ABSTRACTIVE-BASED AUTOMATIC TEXT SUMMARIZATION ON INDONESIAN NEWS USING GPT-2

Aini Nur Khasanah[1*], Mardhiya Hayaty[1]
[1]Computer Science,Universitas Amikom Yogyakarta
email: [*]aini.khasanah@students.amikom.ac.id, mardhiya_hayati@amikom.ac.id

Abstract: Automatic text summarization is challenging research in natural language processing, aims to obtain important information quickly and precisely. There are two main approach techniques for text summary: abstractive and extractive summary. Abstractive Summarization generates new and more natural words, but the difficulty level is higher and more challenging. In previous studies, RNN and its variants are among the most popular Seq2Seq models in text summarization. However, there are still weaknesses in saving memory; gradients are lost in long sentences so resulting in a decrease in lengthy text summaries. This research proposes a Transformer model with an Attention mechanism that can fetch important information, solve parallelization problems, and summarize long texts. The Transformer model we propose is GPT-2. GPT-2 uses decoders to predict the next word using the pre-trained model from w11wo/indo-gpt2-small, implemented on the Indosum Indonesian dataset. Evaluation assessment of the model performance using ROUGE evaluation. The study's results get an average result recall for R-1, R-2, and R-L were 0.61, 0.51, and 0.57, respectively. The summary results can paraphrase sentences, but some still use the original words from the text. Future work increase the amount of data from the dataset to improve the result of more new sentence paraphrases.

Keywords: ATS; generatif pre-trained 2; indosum; transformer

Abstrak: Peringkasan teks otomatis merupakan penelitian menantang dalam pemrosesan bahasa alami, bertujuan untuk mendapatkan informasi penting dengan cepat dan tepat. Ada dua teknik pendekatan utama untuk ringkasan teks: ringkasan abstraktif dan ekstraktif. Peringkasan abstraktif menghasilkan kata-kata baru dan lebih alami, tetapi tingkat kesulitannya lebih tinggi dan lebih menantang. Dalam penelitian sebelumnya, RNN dan variannya termasuk model Seq2Seq yang paling populer dalam peringkasan teks. Namun, masih ada kelemahan dalam menyimpan memori; gradien hilang dalam kalimat yang panjang sehingga mengakibatkan penurunan ringkasan teks yang panjang. Penelitian ini mengusulkan model Transformer dengan mekanisme Attention yang dapat mengambil informasi penting, memecahkan masalah paralelisasi, dan meringkas teks yang panjang. Model Transformer yang dusulkan adalah GPT-2. GPT-2 menggunakan decoder untuk memprediksi kata berikutnya menggunakan model *pre-trained* dari w11wo/indo-gpt2-small, yang diimplementasikan pada *dataset* Indosum Indonesia. Evaluasi penilaian kinerja model menggunakan evaluasi ROUGE. Hasil penelitian mendapatkan hasil rata-rata *recall* untuk R-1, R-2, dan R-L berturut-turut adalah 0,61, 0,51, dan 0,57. Hasil ringkasan dapat memparafrasekan kalimat, tetapi beberapa masih menggunakan kata asli dari teks. Penelitian di masa depan dapat meningkatkan jumlah data dari dataset untuk meningkatkan hasil parafrase kalimat baru lebih banyak.

Kata kunci: ATS; generatif pre-trained 2; indosum; transformer

INTRODUCTION

Automated text summaries are important research in natural language processing (NLP). The purpose of doing a text summary is to summarize the main news information from the long original text (document) to be shorter [1]. Automate text summarization using deep learning, which is currently popular, so that the content can be understood more quickly and eliminate the manual summarising process, which takes much time to read.

Techniques for summarizing text are divided into two, namely: extractive text summaries and abstract text summaries. Extractive summarization is simplifying text by selecting and taking sentences or words from the original source text without making any changes or modifications to those sentences. Meanwhile, translating using the abstract method is more challenging because it has a greater difficulty, and there needs to be more research in this field. Still, the results of summarizing abstractive texts are more natural like the summary created by humans which produces new sentences/words, making it easier to understand, and the cohesion between sentences is higher and produces new sentences/words [2], [3]. Based on the input, text summarization techniques can also be divided into two categories, single-document summarization which focuses on summarizing a single text document, whether it's a long or short document, and multi-document summarization utilizing multiple text documents to produce a summary [4]. In this research, the author focuses on the single-document summarization approach to generate a summary from a single news article with the goal of preserving important information and facilitating understanding.

The most popular summary method uses an RNN, but this method [5] there are still has weaknesses in saving memory; gradients are lost in long sentences so resulting in a decrease in lengthy text summaries. Compared to the Seq2Seq and RNN models [6], [7], Transformer models with an Attention mechanism can eliminate repetition and convolution and can perform parallelization [8] and do long text summaries.

Therefore, according observation, compared to the Seq2Seq model, the Transformer GPT-2 model ditches the encoder block and reduces it to a decoder stack, allowing for predicting the next word in the sequence [9] that is expected to predict new words. The accuracy of the GPT-2 model, which is a autoregressive transformer model is an open-source artificial intelligence model developed by OpenAI. This model has 1.5 billion parameters trained on an 8 million web pages dataset [10] by using pre-training in predicting new words from the given input.

METHOD

Automatic Text Summarization

Model seq2seq [10] is used for NLP tasks such as automatic translation and text generation. The model consists of an encoder and a decoder, where the encoder processes the input and compresses the information into a context vector, the decoder processes the information from the encoder and produces the output. The weakness of the seq2seq model is its inability to remember long sentences, the attention mechanism is used to address this.

In research that was popular back then, there is one variation of the RNN, namely LSTM, which stores relevant in-

formation and forgets information that is not important in helping to produce a summary. However, the LSTM model only takes one input at a time, which is a problem because it is proven not to work for every case, whereas in the seq2seq model implemented with the encoder-decoder framework, there are still problems with parallelization [6].

This study uses a transformer architecture achieving a much lower word error rate (WER) than the RNN architecture, with a score of 2.8% and 7.3% WER for clean and other test data from LibriSpeech [11]. Variation of the LSTM two-way encoder [12] yielding 8.5% ROUGE-1 means that only a fraction of the words of the original text document can be found in the resulting content.
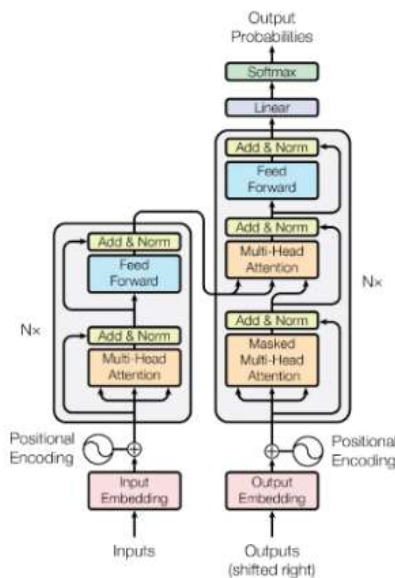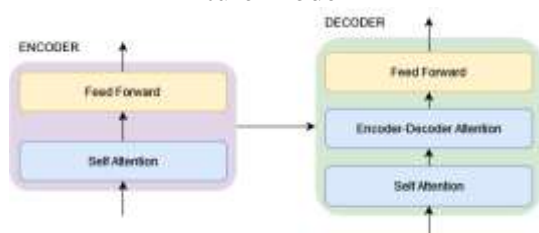

Image 1. Transformer Architecture Model

Image 2. Encoder-Decoder

The Transformer Architecture Model which is shown in Image 1. is in-spired by the encoder-decoder framework and are based on several layers of attention. The Transformer model links the encoder-decoder via the Attention Mechanism, eliminating repetition and convolution completely[8], and maps input sequences to hidden states using attention [13]. Image 2. The function of the encoder is to process the input with a fixed length and sequence, while the decoder is to display the output (target) from the encoder. Each encoder layer uses self-attention to represent the context. Each decoder layer also uses self-attention in two sub-layers. A transformer Model Encoder can work on a sequence of inputs in parallel but the decoder is auto-regressive, each output is affected by the previous output symbol [8].

The Transformer model has a strategy for memorizing the entire sequence with self-attention, processing all inputs at once. So that makes this model very fast when trained with large amounts of data[8]. Attention [14] is an increasingly popular mechanism used in various neural architectures.
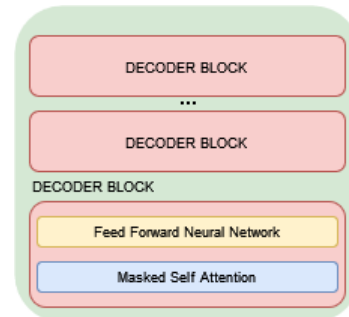

Image 3. Transformer-Decoder

Model Generatif Pre-Trained (GPT-2). The GPT-2 model using block decoder transformers in Image 3. with 1024 tokens and 50,257 vocabularies[15] removing the encoder-transformer, has 4 model sizes consisting of different architectural hyperparameters – small, medium, large, and XL, with parameters 124M, 355M, 774M, and 1.5B respec-

tively [16]. Can predict the next token in sequence because it is trained with Casual Language Modeling (CLM) goals[17]. GPT-2 uses Byte Pair Encoding to generate tokens in vocab, using Byte Pair Encoding is a better strategy to generate code than just using characters[18]. In contrast to BERT, whose self-attention layer covers the token in front by changing the word to a mask, GPT-2, with self-attention calculations, blocks information from tokens to the right of the calculated position. The GPT-2 model includes the single transformer model which treats the summary as the next text to be generated after the input text, such as a TL;DR statement at the end of a long article [19].

The GPT-2 model can quickly achieve better performance for code generation as shown by the continuously increasing BLEU score of 0.22 [18]. In this case[16], the GPT-2 model in the SARS-CoV-2 news dataset gets a score of 0.348 for ROUGE-2 and a score of 0.358 for ROUGE-L. The GPT-2 model can also be implemented in the task of generating automated questions that generate meaningful and diverse questions [9] so that it is expected to be implemented in the text summary. In the LCSTS dataset by implementing GPT-2 an increase of 10.75% ROUGE-1, 13.85% ROUGE-2, and 9.73% ROUGE-L [20].

Here we propose to use the GPT-2 Language Model for generating text summaries automatically by utilizing the pre-training model from w11wo/indo-gpt2-small[21] which has a parameter of 124M, and this pre-training model is built based on the GPT-2 pre-training model English with fine-tuning on the Indonesian Wikipedia dataset.

Experiment

In this section we will propose a GPT-2 transformer model for automatic text summarization and will describe the basic experimental, discuss evaluation metrics and describe the model used for this research, the research flow can be seen in Image 4.
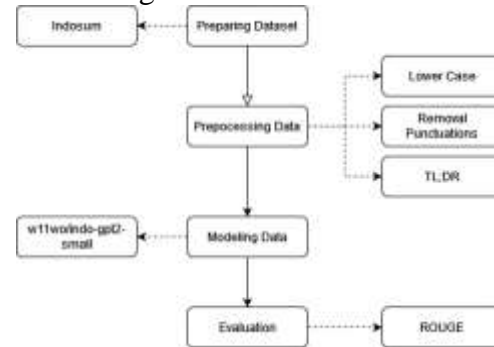

Image 4. Research Flow

In this research, the selected hyperparameters were adjusted to the available and limited hardware resources. This is done to achieve research results efficiently and not consume too many resources. This is important to ensure that the research can be carried out well even with the limited available resources.

The hardware components used in this research can be described as follows : Python 3.6 version; GPU configuration with high RAM capacity; a combination of Transformer and PyTorch as the machine learning model architecture; basic libraries in Python machine learning.

Preparing Dataset. The dataset that we use comes from Indosum[22], [23], an Indonesian language news dataset which has 2 columns, text and summary, with a total of 18774 Indosum data. However, in this research, only 9387 data will be used. The division of the training data used is 8448 (20% of the training data = 1690 data used for validation data, and 6758 for training data), and the test data is 939 data.

Preprocessing Data. This dataset consists of long news articles and short summaries for comparison as targets. The raw dataset is then cleaned using preprocessing techniques. The pre-

processing used is Lower Casing, Elimination of Punctuation, Tags and HTML links. To summarize, the corresponding marker is the TL;DR symbol where this token is placed between the input and summary text. Using pre-processing techniques, making it easier for the machine to train the data.

Evaluation ROUGE.

To evaluate and assess the performance of the model, ROUGE [24] evaluation will be used as the parameter that the researcher considers appropriate to measure the accuracy of text summarization. In this study, the authors will use ROUGE-1, ROUGE-2 and ROUGE-L evaluations in automatic text summarization. ROUGE (Recall-Oriented Understudy for Gisting Evaluation) is a metric used in NLP to evaluate text summarization. This metric compares model-generated summaries to human-generated target summaries. ROUGE-L (1) measures the longest matching sequence of words, without a predefined n-gram length, and does not require sequential matches [25]. ROUGE-n measures the n-gram overlap between two texts, where ROUGE-1 measures the similarity between the summary and the reference based on unigram (single word sequence) and ROUGE-2 measures the similarity between the summary and the reference based on bigram (consecutive two word sequence)[26]. In ROUGE-1, ROUGE-2, and ROUGE-L, there are 3 metrics [26]–[28] used, namely Recall (2) to measure how much of the reference summary is produced by the summarization model, Precision (3) allows to measure how many of the system's summaries are actually relevant or required, and F1-Score (4) calculates the harmonic mean of recall and precision to measure how well a model predicts.

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

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

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

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

RESULT AND DISCUSSION

In this study, using the Indosum dataset, the training model was carried out with the pre-trained model from w11wo/indo-gpt2-small with 15 epochs experiments and batch size 4. This is due to limitations in the author's resources which affects the limitations in testing. The evaluation results obtained from the applied model can show training graphs and validation losses in Image 5. So that the average ROUGE evaluation results are obtained which can be seen in Table 1. It means that the summary contains relevant information from the reference, with a high percentage of the same words between the summary and the reference summary, and the summary does not contain irrelevant information, so the model can predict the summary with enough accuracy.
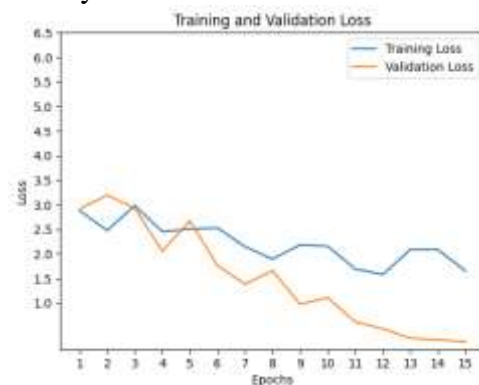

Image 5. Training and Validation Loss

After calculating the ROUGE value, based on Table 1 it can be said that the model sometimes fails to capture im-

portant information with the average recall results obtained R-1, R-2, and R-L respectively 0.61, 0.51, and 0.57.

In the results of research conducted by the author, you can see representative examples of the summary results. The first example in table 2 and table 3 has a high ROUGE score. The second example in table 4 and table 5 has a low ROUGE score. The comparison is made to see the difference between high and low ROUGE scores in the text summary results.

In table 2 and table 3 the results of the ROUGE score are high with summary results that can already be paraphrased even if it's just words **"pilihan presiden (pilpres)"** becomes **"pilpres"** and **"persatuan Indonesia (perindo)"** becomes **"perindo"** but can immediately continue with the next condition which is still in the same context.

Meanwhile in table 4 and table 5 with a low ROUGE score, there are still summary results that are out of context as can be seen in the summary model which mentions **"floorings"** while in the article

and reference summary there are no such words and in the resulting model summary there are still repeated words that do not fit the context such as **"kemustensi"**, **"kompetisi"** and **"bintang".**

Based on the high and low scores in the summary, the author's analysis shows that a low Rouge score can result in irrelevant and contextually inconsistent information because the data used is not much due to the limitations of the devices used for research so some words are not trained to the max. Meanwhile, a high Rouge score produces relevant information with a high percentage of the same words and the same sequence length between the summary and the reference, with a recall value of 0.82 in R1 and RL, which can be said to have fewer paraphrased words. Compared to the results of previous studies [23] by using the Indosum dataset which totaled 20K for text summarization, it was found that the "NeuralSum 300 emb. Size" has R-1, R-2, and R-L performance on the F1-Score of 67.96, 61.65, and 67.24, respectively.

Table 1. Rouge Evaluation Average Results

Score	Evaluasi ROUGE		
	ROUGE-1	**ROUGE-2**	**ROUGE-L**
Recall	0.61	0.51	0.57
Precission	0.55	0.45	0.51
F1-Score	0.56	0.46	0.52

Table 2. High Rouge Scores In Text Summary

Score	Evaluasi ROUGE		
	ROUGE-1	**ROUGE-2**	**ROUGE-L**
Recall	0.82	0.71	0.82
Precission	0.92	0.82	0.92
F1-Score	0.87	0.76	0.87

Table 3. Experimental Results Of High Rouge Scores In Text Summary

Comparison	High Rouge Scores In Text Summary
	Result
Article	Jakarta CNN Indonesia - - Partai Persatuan Indonesia (Perindo) bentukan pengusaha Hary Tanoesoedibjo mewacanakan untuk mengusung Presiden Joko Widodo pada pemilihan presiden (Pilpres) 2019 . Dukungan kepada Jokowi akan diputuskan pada rapat pimpinan nasional (Rapimnas) Perindo dan ditetapkan pada Kongres Perindo akhir tahun ini …….. Sinyal tersebut disampaikan Hary Tanoe seusai acara Penganugerahan Kepala Daerah Inovatif Koran SINDO 2017 di Hotel Westin Jakarta Selasa 1 Agustus 2017 yang juga dihadiri Mendagri Tjahjo Kumolo.
Reference Summary :	Partai Persatuan Indonesia bentukan pengusaha Hary Tanoesoedibjo mewacanakan untuk mengusung Presiden Joko Widodo pada pemilihan presiden 2019. Dukungan kepada Jokowi akan diputuskan pada rapat pimpinan nasional Perindo dan ditetapkan pada Kongres Perindo akhir tahun ini." Ada kesan kuat yang akan diusulkan adalah presiden yang sekarang karena kemungkinan yang menang sekarang " kata Sekretaris Jenderal Partai Perindo Ahmad Rofiq.
Model Summary :	**perindo** bentukan pengusaha hary tanoe tanoe mewacanakan untuk mengusung presiden joko widodo pada **pilpres** 2019. dukungan jokowi akan diputuskan pada rapat pimpinan nasional rapimnas perindo akhir tahun ini. ada kesan kuat yang akan diusulkan adalah presiden yang sekarang karena kemungkinan yang menang sekarang kata sekretaris jenderal partai perindo ahmad rofiq.

Table 4. Low Rouge Scores In Text Summary

Score	Evaluasi ROUGE		
	ROUGE-1	**ROUGE-2**	**ROUGE-L**
Recall	0.17	0.03	0.15
Precission	0.22	0.04	0.19
F1-Score	0.19	0.04	0.17

Table 5. Experimental Results Of Low Rouge Scores In Text Summary

Comparison	Low Rouge Scores In Text Summary
	Result
Article	Kamu yang senang berburu tiket murah untuk berlibur pasti selalu menanti travel fair . Pasalnya di travel fair kamu bisa mencari tiket pesawat dengan harga miring yang menjadi keuntungan tersendiri terutama untuk para traveller yang ingin menghemat dana . Mengerti akan kebutuhan para wisatawan dalam mencari tiket pesawat low budget Singapore Airlines menggelar travel fair yang diadakan di main atrium Gandaria City Jakarta Selatan …... Jakarta - Korea pergi pulang mulai dari Rp 4 juta, Jakarta - Eropa pergi pulang mulai dari Rp 7 juta, Jakarta - Amerika pergi pulang mulai dari Rp 95 juta.
Reference Summary :	Singapore Airlines menggelar travel fair dalam rangka HUTnya ke - 70 yang diadakan di main atrium Gandaria City Jakarta Selatan mulai hari ini .] Edwin Chiang Berikut penawaran menarik : Jakarta - Singapura PP mulai dari Rp 900 ribu Jakarta - Hong Kong dan Thailand PP mulai dari Rp 21 juta Jakarta - Korea PP mulai dari Rp 4 juta Jakarta - Eropa PP mulai dari Rp 7 juta Jakarta - Amerika PP mulai dari Rp 95 juta .
Model Summary :	**pada floorings ada banyak program yang bisa kamu cari untuk berlibur**. salah satunya adalah membuat travel fair atau bermain di singapore airlines. pada acara ini kemustensi **kemustensi** singapore airlines mengadakan **kompetisi** dengan mengundang banyak **bintang** dari berbagai negara. selain membuat travel fair juga akan mengadakan kompetisi dengan mengundang banyak bintang dari berbagai

CONCLUSION

In summary, this study introduced a news summarization algorithm using the GPT2 model and evaluated it with the ROUGE metric. Results varied based on ROUGE scores: low scores led to repetitive and contextually flawed summaries, while high scores indicated effective paraphrasing and context retention. The study identified the need for better pre-processing and more robust hardware. The authors suggest expanding the dataset for improved results in future research, aiming to establish the dataset, and evaluation results can become a GPT2 benchmark for Indonesian text summarization.

BIBLIOGRAPHY

[1]   Q. Guo, J. Huang, N. Xiong, and P. Wang, "MS-pointer network: Abstractive text summary based on multi-head self-attention," *IEEE Access*, vol. 7, pp. 138603–138613, 2019, doi: 10.1109/ACCESS.2019.2941964.

[2]   R. Adelia, S. Suyanto, and U. N. Wisesty, "Indonesian Abstractive Text Summarization Using Bidirectional Gated Recurrent Unit," *Procedia Comput Sci*, vol. 157, pp. 581–588, 2019, doi: 10.1016/j.procs.2019.09.017.

[3]   N. Raphal, H. Duwarah, and P. Daniel, "Survey on Abstractive Text Summarization," in *2018 International Conference on Communication and Signal Processing (ICCSP)*, IEEE, Apr. 2018, pp. 0513–0517. doi: 10.1109/ICCSP.2018.8524532.

[4]   R. Srivastava, P. Singh, K. P. S. Rana, and V. Kumar, "A topic modeled unsupervised approach to single document extractive text summarization," *Knowl Based Syst*, vol. 246, p. 108636, Jun. 2022, doi: 10.1016/j.knosys.2022.108636.

[5]   D. Hu, "An Introductory Survey on Attention Mechanisms in NLP Problems," Nov. 2018, [Online]. Available: http://arxiv.org/abs/1811.05544

[6]   A. Gupta, D. Chugh, and R. Katarya, "Automated News Summarization Using Transformers."

[7]   R. Nallapati, B. Zhou, C. N. dos santos, C. Gulcehre, and B. Xiang, "Abstractive Text Summarization Using Sequence-to-Sequence RNNs and Beyond," Feb. 2016, [Online]. Available: http://arxiv.org/abs/1602.06023

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

[9]   T. Klein and M. Nabi, "Learning to Answer by Learning to Ask: Getting the Best of GPT-2 and BERT Worlds," Nov. 2019, [Online]. Available: http://arxiv.org/abs/1911.02365

[10]   A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, "Language Models are Unsupervised Multitask Learners." [Online]. Available: https://github.com/codelucas/newspaper

[11]   N. Moritz, T. Hori, and J. Le, "Streaming Automatic Speech Recognition with the Transformer Model," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, May

2020, pp. 6074–6078. doi: 10.1109/ICASSP40776.2020.9054476.

[12] G. Szucs and D. Huszti, "Seq2seq Deep Learning Method for Summary Generation by LSTM with Two-way Encoder and Beam Search Decoder," in *2019 IEEE 17th International Symposium on Intelligent Systems and Informatics (SISY)*, IEEE, Sep. 2019, pp. 221–226. doi: 10.1109/SISY47553.2019.9111502.

[13] K. Ahmed, S. Keskar, and R. Socher, "WEIGHTED TRANSFORMER NETWORK FOR MACHINE TRANSLATION."

[14] A. Galassi, M. Lippi, and P. Torroni, "Attention in Natural Language Processing," *IEEE Trans Neural Netw Learn Syst*, vol. 32, no. 10, pp. 4291–4308, Oct. 2021, doi: 10.1109/TNNLS.2020.3019893.

[15] P. J. Liu *et al.*, "Generating wikipedia by summarizing long sequences," *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings*, pp. 1–18, 2018.

[16] H. Batra *et al.*, "CoVShorts: News Summarization Application Based on Deep NLP Transformers for SARS-CoV-2," in *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), ICRITO 2021*, Institute of Electrical and Electronics Engineers Inc., 2021. doi: 10.1109/ICRITO51393.2021.9596520.

[17] G. Lample and A. Conneau, "Cross-lingual Language Model Pretraining," Jan. 2019, doi: 10.48550/arXiv.1901.07291.

[18] L. Perez, L. Ottens, and S. Viswanathan, "Automatic Code Generation using Pre-Trained Language Models," Feb. 2021, [Online]. Available: http://arxiv.org/abs/2102.10535

[19] S. Shearing, A. Gertner, B. Wellner, and L. Merkhofer, "Automated Text Summarization: A Review and Recommendations," 2020.

[20] Q. Zhu, L. Li, L. Bai, and F. Hu, "Chinese text summarization based on fine-tuned GPT2," in *Third International Conference on Electronics and Communication; Network and Computer Technology (ECNCT 2021)*, M. K. Mohiddin, S. Chen, and S. F. EL-Zoghdy, Eds., SPIE, Mar. 2022, p. 109. doi: 10.1117/12.2629132.

[21] "w11wo/indo-gpt2-small · Hugging Face." https://huggingface.co/w11wo/indo-gpt2-small (accessed Feb. 24, 2023).

[22] "kata-ai/indosum: A benchmark dataset for Indonesian text summarization.", Accessed: Jun. 19, 2023. [Online]. Available: https://github.com/kata-ai/indosum

[23] K. Kurniawan and S. Louvan, "IndoSum: A New Benchmark Dataset for Indonesian Text Summarization," in *Proceedings of the 2018 International Conference on Asian Language Processing, IALP 2018*, Institute of Electrical and Electronics Engineers Inc., Jan. 2019, pp. 215–220. doi: 10.1109/IALP.2018.8629109.

[24] T. He *et al.*, "ROUGE-C: A fully automated evaluation method for multi-document summarization," in *2008 IEEE International Conference on Granular Computing*, IEEE, Aug. 2008, pp. 269–274. doi: 10.1109/GRC.2008.4664680.

[25] V. Kieuvongngam, B. Tan, and Y. Niu, "Automatic text summarization of covid-19 medical research articles using bert and gpt-2," *arXiv preprint arXiv:2006.01997*, 2020, [Online]. Available: https://arxiv.org/abs/2006.01997

[26] J. D'Silva and U. Sharma, "Unsupervised Automatic Text Summarization of Konkani Texts using K-means with Elbow Method," *International Journal of Engineering Research and Technology*, vol. 13, no. 9, pp. 2380–2384, 2020, doi: 10.37624/ijert/13.9.2020.2380-2384.

[27] S. Esmaeilzadeh, G. X. Peh, and A. Xu, "Neural abstractive text summarization and fake news detection," *arXiv preprint arXiv:1904.00788*, 2019, [Online]. Available: https://arxiv.org/abs/1904.00788

[28] U. Sharma, "Development of a Konkani Language Dataset for Automatic Text Summarization and its Challenges." [Online]. Available: http://www.irphouse.com