



## Enhancing Text Summarization with a T5 Model and Bayesian Optimization

Arif Ridho Lubis<sup>1\*</sup>, Habibi Ramdani Safitri<sup>1</sup>, Irvan<sup>2</sup>, Muharman Lubis<sup>3</sup>, Muhammad Luthfi Hamzah<sup>4</sup>,  
Al-Khowarizmi Al-Khowarizmi<sup>5</sup>, Okvi Nugroho<sup>6</sup>

<sup>1</sup> Department of Computer Engineering and Informatics, Politeknik Negeri Medan, Medan 20155, Indonesia

<sup>2</sup> Department of Mathematic Education, Universitas Muhammadiyah Sumatera Utara, Medan 20238, Indonesia

<sup>3</sup> School of Industrial Engineering, Telkom University, Bandung 40257, Indonesia

<sup>4</sup> Faculty of Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru 28293, Indonesia

<sup>5</sup> Department of Information Technology, Universitas Muhammadiyah Sumatera Utara, Medan 20238, Indonesia

<sup>6</sup> Kulkas IT, Medan 20219, Indonesia

Corresponding Author Email: [arifridho@polmed.ac.id](mailto:arifridho@polmed.ac.id)

<https://doi.org/10.18280/ria.370513>

### ABSTRACT

**Received:** 9 July 2023

**Revised:** 21 July 2023

**Accepted:** 30 July 2023

**Available online:** 31 October 2023

#### **Keywords:**

*natural language processing, T5, Bayesian optimization, text summarization*

At present the habits and interests of individuals in obtaining information by reading large amounts of information have changed at the stage of reading information more concisely, but these changes have challenges such as the nature of the data which is still unstructured making it difficult to summarize text. This study applies a data cleaning process with text processing and manually annotates to divide the data into summary data and text data so that it can be used for the process of implementing the T5 model and Bayesian optimization. In the implementation of Bayesian optimization using the prior distribution and likelihood parameters. In implementing the T5 model there will be several stages such as processing training and test data then Decodification and Post-Processing processes. The results of this study were obtained using the ROUGE evaluation technique which resulted in an increased evaluation value. The T5 model produces a ROUGE 1 value with an average value of 0.42, ROUGE-2 has a value of 0.55 and ROUGE-L has a value of 0.46 while applying Bayesian optimization produces a ROUGE-1 evaluation with an average value of 0.53 ROUGE-2 has a value of 0.55 and ROUGE-L has a value of 0.59.

## 1. INTRODUCTION

Technological developments, especially on social media platforms, have an impact on text data that has a large volume which is a challenge in processing data that is unstructured and semi-structured [1-3]. In carrying out data processing including current processes and analysis, technology and methods in the field of natural language processing can be used [4, 5]. Currently, the habits of individuals in obtaining information by reading large amounts of information have changed, becoming more interested in reading information more concisely [6-8], these changes occur because activities in social media have character limits such as social media. social Twitter so that many individuals are accustomed to more concise information [9, 10].

The problems that occur are the difficulty of summarizing efficiently in a fast time and the difficulty of getting quality summaries that can be understood by many people so this research requires techniques for summarizing texts or summaries. There are many approaches to summarizing, one of which is the T5 model which performs text-to-text extraction tasks with a transformer architecture that uses pretraining and fine-tuning, but these approaches still have poor summary quality and low evaluation accuracy, so a Bayesian optimization approach is needed that functions to search for the best parameters for text summarizing tasks. This study makes a difference to research studies that have been carried out by applying Bayesian optimization to the

parameters of the T5 model. In conducting a text summary there are advantages, including: providing a summary of research abstracts, can reduce reading time on large amounts of text data, and can identify important sentences or words in a text data [11, 12].

In the technique of performing text summaries automatically, it can be categorized into two, namely extractive and abstractive [13], text summaries in the extractive category extract the most important sentences from the text so that they can form a summary, and processors in extractive techniques sort sentences that are judged by importance or no [14]. After being given the value of the sentences will be combined so as to form a summary [15]. While the category of abstract techniques will form word summaries without using the original input text sentences. The procedure of summarizing with abstract techniques forms a summary with new text data that uses a more understandable and simpler grammar [16, 17]. The text summary is the most complicated challenge in the field of natural language preprocessing because the source text is unstructured and there are no labels. Many researchers manual annotation as Al-Laith et al. [18], Krommyda et al. [19], and Hananto et al. [20] do manual annotation with the aim of obtaining input text data to be used in the field of natural language processing. In carrying out text summary tasks there are many techniques that can be used, one of which is using the most popular models such as the Transformer-based Text-to-Text Transfer Transformer or the T5 model which has advantages in producing good text

summaries. However, several studies conducted by Mastropaolo et al. [21], Fendji et al. [22], and Blekanov et al. [23] revealed that the T5 model still has room to improve performance in the process of carrying out text summary tasks, this study will use Bayesian optimization techniques that will perform a task in increasing the performance of the T5 model. In the Bayesian optimization task, we will use Bayesian probability theory for an iterative model so that it can have the advantage of updating initial knowledge. This research can help in improving the model that ignores text or information that has important value in producing a summary. This study will use an abstract Indonesian text technique with the T5 model which is carried out to increase performance with Bayesian optimization, the contribution of this study is in the form of a collection of Indonesian language data related to research paper abstracts, this research can extract more than 10 words, this research will use an evaluation technique ROUGE score.

Bayesian optimization is an approach technique for searching for the optimum value of a function by using the probabilistic of the overall search and evaluating the function [24, 25], Bayesian will use the theory of Bayesian probability for an iterative model so that it can have the advantage of updating initial knowledge [26, 27]. This research can help in improving the model that ignores text or information that has important value in producing a summary. This study will use an abstract Indonesian text technique with the T5 model which is carried out to increase performance with Bayesian optimization, the contribution of this study is in the form of a collection of Indonesian language data related to research paper abstracts, this research can extract more than 10 words, this research will use an evaluation technique ROUGE score.

## 2. RELATED WORK

Many studies have explored the problems in carrying out text summary tasks using abstractive and extractive methods such as those carried out by Rofiq [28] and Moratanch and Chitrakala [29] who conducted text summary research using extractive techniques that produce summaries by collecting important words, while research uses abstractive methods or techniques as practiced by Khan et al. [30], Masum et al. [31], and Elsaied et al. [32] produce important words then these words have values which are then sorted to produce summary words. In summary, the text has a significant development starting from the use of important features such as frequency, word count, and word similarity [19, 33]. The advantages of the abstractive approach can remove words that are considered unimportant, while the extractive approach performs summaries based on different phrases in the input data [34]. In the Transformer-based Text-to-Text Transfer Transformer or T5 model, a lot of text summaries are carried out with an abstract approach such as that done by Patwardhan et al. [35], Cheng and Yu [36], and Mars [37] which goes through several stages such as tokenization, formation of input-output data, Pretraining and Fine-Tuning, Encoder-Decoder Transformation, and text generation.

The T5 model possesses several advantages, including an end-to-end approach, access to large-scale data with contextual representation, and a more consistent process [38, 39]. The process of the T5 model uses encoders and decoders that can produce text summaries. In measuring the performance of text summary models, many researchers use

performance evaluations such as ROUGE [40-42].

Many studies have carried out a combination of models in carrying out text summary tasks, as was done by Ahuir et al. [43] who carried out text summaries using a combination of a transfer model that had been trained with the T5 model, another study conducted by Chouikhi and Alsuhaibani [44] who carried out text summaries with the T5 model. deep transformers that perform the hyperparameter combinations of the T5 model. Another study by La Quatra and Cagliero [34] used the BERT model to produce text summaries. In conducting text summaries, there are many challenges faced, one of which is the structure of text data, currently the resulting text data is unstructured and semi-structured. Unstructured data needs to be preprocessed before being processed into a text summary model [36, 45]. The use of the T5 model produces evaluations with various ROUGE as done by Chouikhi and Alsuhaibani [44] resulting in a Rouge-1 value of 0.5 and ROUGE-2 0.7. another study conducted by Leitao Martins et al. [45] produced a ROUGE-1 value of 0.3 and a ROUGE-2 value of 0.4. research by Garg [46] resulted in an evaluation value of ROUGE-1 0.4, ROUGE-2 0.5. The use of the T5 model can produce a good evaluation based on the research conducted. However, evaluation improvements need to be carried out, such as Garg's research [46] which carried out evaluation improvements with a pre-trained model, and other research by Rothe et al. [47] implemented BERT-LSTM to improve text summary performance evaluations. In improving the performance of several studies using Bayesian optimization as done by Kolar et al. [48] using Bayesian optimization to improve performance in models for the task of finding faults on a machine, research Moro et al. [49] improve the performance of long documents for text summaries utilizing one of the features of Bayesian optimization. Text summary needs to improve performance with Bayesian optimization, Bayesian optimization improves performance with probability and explores words that are not important.

## 3. DATA ACQUISITION AND PROCESSING

In data acquisition, data will be obtained from the Medan State Polytechnic site where related research data and summaries will be collected in such a way, this study will perform text summary abstracts from research data consisting of abstracts and brief information on the abstracts. To be able to obtain relevant data, web scraping will be carried out, the process flow of which can be seen in Figure 1.

The description of Figure 1 explains that the software will connect to the Medan State Polytechnic and retrieve data such as abstracts and important information which will be stored in a dataset file in CSV format. Data was collected from 2019 to 2023. In collecting data, text mining will be used by utilizing data crawling techniques to retrieve data from the website. The data preprocessing will be carried out with the aim of getting an evaluation of a model that is working well, in the preprocessing stage it will change the data from structured to unstructured. In some literature, it is explained that data preprocessing is a must for processing text data in the field of natural language processing. At the preprocessing stage [50], the text will carry out the process stages such as case folding which changes the letters in the data to lowercase, tokenizing aims to separate each word in the data, stopword aims to collect words that are not important and deletes data and stemming aims to make changes to words that affixes become

standard words or basic words after that the data is analyzed manually to see the word structure if there are still words that are not good then manual annotations will be made to ensure the data is structured then the data will be divided into 2 columns, namely the text data column and the summary column data with the aim of implementing the T5 model and optimizing parameters with Bayesian optimization.

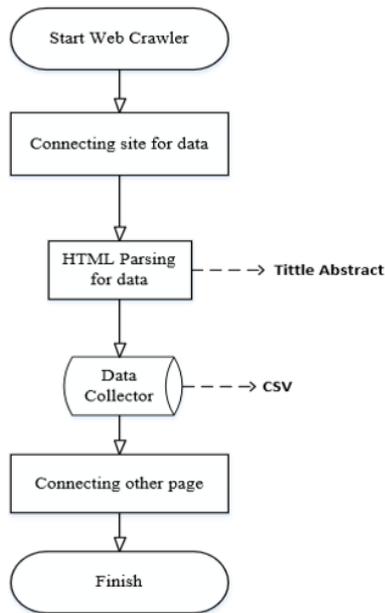


Figure 1. Data scraping chart

#### 4. RESEARCH METHOD

The model that will be used in this study will be presented in this section in the task of summarizing text on large volumes of data. The data set that has been provided will be trained with the T5 model and then its performance will be evaluated with the ROUGE technique after that optimization will be carried out using Bayesian. ROUGE is basically a process of evaluation metrics specifically used to see the accuracy of the resulting summaries, ROUGE is very relevant because of the focus on similarities between text data and the summaries produced by this study using ROUGE-N and ROUGE-L. In optimizing the use of Bayesian, it will process architectural parameters and training and testing hyperparameters, then process the prior distribution of each parameter, then perform probabilistic formation. The following is the proposed model through the research architecture contained in Figure 2.

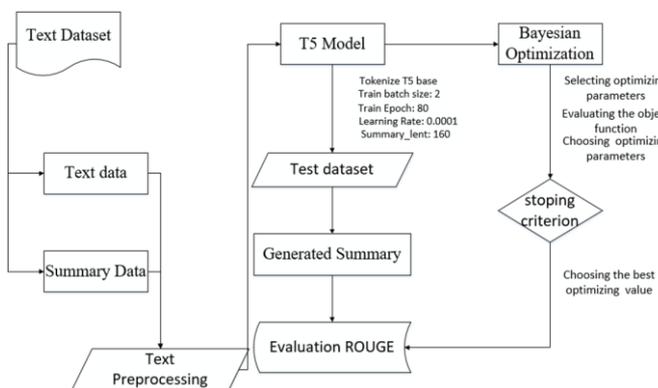


Figure 2. The proposed architecture

Based on the proposed research architecture, it will use a training data percentage of 80% of the total data and 20% is used for testing, using the T5 model has a lot of knowledge and the number of parameters, in this study the epoch value, batch size and learning rate will be determined. The data has an average of 150 characters and a summary length of 24 words.

#### 5. RESULT AND DISCUSSION

In this section, we will discuss the results of implementing Bayesian optimization in the T5 model with the task of performing text summaries. This research will use the Python programming language and libraries related to the T5 model. In the performance evaluation section of this study, techniques from ROUGE will be used. This research will summarize the research data of Politeknik Negeri Medan.

##### 5.1 Text visualization

Text data visualization is data that will be used in applying text summarization as shown in Figure 3.

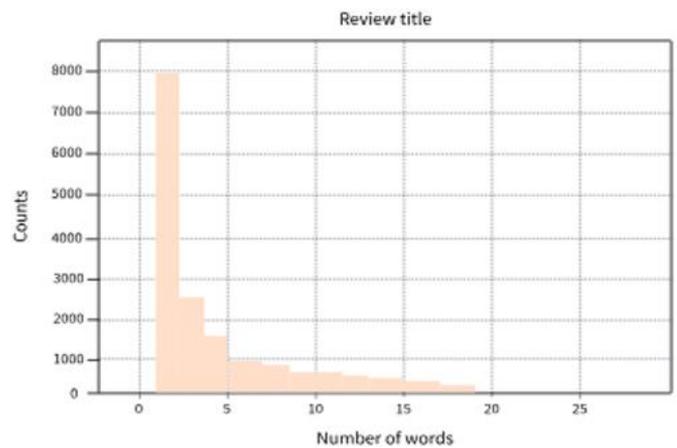


Figure 3. Visualization of text data

##### 5.2 Text summary visualization

The text summary visualization is the data that will be used in applying the text summarization as shown in Figure 4.

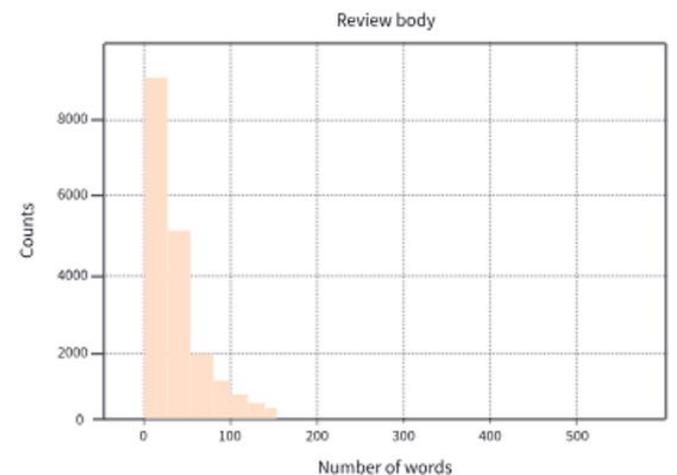


Figure 4. Visualization of the text summary

**Table 1.** Results summary model T5

Text Data	Summary	Model T5 Results
Bawang goreng merupakan salah satu bentuk olahan dari bawang merah yang dalam proses pembuatannya melalui tahap pengirisan. Tahap pengirisan dalam pembuatan bawang goreng di usaha industri kecil masih dilakukan secara manual. Pengirisan bawang merah secara manual memiliki kelemahan seperti memerlukan banyak waktu dan tenaga kerja. Untuk meningkatkan efektivitas dan efisiensi dari penggunaan waktu dan tenaga kerja dalam proses	Bawang goreng ialah bentuk olahan dari bawang merah yang diiris. Tahap pengirisan dalam pembuatan bawang goreng di usaha industri kecil masih dilakukan secara manual. Pengirisan bawang merah secara manual memiliki kelemahan seperti memerlukan banyak waktu dan tenaga kerja.	Bawang goreng dibentuk dengan olahan dari bawang merah yang di iris yang masih untuk industri kecil
Perkembangan teknologi yang semakin pesat, menyebabkan masyarakat menjadikan telepon seluler sebagai kebutuhan primer dalam kehidupan sehari-hari. SMS (Short Message Service) adalah layanan yang disediakan oleh operator seluler untuk mengirim dan menerima pesan singkat	Perkembangan teknologi semakin cepat yang menjadikan smartphone kebutuhan utama dalam kehidupan, salah satu layanan pada smartphone yaitu sms yang dapat menerima pesan	Smartphone merupakan kebutuhan utama dalam kehidupan untuk hal komunikasi seiring dengan perkembangan teknologi
Pohon keputusan (decision tree) adalah salah satu metode dalam analisis klasifikasi. Salah satu algoritma dalam machine learning yang dapat digunakan untuk melakukan analisis klasifikasi dengan metode decision tree yaitu metode Gradient Boosting Machine (GBM), GBM adalah algoritma ensemble yang bekerja dengan cara memperkecil kesalahan model secara bertahap	Pohon keputusan ialah metode dalam melakukan klasifikasi, ada algoritma yang sering digunakan dalam pohon keputusan untuk melakukan klasifikasi seperti decision tree yaitu gradient boosting machine.	Dalam melakukan klasifikasi pada decision tree terdapat algoritma gradient boosting machine
Kartu identitas yang sudah dilengkapi radio frequency identity (RFID) dapat dijadikan alat pendukung sistem kehadiran. Sistem ini terdiri dari unit baca dan unit server. Unit baca bertugas untuk membaca UID dari kartu RFID untuk	Kartu identitas dilengkapi RFID untuk pendukung sistem kehadiran yang terdiri dari unit baca dan unit server	Dalam mendukung sistem kehadiran terhadap alat RFID yang terdapat unit server

**5.3 T5 model**

In the T5 model in carrying out text summary tasks there will be an evaluation using ROUGE, using the T5 model produces results which can be seen in Table 1.

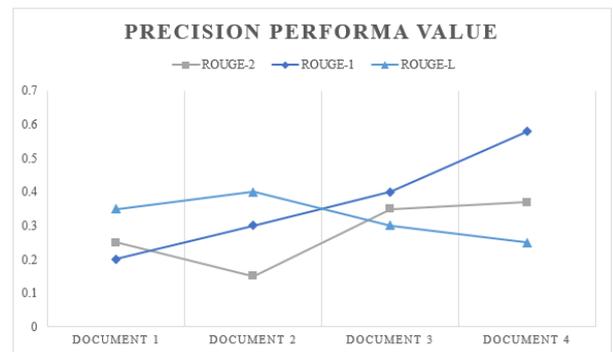
Table 1 shows the results of the T5 model which can perform text summaries based on text data that has gone through the training process, then the T5 model in carrying out text summary tasks is evaluated using the ROUGE technique which produces an evaluation value with the interpretation of the value for ROUGE-N within the scope of ROUGE-1 (Unigram) there is a score range of 0-1 where if 0 means there is no matching unigram and if 1 there is a match between unigram and summary text data, ROUGE-2 (Bigram) where if a value of 0 means that there is no bigram that matches and if the value is 1 there is a match between the bigram and the summary text data and ROUGE-L there is a range of 0-1 with an interpretation if a value of 0 means that there is no subseries that match the summary and if a value of 1 means that there are a similarity between the subseries and the summary data. The following are the results of the evaluation using ROUGE-N and ROUGE-L which are in Table 2.

**Table 2.** ROUGE model T5

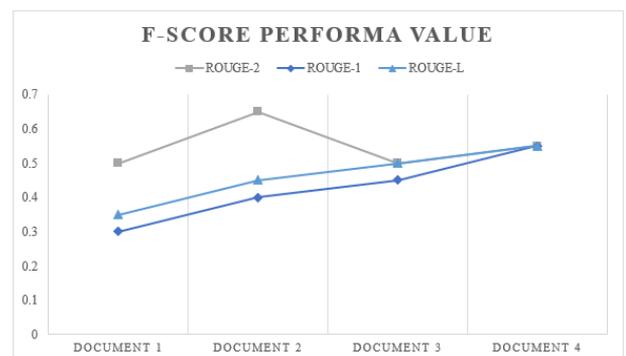
Summary	Results	ROUGE-1	ROUGE-2	ROUGE-L
Document 1	Text summary	0.3	0.5	0.35
Document 2	Text summary	0.4	0.65	0.45
Document 3	Text summary	0.45	0.5	0.5
Document 4	Text summary	0.55	0.55	0.55

Based on the results of Table 2, it shows that there are values for ROUGE-1, ROUGE-2, and ROUGE-L. This evaluation is shown in the following performance visualization of precision and F-score.

After the precision value is shown in Figure 5, an evaluation will be assessed in terms of the F-score with the results shown in Figure 6.



**Figure 5.** Precision performance value



**Figure 6.** F-score performance value

**5.4 Model T5 with Bayesian optimization**

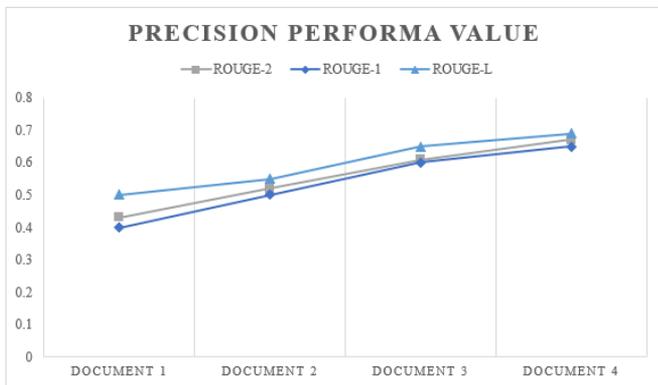
In this section, we will discuss the results of implementing Bayesian optimization in the T5 model with the task of performing text summaries. This research will use the Python

programming language and related libraries to apply Bayesian optimization features to the T5 model. The following are the results of the Bayesian optimization performance for the T5 model which are listed in Table 3.

**Table 3.** ROUGE model T5 with Bayesian optimization

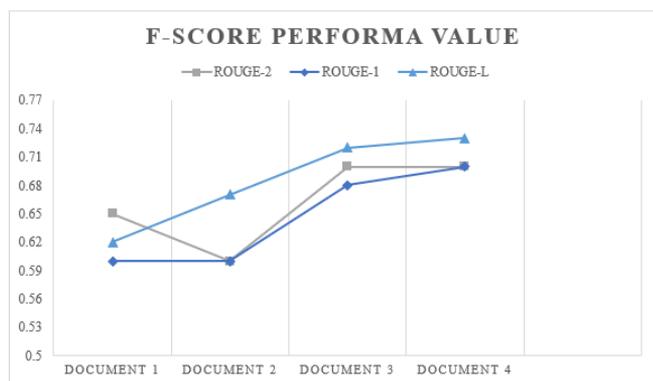
Summary	Results	ROUGE-1	ROUGE-2	ROUGE-L
Document 1	Text summary	0.4	0.43	0.5
Document 2	Text summary	0.5	0.52	0.55
Document 3	Text summary	0.6	0.61	0.65
Document 4	Text summary	0.65	0.67	0.69

Based on the results of Table 3, shows that there are values of ROUGE-1, ROUGE-2, and ROUGE-L for the T5 model by applying Bayesian optimization. This evaluation is shown in the performance visualization of precision and F-score in Figures 7.



**Figure 7.** Precision performance value with Bayesian optimization

After the precision value is shown in Figure 5, an evaluation will be assessed in terms of the F-score with the results shown in Figure 8.



**Figure 8.** F-score performance value with Bayesian optimization

### 5.5 Discussion

In this section, the results obtained from the application of the model in conducting text summaries with the T5 model

will use Bayesian optimization, based on the results obtained with the ROUGE evaluation technique, an increased evaluation value is produced. The T5 model produces a ROUGE-1 value with an average of 0.42, ROUGE-2 has a value of 0.55 and ROUGE-L has a value of 0.46 while applying Bayesian optimization produces a ROUGE-1 evaluation with an average value of 0.53 ROUGE-2 has a value of 0.55 and ROUGE-L has a value of 0.59. based on these results it was concluded that the use of Bayesian optimization can increase the evaluation value of ROUGE, Bayesian optimization uses parameter selection and uses a function that evaluates words that are not valuable and then performs phrases from these words which will then be included in the text summary. The results of this study prove that the T5 model can improve its performance in summarizing text using Bayesian optimization.

### 6. CONCLUSION

This section will discuss the results obtained from the application of the T5 model in carrying out text summary tasks and utilizing Bayesian optimization to improve text summary performance. This study concludes that the T5 model works well in carrying out text summary tasks with an average value of the ROUGE evaluation of 0.4. however, this study conducted experiments with the aim of improving the model using Bayesian optimization in carrying out text summary tasks, the results obtained have good performance, this is evidenced by the value of the ROUGE evaluation of 0.5. has increased from the T5 model without Bayesian optimization. In the T5 model, Bayesian optimization is tasked with finding hyperparameter values, fine-tuning, and optimal domain adaptation such as learning rate values, batch size values, number of epochs, and embedding values. By using Bayesian optimization the T5 model can find the optimal value of the parameters used. This research can be developed in a broader field, especially those that require a text summary model for documents that require quick summaries, in the future the use of Bayesian optimization within the scope of practitioners can improve model performance, better generalization, and increase automation systems. In this study, of course, there are several problems that will be faced and have limitations, including complexity and resource requirements that will carry out the computational process with significant optimization calculations, data requirements in the application of Bayesian require sufficient data so that optimal parameters can be found and overfitting and uncertainty. which is always a big challenge in implementing the T5 model with Bayesian.

### ACKNOWLEDGMENT

We would like to thank Politeknik Negeri Medan, for supporting this research through grant no B/319/PL5/PT.01/04/2023.

### REFERENCES

[1] Lubis, A.R., Nasution, M.K., Sitompul, O.S., Zamzami, E.M. (2021). The effect of the TF-IDF algorithm in times series in forecasting word on social media. Indonesian Journal of Electrical Engineering and Computer Science,

- 22(2): 976-984. <https://doi.org/10.11591/ijeecs.v22.i2.pp976-984>
- [2] Pejić Bach, M., Krstić, Ž., Seljan, S., Turulja, L. (2019). Text mining for big data analysis in financial sector: A literature review. *Sustainability*, 11(5): 1277. <https://doi.org/10.3390/su11051277>
- [3] Hassani, H., Beneki, C., Unger, S., Mazinani, M.T., Yeganegi, M.R. (2020). Text mining in big data analytics. *Big Data and Cognitive Computing*, 4(1): 1. <https://doi.org/10.3390/bdcc4010001>
- [4] Lubis, A.R., Prayudani, S., Fatmi, Y., Nugroho, O. (2022). Classifying news based on Indonesian news using LightGBM. In 2022 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM), Surabaya, Indonesia, pp. 162-166. <https://doi.org/10.1109/CENIM56801.2022.10037401>
- [5] Lubis, A.R., Prayudani, S., Lubis, M., Nugroho, O. (2022). Sentiment analysis on online learning during the covid-19 pandemic based on opinions on twitter using KNN method. In 2022 1st International Conference on Information System & Information Technology (ICISIT), Yogyakarta, Indonesia, pp. 106-111. <https://doi.org/10.1109/ICISIT54091.2022.9872926>
- [6] Lubis, A.R., Nasution, M.K., Sitompul, O.S., Zamzami, E.M. (2020). A framework of utilizing big data of social media to find out the habits of users using keyword. In Proceedings of the 8th International Conference on Computer and Communications Management, Singapore, pp. 140-144. <https://doi.org/10.1145/3411174.3411195>
- [7] Lubis, A.R., Nasution, M.K., Sitompul, O.S., Zamzami, E.M. (2023). A new approach to achieve the users' habitual opportunities on social media. *IAES International Journal of Artificial Intelligence*, 12(1): 41-47. <https://doi.org/10.11591/ijai.v12.i1.pp41-47>
- [8] El-Kassas, W.S., Salama, C.R., Rafea, A.A., Mohamed, H.K. (2021). Automatic text summarization: A comprehensive survey. *Expert Systems with Applications*, 165: 113679. <https://doi.org/10.1016/j.eswa.2020.113679>
- [9] Liang, Z., Du, J., Li, C. (2020). Abstractive social media text summarization using selective reinforced Seq2Seq attention model. *Neurocomputing*, 410: 432-440. <https://doi.org/10.1016/j.neucom.2020.04.137>
- [10] Ma, S., Sun, X., Xu, J., Wang, H., Li, W., Su, Q. (2017). Improving semantic relevance for sequence-to-sequence learning of Chinese social media text summarization. arXiv preprint arXiv:1706.02459.
- [11] Wei, B., Ren, X., Zhang, Y., Cai, X., Su, Q., Sun, X. (2019). Regularizing output distribution of abstractive Chinese social media text summarization for improved semantic consistency. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 18(3): 31. <https://doi.org/10.1145/3314934>
- [12] Abdulateef, S., Khan, N.A., Chen, B., Shang, X. (2020). Multidocument Arabic text summarization based on clustering and Word2Vec to reduce redundancy. *Information*, 11(2): 59. <https://doi.org/10.3390/info11020059>
- [13] Gupta, S., Gupta, S.K. (2019). Abstractive summarization: An overview of the state of the art. *Expert Systems with Applications*, 121: 49-65. <https://doi.org/10.1016/j.eswa.2018.12.011>
- [14] Lin, H., Ng, V. (2019). Abstractive summarization: A survey of the state of the art. In Proceedings of the AAAI Conference on Artificial Intelligence, 33(1): 9815-9822. <https://doi.org/10.1609/aaai.v33i01.33019815>
- [15] Gehrman, S., Deng, Y., Rush, A.M. (2018). Bottom-up abstractive summarization. arXiv preprint arXiv:1808.10792.
- [16] González, J.Á., Segarra, E., García-Granada, F., Sanchis, E., Hurtado, L.F. (2023). Attentional extractive summarization. *Applied Sciences*, 13(3): 1458. <https://doi.org/10.3390/app13031458>
- [17] Bacco, L., Cimino, A., Dell'Orletta, F., Merone, M. (2021). Explainable sentiment analysis: A hierarchical transformer-based extractive summarization approach. *Electronics*, 10(18): 2195. <https://doi.org/10.3390/electronics10182195>
- [18] Al-Laith, A., Shahbaz, M., Alaskar, H.F., Rehmat, A. (2021). Arasencorpus: A semi-supervised approach for sentiment annotation of a large arabic text corpus. *Applied Sciences*, 11(5): 2434. <https://doi.org/10.3390/app11052434>
- [19] Krommyda, M., Rigos, A., Bouklas, K., Amditis, A. (2021). An experimental analysis of data annotation methodologies for emotion detection in short text posted on social media. *Informatics*, 8(1): 19. <https://doi.org/10.3390/informatics8010019>
- [20] Hananto, V.R., Serdült, U., Kryssanov, V. (2022). A text segmentation approach for automated annotation of online customer reviews, based on topic modeling. *Applied Sciences*, 12(7): 3412. <https://doi.org/10.3390/app12073412>
- [21] Mastropaolo, A., Scalabrino, S., Cooper, N., Palacio, D.N., Poshyvanyk, D., Oliveto, R., Bavota, G. (2021). Studying the usage of text-to-text transfer transformer to support code-related tasks. In 2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE), Madrid, pp. 336-347. <https://doi.org/10.1109/ICSE43902.2021.00041>
- [22] Fendji, J.L.E.K., Taira, D.M., Atemkeng, M., Ali, A.M. (2021). WATS-SMS: A T5-based french wikipedia abstractive text summarizer for SMS. *Future Internet*, 13(9): 238. <https://doi.org/10.3390/fi13090238>
- [23] Blekanov, I.S., Tarasov, N., Bodrunova, S.S. (2022). Transformer-based abstractive summarization for Reddit and Twitter: Single posts vs. comment pools in three languages. *Future Internet*, 14(3): 69. <https://doi.org/10.3390/fi14030069>
- [24] Chan, A.S., Fachrizal, F., Lubis, A.R. (2020). Outcome prediction using Naïve Bayes algorithm in the selection of role hero mobile legend. *Journal of Physics: Conference Series*, 1566(1): 012041. <https://doi.org/10.1088/1742-6596/1566/1/012041>
- [25] Paul, D., Goswami, A.K., Chetri, R.L., Roy, R., Sen, P. (2021). Bayesian optimization-based gradient boosting method of fault detection in oil-immersed transformer and reactors. *IEEE Transactions on Industry Applications*, 58(2): 1910-1919. <https://doi.org/10.1109/TIA.2021.3134140>
- [26] Greenhill, S., Rana, S., Gupta, S., Vellanki, P., Venkatesh, S. (2020). Bayesian optimization for adaptive experimental design: A review. *IEEE Access*, 8: 13937-13948. <https://doi.org/10.1109/ACCESS.2020.2966228>

- [27] Lubis, A.R., Nasution, M.K.M., Sitompul, O.S., Zamzami, E.M. (2022). The feature extraction for classifying words on social media with the Naïve Bayes algorithm. *IAES International Journal of Artificial Intelligence*, 11(3): 1041-1048. <https://doi.org/10.11591/ijai.v11.i3.pp1041-1048>
- [28] Rofiq, R.A. (2021). Indonesian news extractive text summarization using latent semantic analysis. In 2021 International Conference on Computer Science and Engineering (IC2SE), Padang, Indonesia, pp. 1-5. <https://doi.org/10.1109/IC2SE52832.2021.9792010>
- [29] Moratanch, N., Chitrakala, S. (2017). A survey on extractive text summarization. In 2017 International Conference on Computer, Communication and Signal Processing (ICCCSP), Chennai, India, pp. 1-6. <https://doi.org/10.1109/ICCCSP.2017.7944061>
- [30] Khan, A., Salim, N., Farman, H., Khan, M., Jan, B., Ahmad, A., Ahmed, I., Paul, A. (2018). Abstractive text summarization based on improved semantic graph approach. *International Journal of Parallel Programming*, 46: 992-1016. <https://doi.org/10.1007/s10766-018-0560-3>
- [31] Masum, A.K.M., Abujar, S., Talukder, M.A.I., Rabby, A.S.A., Hossain, S.A. (2019). Abstractive method of text summarization with sequence to sequence RNNs. In 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, pp. 1-5. <https://doi.org/10.1109/ICCCNT45670.2019.8944620>
- [32] Elsaid, A., Mohammed, A., Ibrahim, L.F., Sakre, M.M. (2022). A comprehensive review of Arabic text summarization. *IEEE Access*, 10: 38012-38030. <https://doi.org/10.1109/ACCESS.2022.3163292>
- [33] Hailu, T.T., Yu, J., Fantaye, T.G. (2020). A framework for word embedding based automatic text summarization and evaluation. *Information*, 11(2): 78. <https://doi.org/10.3390/info11020078>
- [34] La Quatra, M., Cagliero, L. (2022). BART-IT: An efficient sequence-to-sequence model for Italian text summarization. *Future Internet*, 15(1): 15. <https://doi.org/10.3390/fi15010015>
- [35] Patwardhan, N., Marrone, S., Sansone, C. (2023). Transformers in the real world: A survey on NLP applications. *Information*, 14(4): 242. <https://doi.org/10.3390/info14040242>
- [36] Cheng, H.Y., Yu, C.C. (2022). Scene classification, data cleaning, and comment summarization for large-scale location databases. *Electronics*, 11(13): 1947. <https://doi.org/10.3390/electronics11131947>
- [37] Mars, M. (2022). From word embeddings to pre-trained language models: A state-of-the-art walkthrough. *Applied Sciences*, 12(17): 8805. <https://doi.org/10.3390/app12178805>
- [38] Wong, M.F., Guo, S., Hang, C.N., Ho, S.W., Tan, C.W. (2023). Natural language generation and understanding of big code for AI-assisted programming: A review. *Entropy*, 25(6): 888. <https://doi.org/10.3390/e25060888>
- [39] Zhao, H., Zhang, W., Huang, M., Feng, S., Wu, Y. (2023). A multi-granularity heterogeneous graph for extractive text summarization. *Electronics*, 12(10): 2184. <https://doi.org/10.3390/electronics12102184>
- [40] Niculescu, M.A., Ruseti, S., Dascalu, M. (2022). RoSummary: Control tokens for romanian news summarization. *Algorithms*, 15(12): 472. <https://doi.org/10.3390/a15120472>
- [41] Lin, C.S., Jwo, J.S., Lee, C.H. (2023). Adapting static and contextual representations for policy gradient-based summarization. *Sensors*, 23(9): 4513. <https://doi.org/10.3390/s23094513>
- [42] Yang, K., Al-Sabahi, K., Xiang, Y., Zhang, Z. (2018). An integrated graph model for document summarization. *Information*, 9(9): 232. <https://doi.org/10.3390/info9090232>
- [43] Ahuir, V., Hurtado, L.F., González, J.Á., Segarra, E. (2021). Nasca and nases: Two monolingual pre-trained models for abstractive summarization in Catalan and Spanish. *Applied Sciences*, 11(21): 9872. <https://doi.org/10.3390/app11219872>
- [44] Chouikhi, H., Alsuhaibani, M. (2022). Deep transformer language models for Arabic text summarization: A comparison study. *Applied Sciences*, 12(23): 11944. <https://doi.org/10.3390/app122311944>
- [45] Leitao Martins, L.F., Provencher, P.R., Brochu, M., Brochu, M. (2021). Effect of platform temperature and post-processing heat treatment on the fatigue life of additively manufactured AlSi7Mg alloy. *Metals*, 11(5): 679. <https://doi.org/10.3390/met11050679>
- [46] Garg, A., Adusumilli, S., Yenneti, S., Badal, T., Garg, D., Pandey, V., Nigam, A., Gupta, Y., Mittal, G., Agarwal, R. (2021). NEWS article summarization with pretrained transformer. In *Advanced Computing: 10th International Conference*, Panaji, Goa, India, pp. 203-211. [https://doi.org/10.1007/978-981-16-0401-0\\_15](https://doi.org/10.1007/978-981-16-0401-0_15)
- [47] Rothe, S., Maynez, J., Narayan, S. (2021). A thorough evaluation of task-specific pretraining for summarization. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, Punta Cana, Dominican Republic, pp. 140-145. <http://dx.doi.org/10.18653/v1/2021.emnlp-main.12>
- [48] Kolar, D., Lisjak, D., Pająk, M., Gudlin, M. (2021). Intelligent fault diagnosis of rotary machinery by convolutional neural network with automatic hyper-parameters tuning using Bayesian optimization. *Sensors*, 21(7): 2411. <https://doi.org/10.3390/s21072411>
- [49] Moro, G., Ragazzi, L., Valgimigli, L., Frisoni, G., Sartori, C., Marfia, G. (2023). Efficient memory-enhanced transformer for long-document summarization in low-resource regimes. *Sensors*, 23(7): 3542. <https://doi.org/10.3390/s23073542>
- [50] Mahmoudi, O., Bouami, M.F., Badri, M. (2022). Arabic language modeling based on supervised machine learning. *Revue d'Intelligence Artificielle*, 36(3): 467-473. <https://doi.org/10.18280/ria.360315>