020). Pre-trained models can also be utilized as ”Transfer Learning” on more specific tasks, making it possible to train better models even with smaller datasets. An overview of the Transfer Learning process is shown in Figure 2.

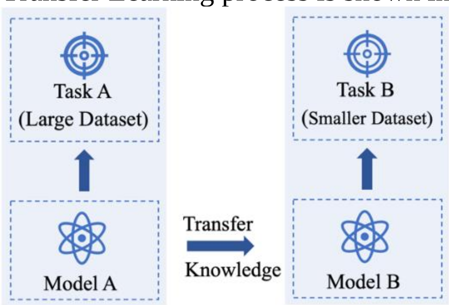


Fig. 2. Transfer Learning process.

One of the pre-trained ones that can be used is BERT (Bidirectional Encoder Representations from Transformers). BERT was built with a vocabulary size of 30.522 from the Book Corpus and Wikipedia datasets. In addition, the learning objective in BERT uses a two-way language modeling or 'masked' language model (Liu et al., 2018). One of the variants of the BERT model is BERT (bert-base-uncased). "Base" refers to the model size, where the BERT-base has a more moderate number of parameters compared to larger versions such as BERT-large. "Uncased" indicates that the model is trained to treat all text in lowercase, regardless of its case or lowercase. BERT models such as bert-base-uncased are very effective in a variety of NLP tasks, including text classification and natural language understanding. "bert-base-uncased" is trained using self-supervised learning techniques utilizing large corpus such as Wikipedia, where the model learns to understand the context of words in sentences by predicting masked language modeling and predicting whether two sentences are sequential or unrelated in text (next sentence prediction). Then, when applying BERT in automated Q&A, BERT is solved, as shown in Figure 3. In Figure 3, specific tasks in BERT include (1) Masked Language Model (MLM), predicting masked words; (2) Next sentence prediction (NSP), checking whether the two input sentences appear sequentially in the text or are not related at all.

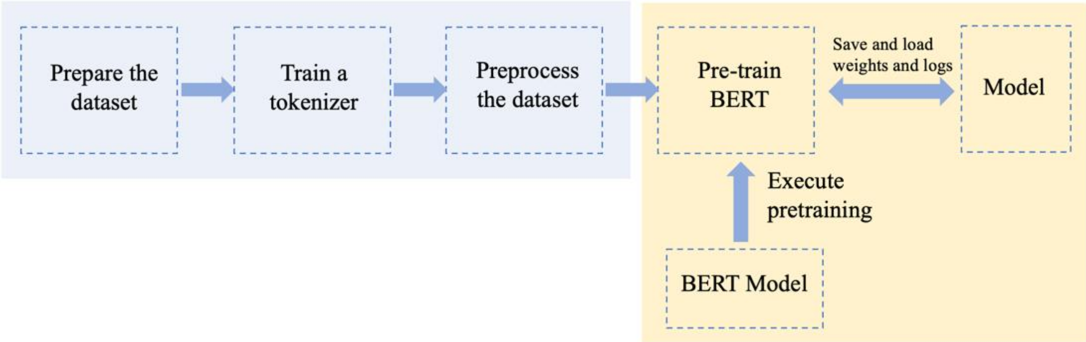


Fig. 3. BERT model pre-trained flow diagram.

### 3.3 Evaluation

In conducting an automatic question-answer performance evaluation, the evaluation of sassiness is carried out by checking similarity. The similarity-checking process goes through the following pseudocode as shown in Figure 4. From pseudocode using the *functions remove stopwords*(), *compute similarity*(), as well as the variable ”question”, which represents a list of strings. Pseudocode captures iteration logic through a list of ”questions”, compares each element with ”text1”, and breaks the loop if a similarity score above the specified threshold is found. Finally, print the values ”text1”, ”text2”, and display the *similarity score*.

```
Pseudocode to check similarity
Set text1 to the result of remove_stopwords ('Question
Context?')
Set a variable i to 0
While i is less than the length of question, do the
following steps:
Set text2 to question[i]
Compute the similarity score between text1 and text2 and
store it in similarity_score
If similarity_score is greater than the desired threshold
(e.g., 0.8), then:
    - Print the value of i
    - Exit the loop
Increment the value of i by 1
Print the values of text1 and text2
Print "Similarity score:" followed by the value of
similarity_score
```

Fig. 4. Pseudocode to check similarity.

## 4. Results and Discussions

Pre-trained models have many advantages. Pre-trained processes contain knowledge from multiple data sources. This process helps in tasks that require language understanding or general knowledge. However, there are situations where training a model from scratch, known as ”training from scratch,” can be more appropriate, especially when the available data is very specific or very different from the general data used for pre-training. This condition is also the basis for deeper study. What if pre-trained is done with a small dataset? So, pretraining needs to be tested on automatic questions and answers in Indonesian. In this research, to build a strong model for question answering, pre-trained BERT was trained with several scenarios. The scenario is built by setting several epoch values, namely 50, 100, and 150, as shown in Table 1.

Table 1 – Scenario obtaining the BERT model.

| Epoch | Training | | Validation | |
|---|---|---|---|---|
| | Accuracy | Loss | Training | Loss |
| 50 | 1 | 0.0249 | 1 | 0.0177 |
| 100 | 1 | 0.0069 | 1 | 0.0048 |
| 150 | 1 | 0.00082 | 1 | 0.00057 |

Table 1 shows the results of the pre-trained BERT model training in the context of automated question-and-answer tasks in Indonesian. This table shows three different training scenarios based on the number of epochs (training iterations): 50, 100, and 150. In a scenario with 50 epochs, the training accuracy reaches 0.0249 with a training loss of 0.0177. Meanwhile, the validation accuracy is 0.0249, with a loss of 0.0177 after one epoch. Then, a scenario with 100 epochs shows a significant performance improvement. The training accuracy dropped to 0.0069 with a training loss of 0.0048 after one epoch, while the validation accuracy and validation loss remained at 0.0069 and 0.0048 respectively. Finally, in a scenario with 150 epochs, the model achieves very high performance with a training accuracy of 0.00082 and a training loss of 0.00057 after one epoch. Validation accuracy and validation losses also show excellent performance with the same number after one epoch. Based on the results of the scenario, the BERT model shows that the more epochs are performed, the better the model performs, which indicates that the model is highly responsive to additional processing and adaptation to the Indonesian dataset for automated question-and-answer tasks.

The training process begins with the initialization of AutoTokenizer, a key component that automates the selection of the appropriate tokenizer for the chosen BERT model, in this case, AutoTokenizer.from pretrained (bert−base−uncased). This is followed by the use of TFAutoModelForSequenceClassification to initialize the BERT model and fine-tune the custom dataset. The role of sequence classification in the TFAutoModelForSequenceClassification model of the BERT architecture is to classify text or sequences into predefined classes or labels. The results of this process, as shown in Table 1, highlight the best loss at the 150th training epoch, with a loss value of 0.00082, and a validation loss of 0.00057. These results form the basis for the subsequent testing of the model's question-answering capabilities.

Table 2 – Question similarity analysis based on dataset.

| Similarity Condition (%) | Label | Question | Similarity Score |
|---|---|---|---|
| >80 | Jenis Kampus Merdeka | jenis program kampus merdeka? | 0.99 |
| >80 | Biaya Studi Independen | mengikuti program studi independen membayar biaya studi? | 0.98 |
| >80 | Persyaratan Magang | persyaratan magang bersertifikat? | 0.99 |
| <20 | Jenis kampus merdeka | mahasiswa mengikuti program studi independen? | 0.01 |
| <20 | Biaya Studi Independen | berapa studi independen? | 0.01 |
| <20 | Persyaratan Magang | Hai! | 0.08 |

Table 3 – Question similarity analysis based outside dataset.

| Similarity Condition (%) | Label | Question | Similarity Score |
|---|---|---|---|
| >50 | Pengertian | ittpbot? | 0.53 |

| | | | |
|---|---|---|---|
| | Machine Learning | | |
| >50 | Perbedaan antara program kampus merdeka dengan pendidikan konvensional | jenis program kampus merdeka? | 0.91 |
| >50 | Peran teknologi dalam mendukung pelaksanaan kampus merdeka | kampus merdeka? | 0.72 |
| <50 | Pengertian Machine Learning | persyaratan studi independen? | 0.11 |
| <50 | Perbedaan antara program kampus merdeka dengan pendidikan konvensional | mahasiswa mengikuti program studi independen? | 0.06 |
| <50 | Peran teknologi dalam mendukung pelaksanaan kampus merdeka Magang | tujuan kampus merdeka | 0.09 |

The reliability of the model, trained with the best loss, is tested in two question similarity scenarios. These scenarios involve questions within the dataset's scope and questions outside the dataset's scope. The similarity of question-and-answer texts is assessed using BERT, as shown in Tables 2 and 3. These tables demonstrate how the similarity or relevance of the questions to the labels is evaluated based on the similarity score. For instance, table 2 reveals that questions with a similarity level above 80% with their labels are highly relevant. A prime example is the question about the 'type of independent campus program?', which has a similarity score of 0.99, indicating a perfect fit with the label 'Type of Independent Campus', as illustrated in Figure 5.

On the other hand, questions that had a similarity rate of less than 20% with their labels indicated that these questions had low or no relevance to the topic. For example, the question about "students participating in independent study programs?" has a similarity score of only 0.01 against the label "Independent Campus Type", indicating that this question is irrelevant or inappropriate for the topic. Table 3 provides a similar analysis but with a slightly lower similarity threshold, above 50% for high relevance and below 50% for low relevance. For example, the question about "types of independent campus programs?" has a similarity score of 0.91 against the label "Difference between independent campus programs and conventional education", indicating a high relevance between the question and the topic, as shown in Figure 6.

Overall, this illustrates that the similarity score can be used as an indicator of reliability in assessing the relevance between the question and the specified topic or label. A high score indicates a highly relevant question, while a low score indicates a mismatch between the question and the topic. This analysis is essential to ensure that the question-and-answer system or information search can provide accurate answers and match the questions asked. So, from testing both, the performance of both shows: (1) Accuracy on the dataset, automatic questions, and using BERT shows good ability in understanding and responding to questions in the dataset. BERT is a powerful language model and can perform better context understanding; (2) Limitations of existing knowledge: Although automatic question and answer using BERT can provide accurate answers in the dataset, the main limitation lies in the knowledge it has, as seen in Figure 6. Suppose in Figure 6, the user asks a question outside the scope of the data set. In this case, the automated Q&A may need the relevant information to provide the correct answer. So, question-and-answer automation using BERT has good capabilities in answering questions on the dataset but is limited to the knowledge that has been provided.

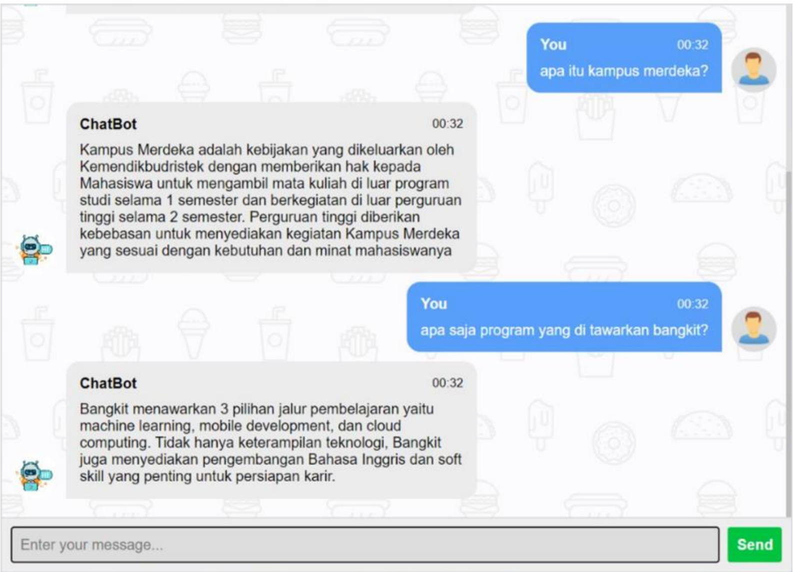


Fig. 5. Automatic question-answer with BERT pre-train using Indonesian.

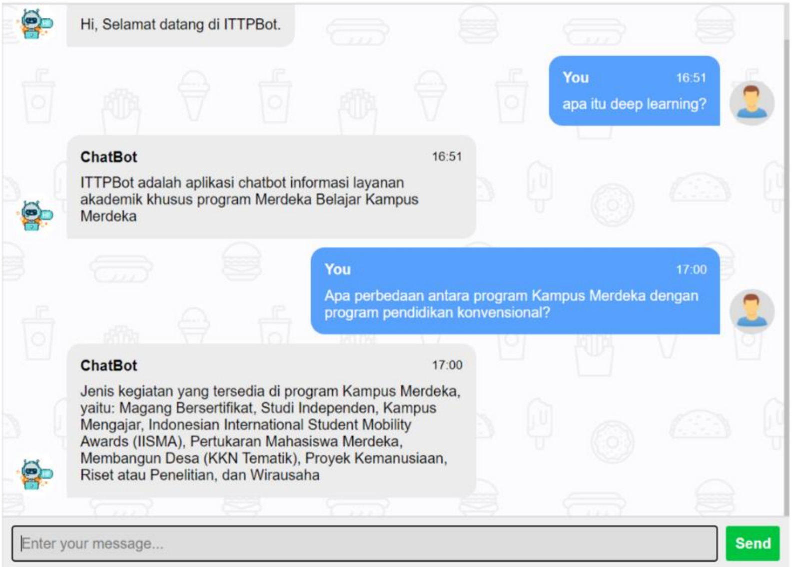


Fig. 6. Similarity checking outside the context of the dataset.

Contrast with the approach of using BERT in an automated question-and-answer system with an Indonesian dataset that utilizes LSTM (Jhaerol & Sudianto, 2023) and other Indonesian chatbot implementations (Lasama et al., 2024; Lubis & Sumartono, 2023), BERT overcomes the challenge of 'long-range dependencies' in text. This means the model can consider information from an entire sentence or document simultaneously, rather than focusing solely on data sequences as LSTMs do, potentially overlooking the broader context. This unique capability of BERT allows it to provide more relevant and accurate answers by considering a more in-depth context. Moreover, BERT's versatility is not limited to the sequence of questions; it can also handle more complex or diverse questions, including those that require a deep understanding of content from data sources such as MBKM. This broad applicability of BERT opens up the possibility of answering questions in various contexts and aspects. In summary, using BERT in automated question-and-answer systems significantly contributes by expanding the scope and depth of context understanding in answering questions. These results not only enhance the quality and relevance of answers but also present a more powerful chatbot in processing natural language, especially in domains that require a deep understanding of context such as in the utilization of MBKM data, thereby underlining the model's potential in your research.

## 5. Conclusion

The BERT model adapted for question-and-answer tasks in Indonesian succeeded in providing good performance in answering questions. The best result is a training loss value of 0.00082, then a validation loss of 0.00057 after 150 epochs through a comprehensive adjustment approach and training process. Further evaluation revealed that the BERT model could provide answers based on the knowledge contained in the dataset, achieving an average concordance above 80%, which shows its potential use in various fields. Although the results of this research are positive, several suggestions can be applied for further development. Further exploration of BERT's various architectural and model parameters can significantly improve performance. Then, expanding the training data set with various questions and contexts will enrich the model's understanding of the complex Indonesian language. Using data augmentation techniques can also increase the diversity of the data set.

**Acknowledgement**
Would like to thank LPPM Institut Teknologi Telkom Purwokerto for supporting this research.

## References

Alotaibi, R., Ali, A., Alharthi, H., & Almehamdi, R. (2020). AI Chatbot for Tourist Recommendations: A Case Study in the City of Jeddah, Saudi Arabia. *International Journal of Interactive Mobile Technologies (iJIM)*, *14*(19), Article 19. https://doi.org/10.3991/ijim.v14i19.17201

Arnumukti, M. L., Sudianto, S., & Athiyah, U. (2023). Product Layout Recommendations based on Customer Behavior and Data Mining. *2023 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT)*, 330–334. https://doi.org/10.1109/COMNETSAT59769.2023.10420563

Cahyawijaya, S., Winata, G. I., Wilie, B., Vincentio, K., Li, X., Kuncoro, A., Ruder, S., Lim, Z. Y., Bahar, S., Khodra, M. L., Purwarianti, A., & Fung, P. (2021). *IndoNLG: Benchmark and Resources for Evaluating Indonesian Natural Language Generation*.

Clark, J. H., Choi, E., Collins, M., Garrette, D., Kwiatkowski, T., Nikolaev, V., & Palomaki, J. (2020). TYDI QA: A Benchmark for Information-Seeking Question Answering in in Typologically Diverse Languages. *Transactions of the Association for Computational Linguistics*, *8*, 454–470. https://doi.org/10.1162/tacl_a_00317

Cortes, E. G., Woloszyn, V., Barone, D., Möller, S., & Vieira, R. (2022). A systematic review of question answering systems for non-factoid questions. *Journal of Intelligent Information Systems*, *58*(3), 453–480. https://doi.org/10.1007/s10844-021-00655-8

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* (arXiv:1810.04805). arXiv. https://doi.org/10.48550/arXiv.1810.04805

Dikti, K. M. (2020). *Buku Panduan Merdeka Belajar—Kampus Merdeka*.

Galassi, A., Lippi, M., & Torroni, P. (2021). Attention in Natural Language Processing. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(10), 4291–4308. https://doi.org/10.1109/TNNLS.2020.3019893

Imamura, K., & Sumita, E. (2019). Recycling a Pre-trained BERT Encoder for Neural Machine Translation. In A. Birch, A. Finch, H. Hayashi, I. Konstas, T. Luong, G. Neubig, Y. Oda, & K. Sudoh (Eds.), *Proceedings of the 3rd Workshop on Neural Generation and Translation* (pp. 23–31). Association for Computational Linguistics. https://doi.org/10.18653/v1/D19-5603

Indah Rahajeng, M., & Purwarianti, A. (2021). Indonesian Question Answering System for Factoid Questions using Face Beauty Products Knowledge Graph. *Jurnal Linguistik Komputasional (JLK)*, *4*(2), 59. https://doi.org/10.26418/jlk.v4i2.62

Jhaerol, M. R., & Sudianto, S. (2023). Implementation of Chatbot for Merdeka Belajar Kampus Merdeka Program using Long Short-Term Memory. *Jurnal Nasional Pendidikan Teknik Informatika : JANAPATI*, *12*(2), Article 2. https://doi.org/10.23887/janapati.v12i2.58794

Koto, F., Rahimi, A., Lau, J. H., & Baldwin, T. (2020). IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP. *Proceedings of* *the 28th International Conference on Computational Linguistics*, 757–770. https://doi.org/10.18653/v1/2020.coling-main.66

Kovaleva, O., Romanov, A., Rogers, A., & Rumshisky, A. (2019). *Revealing the Dark Secrets of BERT* (arXiv:1908.08593). arXiv. http://arxiv.org/abs/1908.08593

Lasama, J., Sudianto, S., Ramadhani, R., Hilmawan, M. D., Aldean, M. Y., & Satria, M. A. H. (2024). English Indonesia-Chan: OPUS-MT Powered Chatbot. *Jurnal Teknik Elektro dan Komputasi (ELKOM)*, *6*(1), 105–111. https://doi.org/10.32528/elkom.v6i1.18613

Lee, J.-S., & Hsiang, J. (2020). Patent classification by fine-tuning BERT language model. *World Patent Information*, *61*, 101965. https://doi.org/10.1016/j.wpi.2020.101965

Liu, P. J., Saleh, M., Pot, E., Goodrich, B., Sepassi, R., Kaiser, L., & Shazeer, N. (2018). *Generating Wikipedia by Summarizing Long Sequences* (arXiv:1801.10198). arXiv. https://doi.org/10.48550/arXiv.1801.10198

Lubis, A., & Sumartono, I. (2023). *Implementasi Layanan Akademik Berbasis Chatbot untuk Meningkatkan Interaksi Mahasiswa*. *3*(5).

Nasa-Ngium, P., Nuankaew, W. S., & Nuankaew, P. (2023). Analyzing and Tracking Student Educational Program Interests on Social Media with Chatbots Platform and Text Analytics. *International Journal of Interactive Mobile Technologies (iJIM)*, *17*(05), Article 05. https://doi.org/10.3991/ijim.v17i05.31593

Naufal, A. B., Sudianto, S., & Fachri, M. A. A. (2023). *Implementation of Chatbot System on Tourism Objects in Banyumas Regency with AIML and Chatterbot*. *5*.

Nguyen, T. T., Le, A. D., Hoang, H. T., & Nguyen, T. (2021). NEU-chatbot: Chatbot for admission of National Economics University. *Computers and Education: Artificial Intelligence*, *2*. https://doi.org/10.1016/j.caeai.2021.100036

Pratama, T., & Rjito, S. (2021). IndoXLNet: Pre-Trained Language Model for Bahasa Indonesia. *International Journal of Engineering Trends and Technology*, *70*(5), 367–381. https://doi.org/10.14445/22315381/IJETT-V70I5P240

Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, *63*(12), 54–63. https://doi.org/10.1145/3381831

Sudianto, Herdiyeni, Y., & Prasetyo, L. B. (2023a). Machine learning for sugarcane mapping based on segmentation in cloud platform. *AIP Conference Proceedings*, *2482*(1), 020001. https://doi.org/10.1063/5.0132180

Sudianto, S., Herdiyeni, Y., & Prasetyo, L. B. (2023b). Classification of Sugarcane Area Using Landsat 8 and Random Forest based on Phenology Knowledge. *JOIV : International Journal on Informatics Visualization*, *7*(3–2), Article 3–2. https://doi.org/10.30630/joiv.7.3-2.1401

Sudianto, S., Herdiyeni, Y., & Prasetyo, L. B. (2023c). Early Warning for Sugarcane Growth using Phenology-Based Remote Sensing by Region. *International Journal of Advanced Computer Science and Applications*, *14*(2). https://doi.org/10.14569/IJACSA.2023.0140259

Wilie, B., Vincentio, K., Winata, G. I., Cahyawijaya, S., Li, X., Lim, Z. Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S., & Purwarianti, A. (2020). *IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding*.