ILKOM Jurnal Ilmiah Vol. 15, No. 1, April 2023, pp.124-131
Accredited 2nd by RISTEKBRIN No. 200/M/KPT/2020; E-ISSN 2548-7779 | P-ISSN 2087-1716

Research Article

Abstractive text summarization using Pre-Trained Language model "Text-to-Text Transfer Transformer (T5)"

Qurrota A'yuna Itsnaini [a,1]; Mardhiya Hayaty [a,2,*]; Andriyan Dwi Putra [a,3]; Nidal A.M Jabari [b,4]
[a] *Universitas Amikom Yogyakarta, Jl. Ring Road Utara, Yogyakarta and 55281, Indonesia*
[b] *Palestine technical university Kadoorie, , Palestine*
[1] *qurrota.itsnaini@students.amikom.ac.id;* [2] *mardhiya_hayati@amikom.ac.id;* [2] *andriyan.Putra@amikom.ac.id;* [4] *nidal.jabari@ptuk.edu.ps*
* *Corresponding author*

Article history: Received November 29, 2022; Revised December 02, 2022; Accepted January 21, 2023; Available online April 07, 2023

Abstract

Automatic Text Summarization (ATS) is one of the utilizations of technological sophistication in terms of text processing assisting humans in producing a summary or key points of a document in large quantities. We use Indonesian language as objects because there are few resources in NLP research using Indonesian language. This paper utilized PLTMs (Pre-Trained Language Models) from the transformer architecture, namely T5 (Text-to-Text Transfer Transformer) which has been completed previously with a larger dataset. Evaluation in this study was measured through comparison of the ROUGE (Recall-Oriented Understudy for Gisting Evaluation) calculation results between the reference summary and the model summary. The experiments with the pre-trained t5-base model with fine tuning parameters of 220M for the Indonesian news dataset yielded relatively high ROUGE values, namely ROUGE-1 = 0.68, ROUGE-2 = 0.61, and ROUGE-L = 0.65. The evaluation value worked well, but the resulting model has not achieved satisfactory results because in terms of abstraction, the model did not work optimally. We also found several errors in the reference summary in the dataset used.

Keywords: Automatic Text Summarization; Transformer; Pre-Trained Model; T5; ROUGE

Introduction

Automatic text summarization (ATS) is a form of text processing in the field of Natural Language Processing that utilizes the sophistication of the internet and technology to summarize text automatically[1]. ATS functions to produce a clear summary while maintaining the main information and the overall meaning contained therein [2]. Text summarization is grouped into two methods, extractive and abstractive. Extractive text summarization uses the calculation of the score of words in sentences [3]. At the same time, the summary text abstraction aims to produce a summary by interpreting and analyzing the text as a whole so that it has less text but still contains the essential information conveyed in the original text [2].

One of the studies was carried out using the Sequence-to-Sequence method with enhanced features for single documents [4]. It used a non-local network feature that functions to improve the traditional Sequence-to-Sequence structure, thus proving that the model proposed in this study has more effective results compared to the basic model, with an increase of 5.6%, 5.3%, 6.2% in three ROUGE matrix of R-1, R-2, and R-L values.

Another research was carried out using the transformer method and datasets from 'Wikipedia' and 'The Hindu' [5]. The data was processed using the Bidirectional Encoder Representation Transformer (BERT) model, which produced ROUGE-1, ROUGE-2, and ROUGE-L values of 41.72, 19.39, and 38.76, respectively.

Based on our observation, transformer architecture has a better effect than other methods in summarizing large-scale texts. Transformer architecture performs text summarization automatically using self-attention to calculate different input and output representations [6]. For this reason, we would implement a transformer architecture using PTLMs (Pre-Trained Language Models) that have been trained with large-scale data, namely the T5 (Text-to-Text

https://doi.org/10.33096/ilkom.v1xix.xxx.x-x

Transfer Transformer) model, into Indonesian language documents. This study aimed to prove the effectiveness of the T5 pre-trained model in ATS of large-scale Indonesian online news data.

Method

A. Automatic Text Summarization (ATS)

The development of text summarization has been improving, it can be seen from the increasing number of research discussing ATS. ATS is divided into two methods, extractive and abstractive. Research on extractive forms using the Aspect Based Sentiment Analysis (ABSA) process was carried out using the SumEval dataset[7]. The data was extracted through the self-attention stage, combined with the original sentences that were previously extracted through Convolutional Neural Networks (CNN). In the final stage, the aspect-category or aspect-term data was connected to obtain the feature sentiment results. The results on the aspect category with the Gated CNN self-attention model yielded scores of 81.40 and 79.77. Meanwhile, the aspect term with the same model produced 71.96 and 62.54.

The research was carried out in an abstract form using the sequence-to-sequence method of summarizing English social media texts [8]. An intervention was conducted to the encoder to filter information better. The Evaluation acores of the F1 ROUGE-1, ROUGE-2, and ROUGE-L were 2.6%, 2.1%, and 2.5%, respectively. These results showed that the selective method given to the encoder worked optimally. The transformer approach for Indonesian-language datasets was used to detect rude comments on online news in Indonesia[9]. The model used was the BERT model and the BERT Multilingual pre-train to serve as a baseline, so that the results from the Scratch model that were trained obtained an F-1 average score of 50%, which was then compared to the Multilingual BERT model of 54%.

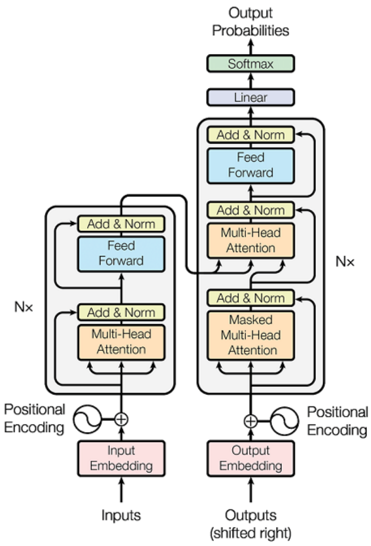

Figure 1. Transformer architecture with an attention mechanism

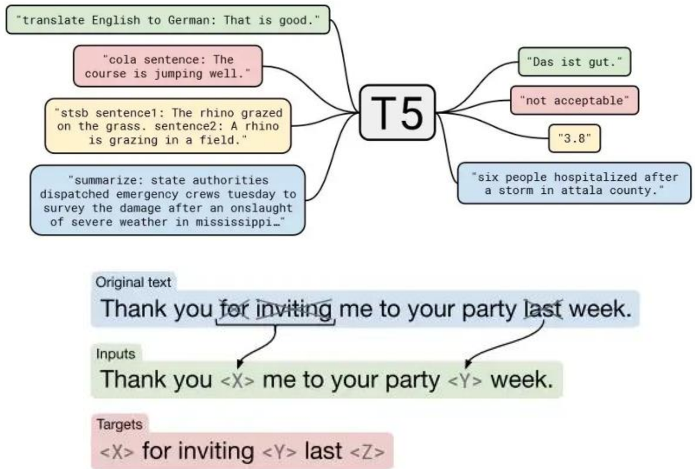

Figure 2. How the T5 models work

Transformer. Transformer architecture changed the sequence using two necessary essential encoders and a decoder. The function of the encoder was to capture the input sequence and process information with a fixed length and sequence, after that the decoder displayed the output previously processed on the encoder, shown in **Figure 1**. Both of these parts could work optimally when using the attention mechanism (Attention Mechanism) [10][11]. The attention Mechanism functioned to connect the encoder and decoder in processing the information provided by giving attention to long input sentences [12]. So that information can be disseminated to all sequences of input sentences which were carried out selectively by the decoder with an attention mechanism [13].

The research was conducted to summarize news about COVID-19 using a transformer architecture [14] by utilizing the encoder-decoder model and obtaining ROUGE-1 and ROUGE-2 scores of 0.58 and 0.42, with a training time of 11438 seconds, respectively.

Text-to-Text Transfer Transformer (T5). The T5 model is a framework developed on top of popular architectures such as BERT, GPT, etc., by utilizing text-to-text transfer learning, as shown in **Figure 2**, with examples of translation tasks, text similarities, and text summaries. T5 is an end-to-end trained transformer model with text as input and modified text as output[15]. In its completion, the T5 model proved successful and capable of performing three tasks [16]: generating, classification, and regression.

The application of the T5 model to the BBC News dataset to determine which model is superior to the ROUGE evaluation value. Research comparing several transformer models [3], one of which is the T5 model, which is proven to be superior to other transformer models. The score of ROUGE-1, ROUGE-2, and ROUGE-L are 0.47, 0.33, and 0.42 respectively.

Transformers, a tool or platform for downloading and training pre-trained models, has been trained on previous datasets using Hugging Face[17]. Pre-training of models has been carried out on larger corpus or datasets to improve model performance[18]. The pre-trained Model T5 has several types according to the size of the parameters [15], including :

- **T5-Small** **:** 60 M of parameters
- **T5-Base** **:** 220 M of parameters
- **T5-Large** **:** 770 M of parameters
- **T5-3B** **:** 3 B of parameters
- **T5-11B** **:** 11 B of parameters

Models with smaller parameters can help train data with limited hardware specifications. The pre-training model used in this study uses the "cahya/t5-base-indonesian-summarization-cased" model for Indonesian texts with fine-tuning on the id_liputan6 dataset. Specifically, the T5-base model trains a model with a pre-trained text-to-text format with parameters of 220 M [15] [19]. T5-base is applied to the encoder in the form of Indonesian news text.

B. Experiment

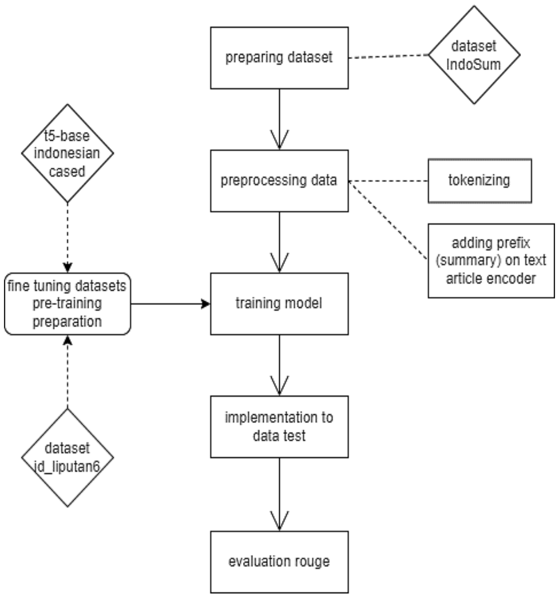

Figure 3. Research Flowchart Diagram

As part of the research, we outline the research flow in **Figure 3** by pre-processing the text of the article in the encoder by adding the **"summarize:"** prefix to facilitate the task intended by the input. The hyperparameters used in this research adjust to the hardware resources' limitations. We implement tokenizing with a maximum of 512 words for each encoder and 128 words for the decoder.

The hardware specifications used can be described as follows:

- Python 3.6 and up
- Transformers dan PyTorch
- Python Machine Learning base libraries
- High RAM GPU settings

1) *Preparing Dataset*

This study only used 9,387 data from IndoSum's 18,774 Indonesian language news. This was due to our limited resources in conducting this research. But this did not affect the evaluation process in modelling. So, the research objectives could still be achieved. We divide the data into training data (90% = 8448) and testing data (10% = 939).

To validate during training, we divided 10% of the training data, namely 845 data, for data validation. IndoSum's news data was taken from major Indonesian-language news portals, Kumparan and CNN Indonesia [20].

2) *ROUGE Evaluation*

The evaluation used in this research was the ROUGE (Recall-Oriented Understudy for Gisting Evaluation) matrix[21]. Rouge is an evaluation tool used for multi-document summarization and has great advantage in text summarization [22]. Rouge generally compares the human summary as a model with the machine summary. In this case, the evaluation used in the current study was the number of unigrams for ROUGE-1, namely between the system summary and the reference, the number of bigrams for ROUGE-2 between the system and the reference summary, and ROUGE-L by comparing the Longest Common Subsequence (LCS) between the summary results machine with summary reference to the dataset. In equation (1), it was explained that the value of LCS or the longest word compared to M was the number of words in the reference summary [23].

$$ROUGE - L = \frac{\text{LCS}}{\text{M}} \tag{1}$$

Each ROUGE value included the Recall (2), Precision (3), and F1-Score (4) matrices.

$$Recall = \frac{\text{W overlapped}}{\text{W ref}} \tag{2}$$

$$Precision = \frac{\text{W overlapped}}{\text{W cand}} \tag{3}$$

$$F1 - Score = \frac{2\ x\ Recall\ x\ Precision}{(Recall + Precision)} \tag{4}$$

Results and Discussion

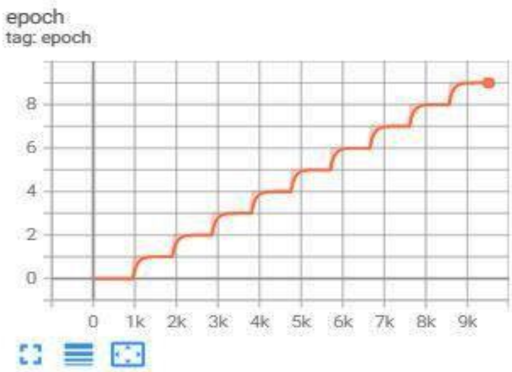

Figure 4. The number of steps across the epochs

The present research used a batch size set to 8 because of the limited computational unit during training. **Figure 4** shows the number of epochs carried out during training, namely 10 (y labels) with a total of 9,509 steps (x labels) from 8448 Indonesian news data.

After training, the model was tested using a dataset test to determine the model’s success in understanding news sentences in the training data. The experimental summary results on the test data was shown in **Table 1**.

Table 1. Example of Summarized Results on Data Test

Article:

Jakarta CNN Indonesia - - Anthony Martial menunjukkan aksi individu yang luar biasa saat **Manchester United menang atas Real Madrid lewat drama adu penalti dalam laga pramusim di California Amerika Serikat** . Gol Manchester United di laga tersebut dicetak oleh Jese Lingard di masa injury time babak pertama . Namun Martial berperan besar di balik terciptanya gol tersebut . Setelah menerima sodoran bola dari Marouane Fellaini Martial menggiring bola sambil melakukan penetrasi ke kotak penalti Real Madrid . Martial dibayangi oleh Dani Carvajal dengan ketat ditambah pengawasan dari Luka Modric. Namun Martial mampu meliuk-liuk dengan perpindahan bola yang cepat dari kaki kanan ke kaki kiri berulang kali . Setelah Carvajal terkecoh Raphael Varane coba membantu . Tetapi Martial dengan dingin mengirim bola pada Lingard yang ada di muka gawang . Lingard pun tanpa kesulitan menceploskan bola ke dalam gawang . Gol Lingard ini sendiri sukses dibalas Madrid lewat penalti Casemiro di babak kedua . Tetapi pada akhirnya Manchester United menang lewat drama adu penalti . Dalam drama adu penalti dua algojo Manchester United Henrikh Mkhitaryan dan Daley Blind berhasil melaksanakan tugasnya sedangkan di kubu Los Blancos hanya Luismi Quezada yang berhasil menjadi eksekutor . **Kemenangan "" Setan Merah "" atas Los Blancos** membuat tren positif skuat asuhan Jose Mourinho di tur pramusim berlanjut . Ini adalah kemenangan keempat Manchester United di tur pramusim setelah sebelumnya mereka menaklukkan LA Galaxy (5 - 2) Real Salt Lake (2 - 1) dan Manchester City (2 - 0) .

Jakarta CNN Indonesia - Anthony Martial showed extraordinary individual performance when Manchester United won over Real Madrid via a penalty shoot-out in a pre-season match in California, United States of America. Jese Lingard scored Manchester United's goal in the match in the first half injury time. But Martial played a big role behind the creation of these goals. After receiving a ball from Marouane Fellaini, Martial dribbled the ball while penetrating the Real Madrid penalty box. Martial was being closely shadowed by Dani Carvajal plus supervision from Luka Modric. However, Martial was able to weave the ball quickly from right to left foot repeatedly. After Carvajal was tricked Raphael Varane tried to help. But Martial coolly sent the ball to Lingard who was in front of the goal. Lingard also without difficulty struck the ball into the goal. Madrid successfully answered Lingard's goal itself through Casemiro's penalty in the second half. But in the end Manchester United won via penalty shootout . In the penalty shoot-out, two Manchester United executors Henrikh Mkhitaryan and Daley Blind managed to carry out their duties while in the Los Blancos camp only Luismi Quezada managed to become the executor. The victory of ""Red Devils"" over Los Blancos has made the positive trend of Jose Mourinho's squad on the pre-season tour continue. This was Manchester United's fourth win on the pre-season tour after they previously beat LA Galaxy (5-2) Real Salt Lake (2-1) and Manchester City (2-0).

Reference Summary:

Anthony Martial menunjukkan aksi individu yang luar biasa saat **Manchester United menang atas Real Madrid lewat drama adu penalti dalam laga pramusim di California Amerika Serikat** . Gol Manchester United di laga tersebut dicetak oleh Jese Lingard di masa injury time babak pertama . Namun Martial berperan besar di balik terciptanya gol tersebut .

Anthony Martial showed extraordinary individual performance when Manchester United won over Real Madrid via a penalty shoot-out in a pre-season match in California, United States of America. Jese Lingard scored Manchester United's goal in the match in the first half injury time. But Martial played a big role behind the creation of these goals.

Model Summary :

Kemenangan atas Real Madrid lewat drama adu penalti dalam laga pramusim di California Amerika Serikat membuat tren positif skuat asuhan Jose Mourinho di tur pramusim berlanjut. Ini adalah kemenangan keempat Manchester United di tur pramusim setelah sebelumnya mereka menaklukkan LA Galaxy (5 - 2) Real Salt Lake (2 - 1) dan Manchester City (2 - 0).

The victory over Real Madrid via a penalty shootout in a pre-season match in California, United States of America, has made the positive trend of Jose Mourinho's squad continue on the pre-season tour. This is Manchester United's fourth win on the pre-season tour after previously they beat LA Galaxy (5 - 2) Real Salt Lake (2 - 1) and Manchester City (2 - 0).

Table 2. Results of the ROUGE Evaluation Value from **Table. 1**

	Rouge-1	**Rouge-2**	**Rouge-3**
Recall	0.4	0.28	0.4
Precision	0.36	0.25	0.36
F1-Score	0.38	0.27	0.8

Results of the ROUGE Evaluation Value from **Table. 1** see in **Table 2** Several errors were found in the summary of the news text above, including:

- Paraphrasing produced by the model was not optimal. An example of a successful paraphrase was **"Kemenangan"** at the beginning of a predicted summary sentence. **"Kemenangan"** was a paraphrase of the sentence **"Manchester United menang atas Real Madrid".** In addition, the model summary results were only in the form of main sentences that exactly match the encoder. This might be due to the absence of Word Embedding before training. So, the model cannot read words that had the same context.
- The model managed to read and save the meaning of the sentence **"Kemenangan Setan Merah atas Blancos"** as the meaning of **"Kemenangan Manchester United atas Real Madrid".**
- In the text of the article, some words that were not labelled were omitted and randomly removed. The model only took the main sentence at the beginning and end of the paragraph.
- Score Recall on ROUGE-L in Table. II showed the number 0.4 of the calculation results of the LCS or the longest word in the engine resulting in summary, namely the predicted summary compared to the summary of the dataset. The results were low because the difference between the two was noticeable.

We also found several errors in the dataset, where the summary reference of the dataset was only in the form of rewritten main sentences; no changes to words or paraphrases were made from human summaries. An example of the dataset is shown in **Table 3**.

Table 3. Examples of Errors in the Dataset

Article :

Merdeka.com - Mantan manajer legendaris Manchester United Sir Alex Ferguson menyebut agen Paul Pogba Mino Raiola sebagai seorang tukang bohong . Kecaman itu sendiri tidak lepas dari hubungan buruk Ferguson dengan pria Italia tersebut . Kedua orang ini jadi tidak akur akibat transfer Pogba . Gelandang asal Prancis itu dulu memperkuat MU sebelum akhirnya lari ke Juventus pada tahun 2012 . Parahnya ia lari dengan status bebas transfer . Ferguson sendiri pernah menyebut Raiola sebagai salah satu orang yang tidak disukainya . Dan ketika ditanya terkait apa yang terjadi pada tahun 2012 silam Ferguson pun tidak tahan untuk tidak mencela Raila dan menyalahkannya atas kepergian Pogba ke Turin lima tahun lalu . " Paul Pogba ? Ia memiliki agen yang buruk . Seorang tukang bohong " ketus Ferguson seperti dilansir Sportsmole . " Paul Pogba sudah kami kenal dengan baik . Kami tahu ia pemain bagus ia masih pemain yang bagus " tegasnya . " kami menawarinya kontrak terbaik " sambungnya .

Merdeka.com - Legendary Manchester United ex-manager Sir Alex Ferguson has accused Paul Pogba's agent Mino Raiola of being a liar. The criticism itself is inseparable from Ferguson's bad relationship with the Italian man. These two people don't get along because of the Pogba transfer. The midfielder from France used to strengthen MU before finally moving to Juventus in 2012. To make matters worse, he ran on a free transfer. Ferguson himself once referred to Raiola as one of the people he did not like. And when asked about what happened in 2012, Ferguson could not help but criticize Raila and blame him for Pogba's departure for Turin five years ago. " Paul Pogba ? He has a bad agent . A liar " said Ferguson as reported by Sportsmole . " We know Paul Pogba well . We know he is a good player he is still a good player " he said . " We offered him the best possible contract " he continued

Reference Summary:

Mantan manajer legendaris Manchester United Sir Alex Ferguson menyebut agen Paul Pogba Mino Raiola sebagai seorang tukang bohong . Kecaman itu sendiri tidak lepas dari hubungan buruk Ferguson dengan pria Italia tersebut . Kedua orang ini jadi tidak akur akibat transfer Pogba . Gelandang asal Prancis itu dulu memperkuat MU sebelum akhirnya lari ke Juventus pada tahun 2012 . Parahnya ia lari dengan status bebas transfer .

Legendary Manchester United exmanager Sir Alex Ferguson has accused Paul Pogba's agent Mino Raiola of being a liar. The criticism itself is inseparable from Ferguson's bad relationship with the Italian man. These two people don't get along because of the Pogba transfer. The midfielder from France used to strengthen MU before finally moving to Juventus in 2012. To make matters worse, he ran on a free transfer.

The reference summary of the dataset greatly influences the ROUGE evaluation value generated. It can be seen clearly in **Table 3** that the reference summary only contained the repetition of the first sentence of the paragraph. There was not a single change or layout for each word. This was an important thing to consider before doing research.

Table 4. Rouge Value of Overall Data Test

	R-1	R-2	R-L
RECALL	0.686	0.612	0.654
PRECISION	0.741	0.662	0.707
F-1 SCORE	0.71	0.633	0.677

After calculating the ROUGE value for each line of test data, which amounts to 939, we get the overall ROUGE value taken based on the average value per line. Based on the **Table 4**, the score value can be stated as high. However, errors in the dataset and model results were also essential in calculating this evaluation. We assumed the high results were generated because some reference summaries only take the paragraph's main sentences. Meanwhile, the model produced in this study can only paraphrase a few of the many words in the news article, so several summary models were also generated from the main sentences of the paragraph.

Conclusion

Model T5 (Text-to-Text Transfer Transformer) is the newest PLTM that requires highly qualified research resources. The pretrained T5 model resulting from this study obtained an average evaluation value of ROUGE-1 of 0.68, ROUGE-2 of 0.61, and ROUGE-L of 0.65. This showed that the model succeeds in paraphrasing sentences even though only a few and not optimally; at least the model did not change the meaning of the original article so that the summary was easy to understand without having to read the original article.

For further research, some suggestions that can be done and even prepared are:

- Using datasets that have more data and maximum quality reference summaries. With more data, it is hoped that the model can store much vocabulary to be later implemented on the test data with paraphrases that still make sense and don't change the meaning of the context in it.
- Doing word embedding after preprocessing. It is meant to convert the word into a vector and save it to see whether the next word has the same context. Then the paraphrase will work more optimally.
- Using a pre-trained model with more extensive parameters, such as T5-Large and some above.
- Adding batch size and epoch training to produce better model results.

References

[1] A. P. Widyassari *et al.*, 'Review of automatic text summarization techniques & methods', *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 4, pp. 1029–1046, 2022, doi: 10.1016/j.jksuci.2020.05.006.

[2] M. Allahyari *et al.*, 'Text summarization techniques: A Brief Survey', *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 10, 2017, doi: 10.14569/ijacsa.2017.081052.

[3] A. Gupta, D. Chugh, Anjum, and R. Katarya, 'Automated news summarization using Transformers', *Lect. Notes Electr. Eng.*, vol. 840, pp. 249–259, 2022, doi: 10.1007/978-981-16-9012-9_21.

[4] Z. Hao and B. Xue, '2020 5th Asia-Pacific Conference on Intelligent Robot Systems, ACIRS 2020', *2020 5th Asia-Pacific Conf. Intell. Robot Syst. ACIRS 2020*, pp. 163–167, 2020.

[5] M. Ramina, N. Darnay, C. Ludbe, and A. Dhruv, 'Topic level summary generation using BERT induced Abstractive Summarization model', *Proc. Int. Conf. Intell. Comput. Control Syst. ICICCS 2020*, no. Iciccs, pp. 747–752, 2020, doi: 10.1109/ICICCS48265.2020.9120997.

[6] A. Vaswani *et al.*, 'Attention is all you need', *Adv. Neural Inf. Process. Syst.*, vol. 2017-Decem, no. Nips, pp. 5999–6009, 2017.

[7] J. Yang and J. Yang, 'Aspect based sentiment analysis with Self-Attention and Gated Convolutional Networks', *Proc. IEEE Int. Conf. Softw. Eng. Serv. Sci. ICSESS*, vol. 2020-Octob, pp. 146–149, 2020, doi: 10.1109/ICSESS49938.2020.9237640.

[8] Z. Liang, J. Du, and C. Li, 'Abstractive social media text summarization using selective reinforced Seq2Seq attention model', *Neurocomputing*, vol. 410, pp. 432–440, 2020, doi: 10.1016/j.neucom.2020.04.137.

[9] A. D. Rendragraha, M. A. Bijaksana, and A. Romadhony, 'Pendekatan metode Transformers untuk deteksi bahasa kasar dalam komentar berita online Indonesia', *e-Proceeding Eng.*, vol. 8, no. 2, pp. 3385–3395, 2021.

[10] N. Kitaev, Ł. Kaiser, and A. Levskaya, 'Reformer: The Efficient Transformer', *ICLR 2020*, pp. 1–12, 2020, [Online]. Available: http://arxiv.org/abs/2001.04451.

[11] S. Sukhbaatar, E. Grave, P. Bojanowski, and A. Joulin, 'Adaptive attention span in transformers', *ACL 2019 - 57th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf.*, no. 2015, pp. 331–335, 2020, doi: 10.18653/v1/p19-1032.

[12] D. Britz, A. Goldie, M. T. Luong, and Q. V. Le, 'Massive exploration of neural machine translation architectures', *EMNLP 2017 - Conf. Empir. Methods Nat. Lang. Process. Proc.*, pp. 1442–1451, 2017, doi:

10.18653/v1/d17-1151.

[13] Y. Belinkov, S. Gehrmann, G. Ai, and E. Pavlick, ‘Tutorial Proposal: Interpretability and Analysis in Neural NLP. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Tutorial Abstracts’, *ACL Meet. Assoc. Comput. Linguist.*, pp. 1–5, 2020, [Online]. Available: https://doi.org/10.18653/v1/P17.

[14] N. Hayatin, K. M. Ghufron, and G. W. Wicaksono, ‘Summarization of COVID-19 news documents deep learning-based using transformer architecture’, *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 19, no. 3, pp. 754–761, 2021, doi: 10.12928/TELKOMNIKA.v19i3.18356.

[15] C. Raffel *et al.*, ‘Exploring the limits of transfer learning with a unified text-to-text transformer’, *J. Mach. Learn. Res.*, vol. 21, pp. 1–67, 2020.

[16] A. Alomari, N. Idris, A. Q. M. Sabri, and I. Alsmadi, ‘Deep reinforcement and transfer learning for abstractive text summarization: A review’, *Comput. Speech Lang.*, vol. 71, no. August 2021, p. 101276, 2022, doi: 10.1016/j.csl.2021.101276.

[17] T. Wolf *et al.*, ‘Transformers: State-of-the-Art Natural Language Processing’, pp. 38–45, 2020, doi: 10.18653/v1/2020.emnlp-demos.6.

[18] K. Song, X. Tan, T. Qin, J. Lu, and T. Y. Liu, ‘MASS: Masked sequence to sequence pre-training for language generation’, *36th Int. Conf. Mach. Learn. ICML 2019*, vol. 2019-June, pp. 10384–10394, 2019.

[19] A. Roberts, C. Raffel, and N. Shazeer, ‘How much knowledge can you pack into the parameters of a language model?’, *EMNLP 2020 - 2020 Conf. Empir. Methods Nat. Lang. Process. Proc. Conf.*, pp. 5418–5426, 2020, doi: 10.18653/v1/2020.emnlp-main.437.

[20] K. Kurniawan and S. Louvan, ‘INDOSUM : A New Benchmark Dataset for Indonesian Text Summarization’, *2018 Int. Conf. Asian Lang. Process.*, pp. 215–220, 2018.

[21] S. Ontoum and J. H. Chan, ‘Automatic Text Summaration of COVID-19 Scientific Research Topics Using Pre-trained Model from HuggingFace®’, pp. 1–8, 2022.

[22] T. He *et al.*, ‘ROUGE-C: A fully automated evaluation method for multi-document summarization’, *2008 IEEE Int. Conf. Granul. Comput. GRC 2008*, pp. 269–274, 2008, doi: 10.1109/GRC.2008.4664680.

[23] Y. Yuliska and K. U. Syaliman, ‘Literatur Review Terhadap Metode, Aplikasi dan Dataset Peringkasan Dokumen Teks Otomatis untuk Teks Berbahasa Indonesia’, *IT J. Res. Dev.*, vol. 5, no. 1, pp. 19–31, 2020, doi: 10.25299/itjrd.2020.vol5(1).4688.