



## Indonesian News Text Summarization Using MBART Algorithm

Rahma Hayuning Astuti<sup>1</sup>, Muljono<sup>2\*</sup>, Sutriawan<sup>3</sup>

<sup>1,2</sup>Department of Computer Science, Universitas Dian Nuswantoro, Indonesia

<sup>3</sup>Department of Computer Science, Universitas Muhammadiyah Bima, Indonesia

### Abstract.

**Purpose:** Technology advancements have led to the production of a large amount of textual data. There are numerous locations where one can find textual information sources, including blogs, news portals, and websites. Kompas, BBC, Liputan 6, CNN, and other news portals are a few websites that offer news in Indonesian. The purpose of this study was to explore the effectiveness of using mBART in text summarization for Bahasa Indonesia.

**Methods:** This study uses mBART, a transformer architecture, to perform fine-tuning to generate news article summaries in Bahasa Indonesia. Evaluation was conducted using the ROUGE method to assess the quality of the summaries produced.

**Results:** Evaluation using the ROUGE metric showed better results, with ROUGE-1 of 35.94, ROUGE-2 of 16.43, and ROUGE-L of 29.91. However, the performance of the model is still not optimal compared to existing models in text summarization for another language.

**Novelty:** The novelty of this research lies in the use of mBART for text summarization, specifically adapted for Bahasa Indonesia. In addition, the findings also contribute to understanding the challenges and opportunities of improving text summarization techniques in the Indonesian context.

**Keywords:** Abstractive text summarization, MBART, ROUGE, News

**Received** December 2023 / **Revised** February 2024 / **Accepted** February 2024

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).



### INTRODUCTION

Current technological changes have resulted in online users facing information overload due to the rapidly growing amount of textual information on websites, making it difficult to read through the information [1]. Web-derived sources on the internet, such as blogs, social media networks, news, and so on, are a huge source of textual data [2]. There are many websites that provide news in Indonesia, such as Kompas, BBC, Liputan 6, CNN, and so on. These media produce news and articles every day [3]. More and more online documents require summarization in order to help online users understand information. Text summarization of online documents is done so that users do not spend time looking for the information needed [4].

Text summarization is the process of summarizing a long text into a short text while maintaining the main idea. In natural language processing (NLP) and information retrieval, automatic text summarization is one of the fundamental tasks [5]. Text summarization can be applied in various industrial fields, such as news, aggregators, blogs, product descriptions, and others [6]. Text summarization can make it easier for search engines to search for content compared to searching in full text. Digital businesses such as e-commerce can also benefit from text summarization to display a brief description of the product. Text summarization can also help journalists display news headlines [7].

Text summarization can be classified into three categories: extractive, abstractive, and hybrid. Extractive summarization is done by finding important parts of the content and forming a subset of sentences from sentences contained in the original document [8]. Extractive summarization does not add words to existing content and cannot combine two or more sentences to summarize content. Extractive summarization works on the basis of combining words or phrases from the corpus for summary [6], [9], [10]. Hybrid summarization combines extractive with abstractive. The method has the drawback of producing lower-quality abstractive summaries compared to the pure abstractive approach [2], [11], [12].

---

\* Corresponding author.

Email addresses: [muljono@dsn.dinus.ac.id](mailto:muljono@dsn.dinus.ac.id) (Muljono)

DOI: [10.15294/sji.v11i1.49224](https://doi.org/10.15294/sji.v11i1.49224)

Abstractive summarization performs summarization by understanding the given sentence and developing relevant summary sentences on its own. Abstractive summarization is also more flexible in generating summaries [13], [14], [15]. Unlike extractive summarization, which can produce poor sentences, abstractive summarization can produce grammatically correct sentences [6], [16], [17]. The abstractive method paraphrases and rearranges sentences into a summary [18], [19]. In this research, we will use the abstractive summarization method. This is due to the advantages possessed by the abstractive method.

In using the abstractive summarization method, there are various algorithms that can be used, one of which is the multilingual version of the BART algorithm. BART is one of the pre-trained systems based on transformer architecture [20],[21]. Currently, there is a multilingual version of BART, or what can be called mBART. One of the languages that can be processed is Indonesian. The use of mBART in text summarization can produce a good model. Some research on the use of mBART has been conducted in several languages, such as Russian [22],[23], Vietnamese [24],[25],[26] and various other languages. From some of these studies, the evaluation results can produce good values, such as in Vietnamese language research [24], [26], [27] which get a rough-value of 55.21, a rough- 2 of 25.69, and a rough- L of 37.33 for a dataset called WikiLingua, and for the Vietnews dataset, a rough-value of 59.81, a rough- 2 of 28.28, and a rough- L of 38.71. In this study, research was conducted on the use of mBART in the text summarization of Indonesian news.

Further discussion in this paper included the following: Section 2 contains a description of the research method regarding the application of the MART algorithm in the text summarization of Indonesian news. Section 3 contains the results and discussion, followed by the conclusion drawn in Section 4.

## METHODS

### Proposed method

The proposed research method for text summarization using mBART is described in Figure 1, covering the steps of pre-processing, fine-tuning, training, evaluation, and summary prediction. It starts with pre-processing to fine-tune the XL-SUM dataset of Hugging Face: <https://huggingface.co/datasets/csebuetnlp/xlsum/viewer/indonesian> [28],[3]. The method ensures the readiness of the data for the next steps. Using the Google Colab Python tool to retrieve and split the dataset, the next step, fine-tuning the mBART model, was followed by training to optimize model performance. Evaluation, using the ROUGE metric, assesses the quality of the summary against the reference or original text. Before entering the summary prediction stage, the ability of the model to produce short and precise summaries is tested.

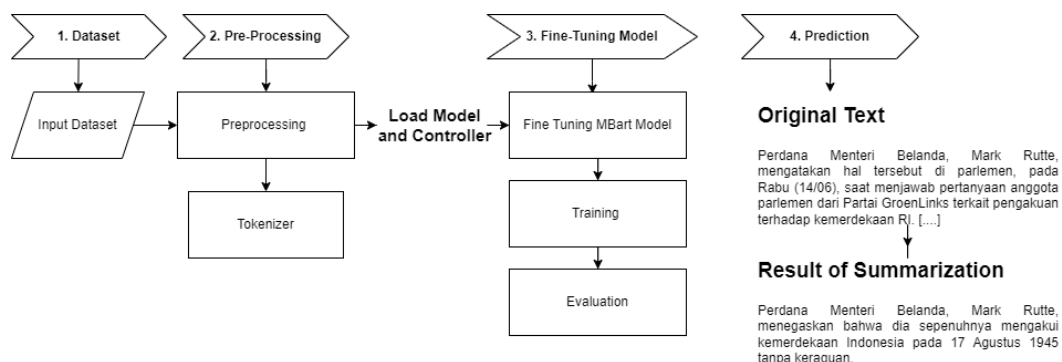


Figure 1. Flowchart of research flow

### Data collection method

#### Data type

The type of data required in this research is qualitative data. Qualitative data is non-numeric data. Text, sentences, words are examples of qualitative data and are needed as datasets for model building. The model created is a model about text summarization of Indonesian news. Therefore, the data needed is in the form of news articles in Indonesian.

### Literature study

In carrying out research, the first step is to combine a number of literature studies or research related to having similar problems or topics. Scientific articles, journals, books are one of the literature sources that can be used. The selection of literature studies is based on the same problem, namely text summarization. The type of text summarization chosen is the abstractive method using mBART.

### Dataset text summarization

The text summarization dataset collection method can be done by several methods. The first method is manually scraping the website and then making a summary manually. The first method is very ineffective because it requires a long time to make datasets. The second method is by asking permission from similar research, namely text summarization, to be used. The third method is to use a public dataset. The third method can get datasets from open dataset websites such as Kaggle, UCI Machine Learning Repository, and so on. In this study, we used the third method, namely the method of using a public dataset, namely XL-SUM, which contains 44 languages, one of which is Indonesian [3]. XL-SUM is a large and diverse dataset that includes 1 million pairs of professionally annotated article summaries taken from the BBC using a series of carefully designed heuristics. The dataset used is only the Indonesian language part of XL-SUM. The selection of the XL-SUM dataset is due to the ease of using and accessing the dataset, namely by accessing it through the library on the HuggingFace website.

### Evaluation of results

In this study, automatic evaluation of results was carried out using ROUGE in accordance with previous research on the same topic. ROUGE (Recall-Oriented Understudy for Gisting Evaluation) is a package that includes several automatic evaluation methods that calculate the similarity between summaries. There are 4 types of ROUGE calculations, namely ROUGE-N, ROUGE-L, ROUGE-W, and also ROUGE-S [29]. In this research, the types of ROUGE which were used are ROUGE-N (ROUGE-1, ROUGE-2) and ROUGE-L. The results of ROUGE used as an evaluation of the model that has been made and as a comparison with previous research.

## RESULTS AND DISCUSSIONS

### Dataset preparation

The dataset used in this research is a dataset in the form of Indonesian news articles. The dataset used is XL-SUM [3]. XL-SUM has 44 article languages, with one of them being Indonesian. In the XL-SUM dataset, the dataset has been divided into train, test, and validation. There are 47,800 Indonesian-language datasets in XL-SUM, with a division of datasets for training of 38,200 and testing and validation of 4780. The dataset contains articles from BBC News. The dataset contains an ID, a URL, a title, a summary, and text, as shown in Table 1.

Table 1. Dataset structure

Category	Text
Id	media-49647079
Url	<a href="https://www.bbc.com/indonesia/media-49647079">https://www.bbc.com/indonesia/media-49647079</a>
Title	<i>Gajah mengamuk saat upacara keagamaan, 17 orang terluka</i>
Summary	<i>Seekor gajah mendadak mengamuk saat prosesi tahunan agama Buddha di Kolombo, Sri Lanka, sehingga membuat peserta upacara tunggang-langgang. Setidaknya 17 orang terluka.</i>
Text	<i>Dilaporkan dua orang terluka cukup serius, sementara sisanya sudah diperbolehkan pulang setelah mendapatkan perawatan. Video gajah yang mengamuk itu kontan viral di media sosial. Si gajah berlari tak tentu arah, menabrak serta menginjak sebagian peserta upacara. Simak juga: Belum diketahui penyebab gajah tersebut tiba-tiba mengamuk. Diduga gajah itu kaget oleh sesuatu di antara para peserta dan pengunjung. Media setempat melaporkan gajah lain yang juga mengamuk di prosesi berbeda. Gajah hias merupakan daya tarik tersendiri dalam upacara keagamaan di Sri Lanka. Bagi warga Sri Lanka, memiliki gajah adalah simbol status. Beberapa kuil di Sri Lanka juga memiliki gajah untuk keperluan upacara.</i>



In table 4.2, after the text has undergone tokenization, there are 3 outputs, namely 'input\_ids', 'attention\_mask', and 'labels'. 'Input\_ids' comes from the 'text' category in the dataset, while 'labels' comes from the 'summary' category. 'Attention\_mask' is an optional argument that is used when merging sequences. There are two values in 'attention\_mask', which are 0 or 1. 1 in 'attention\_mask' indicates tokens that need attention, and 0 indicates tokens that do not need attention. The numeric results in the 'input\_ids' and 'labels' sections are token IDs derived from breaking the input sentence into tokens so that the tokens can be processed at a later stage.

#### ***Loan model and data collator***

Before fine-tuning, the model would be downloaded first, which could be accessed on the HuggingFace website. MBART has several versions available, such as mbart-large-cc25, mbart-large-50, and others. In this research, we used mbart-large-50, which can be used in Indonesian. The model that has been downloaded is 2.44 GB. In addition, the padding process is also carried out using DataCollatorForSeq2Seq. This process is done in order to effectively perform the process of padding because it can be done dynamically to pad the longest sentence in a batch during inspection. The padding process is needed so that the tokens that have been processed before can have the same length, and then the tokens would be entered into the model.

#### ***Fine-tuning model***

The model used is pre-trained using the Transformer architecture. A pre-trained model is one that has been trained in advance with other datasets. In this model, fine-tuning would be done. Fine-tuning is a technique used to adjust the model to a new dataset. MBART50 is one of the pre-trained so that fine-tuning will be done [33], [34].

In the process of fine-tuning this model, there is no change in the model architecture but only an adjustment of the model to the dataset used. This fine-tuning would be adjusted to the dataset used, namely the XL-SUM dataset for the Indonesian part. In this process, batch\_size = 4, learning\_rate = 2e-5, optimization = "adamw\_torch", weight\_decay = 0.01, save\_total\_limit = 3, and num\_train\_epochs = 1. These hyperparameters will be used in the next model training.

#### ***Training model***

After pre-processing the dataset and selecting the hyperparameter model, the next step is the training process. The method used in this study is a pre-trained model based on the architecture of the transformer, namely mBART50. MBART50 is an extension of mBART25, so it has similarities in its model architecture [35]. The mBART25 architecture is based on a sequence-to-sequence transformer architecture with 12 layers in each encoder and decoder and 1024 model dimensions in 16 heads (~680 million parameters). In addition, the mBART25 architecture has an additional normalization layer on top of each encoder and decoder [36], [37]. The difference between the mBART25 architecture and mBART50 is that in mBART50, an embedding layer is added with a randomly initialized vector for an extra set of 25 new language tokens [33].

MBART50, which is multilingual for BART, then in the performance way for fine-tuning summarization, is to copy the information from the input but manipulate it, which is closely related to the purpose of denoising pre-training. In this case, the input encoder is the input sequence, and the decoder produces the output autoregressively [20].

This training process uses Pytorch. The data used in the training process is about 38200 articles for the training part, and the validation dataset is 4780 articles. The training process takes 1 hour and 28 minutes for 1 epoch using the A100 GPU with 40GB of RAM for the GPU on Google Colab Pro. In the training process, it produces the loss in Table 3.

Table 3. Results in loss

<b>Training Loss</b>	<b>Validation Loss</b>
2.136400	2.009753

Model training in the training process runs well. Then the model that has gone through the training process will be stored locally, namely in Google Drive. This storage is done to be able to use the model again during the process of trying to summarize the text.

**Evaluation and comparison**

The method used to evaluate the model is rough. ROUGE, or Recall-Oriented Understudy for Gisting Evaluation, is a set of metrics used in automatic text summarization evaluation [38]. ROUGE consists of several automatic evaluation methods to determine the similarity of summaries [29]. In the results of this study, the roughness used is based on roughness N (roughness 1 and roughness 2) and roughness L. ROUGE: N is a calculation of the number of corresponding n-grams between the text generated by the model and the reference. N-grams are collections derived from tokens or words [39]. A unigram is an n-gram derived from one word. Bigram is an n-gram derived from two consecutive words. While ROUGE-L performs calculations on the longest common subsequences (LCS) between the output derived from the model and the reference [40],[41]. In the research results using mBART50, the resulting ROUGE is in Table 4.

Table 4. Fine-tuning results of mBART50

Model	ROUGE - 1	ROUGE - 2	ROUGE - L
MBART50	35.94	16.43	29.91

Previously, one of the hyperparameters, learning\_rate, was also tested to get the best results. In this test, 3 types of learning rates were carried out, namely 1e-4, 1e-5, and 2e-5.

Table 5. Comparison results with learning rate

Learning_rate	ROUGE - 1	ROUGE - 2	ROUGE - L
1e-4	34.76	15.70	29.05
1e-5	35.21	15.94	29.36
2e-5	35.94	16.43	29.91

From these results, it can be seen that the difference is not far between one learning rate and another, so the best results are taken from using a learning rate of 2e-5 to compare with other model comparisons and benchmarks.

In this research, experiments were also conducted using another similar algorithm, mT5. MT5 is a multilingual model with an encoder-decoder architecture based on T5. In mT5, there are 5 versions of the model, namely small (≈ 300 million parameters), base (580 million), large (1.2 billion), XL (3.7 billion), and XXL (13 billion) [42]. In conducting the comparison for this experiment, the mT5 version used is mT5 base [3],[43]. The MT5 base is used because of the size of the model before fine-tuning, which is almost balanced with the mBART 50, which is 2.33 GB.

Table 6. Comparison between mBART50 and mT5 base

Model	Model Size	R - 1	R - 2	R - L
MBART50	2.44 GB	35.94	16.43	29.91
MT5 base	2.33 GB	24.52	8.85	20.59

From the results of the table, it can be concluded that mBART50 has better results than mT5 base, with a difference of 11.4272 for ROUGE 1, ROUGE 2 of 7.5777, and ROUGE L of 9.319, even though their model sizes are almost the same. In addition, a comparison is also made with the results of benchmark research using similar algorithms.

Table 7. Comparison of results with benchmark

Model	Language	Dataset	R-1	R-2	R-L
mBART25	AR [43]	AAW	29.5	14.4	26.0
		AFP	54.8	37.3	52.8
		AHR	39.1	26.1	37.5
		HYT	32.0	16.2	29.3
		NHR	31.0	16.2	29.2
		QDS	62.4	54.1	61.7
		XIN	65.1	53.4	64.2
		MIX	39.0	25.6	37.1

		XL-S	32.1	12.5	27.6
		XL-T	29.8	11.7	26.9
		Macro Averages	41.5	28.1	39.2
mBART	IT [44]	MLSum-It	19.3	6.4	16.3
		Fanpage	36.5	17.4	26.1
		IlPost	38.9	21.3	32.0
mBART	VI [24]	WikiLingua	55.2	25.6	37.3
		Vietnews	59.8	28.2	38.7
mBART	RU [45]	RIA	42.8	25.5	39.9
mBART	RU [22]	Gazeta	32.1	14.2	27.9
<u>mBART50</u>	<u>ID</u>	<u>XL-SUM (ID)</u>	<u>35.9</u>	<u>16.4</u>	<u>29.9</u>

Based on Table 7, the code AR refers to Arabic, IT to Italian, VI to Vietnamese, RU to Russian, and ID to Indonesian. Table 6 also shows that the ROUGE-1, ROUGE-2, and ROUGE-L scores obtained by researchers are still relatively low and have not exceeded some other studies. For example, AR XL-S was trained using the mBAR25 model and achieved ROUGE evaluation scores of R1 = 32.1, R2 = 12.5, and RL = 27.6. However, when trained with the XL-T dataset, the ROUGE score remained around R1 = 29.8, R2 = 11.7, and RL = 26.9 [43]. Likewise, IT MLSum-It achieved ROUGE scores of R1 = 19.3, R2 = 6.4, and R3 = 16.3 by using the mBART model [44]. Another study that tested the mBART model with the RU Gazeta dataset (Russian language) obtained ROUGE evaluation results with values of R1 = 32.1, R2 = 14.2, and RL = 27.9 [22]. Our evaluation of the three AR language models with the XL-S and XL-T datasets, the IT language model with the MLSum-It dataset, and the RU language model with the Gazeta dataset shows that each trained model faces various problems caused by several factors. These factors include the pre-processing stage, which only uses tokenizers and data collectors as padding, and the use of only one epoch for training. These constraints are triggered by resource limitations such as limited GPU usage on Google Colab Pro and the number of datasets that affect the final ROUGE score of the model.

We evaluate and compare prediction models that have been trained and tested to summarize news articles in Bahasa Indonesia using the Google Colab Pro tool. Our proposed model, mBART50, was trained using the XL-Sum ID dataset, which is a news dataset collected from BBC News Bahasa Indonesia. During the training stage, we found that the evaluation of the mBART50 model is highly dependent on the data used for training, especially at the pre-processing stage to clean up noise and irrelevant words in the documents. At the fine-tuning stage, the mBART50 model can summarize documents well without losing their original meaning, as shown by the evaluation using the ROUGE metric with scores of R1 = 35.9, R2 = 16.4, and R3 = 29.9. However, we also realize that the quality of the summary produced by the mBART Model also depends on the characteristics of the dataset and the type of language used in text summarization, which may affect the overall quality of the summary.

Table 8. Text summary experiment

Url	Original article	Summary produced
<a href="https://www.bbc.com/indonesia/dunia-65911860">https://www.bbc.com/indonesia/dunia-65911860</a>	<p>Perdana Menteri Belanda, Mark Rutte, mengatakan hal tersebut di parlemen, pada Rabu (14/06), saat menjawab pertanyaan anggota parlemen dari Partai GroenLinks terkait pengakuan terhadap kemerdekaan RI.</p> <p>Rutte berjanji akan berkonsultasi dengan Presiden Indonesia, Joko Widodo, untuk mencapai interpretasi bersama tentang hari kemerdekaan itu.</p> <p>"Kami sepenuhnya sudah mengakui 17 Agustus zonder voorbehaud [tanpa keraguan]. Saya masih akan mencari jalan keluar bersama presiden [Indonesia, Joko Widodo] untuk mencari cara terbaik agar bisa diterima kedua pihak," ujar PM Rutte sebagaimana dikutip media Historia.</p> <p>[...]</p>	<p>Perdana Menteri Belanda, Mark Rutte, menegaskan bahwa dia sepenuhnya mengakui kemerdekaan Indonesia pada 17 Agustus 1945 tanpa keraguan.</p>

<p>https://www.cnnindonesia.com/teknologi/20230710091418-199-971437/bunga-bangkai-di-kebun-roya-cibodas-mekar-lagi</p>	<p>Dilansir situs resmi Badan Riset dan Inovasi Nasional (BRIN), bunga ini terlihat mekar sempurna pada pukul 09.30 WIB dengan tinggi dan diameter spata 54 cm. Hampir sama dengan bunga yang sebelumnya mekar pada April 2023 lalu, ketinggian bunga yang saat ini mekar pun tidak mencapai 3 meter. Peneliti Badan Riset dan Inovasi Nasional (BRIN)Destri mengatakan, bunga yang kali ini mekar memiliki pertumbuhan yang tidak maksimal. "Kemungkinan ada penurunan kualitas media sehingga bunga yang mekar kali ini tidak setinggi dengan yang mekar sebelumnya (2015 dan 2019) yaitu mencapai hampir 3 meter," ujar Destri. Selanjutnya, diperlukan media tanam yang bagus agar bunga tumbuh subur. Media tanam yang bagus diharuskan memiliki porositas dan kandungan bahan organik yang sesuai dengan kebutuhan tanaman koleksi.</p> <p>[...]</p>	<p>Bunga bangkai jenis <i>Amorpophallus titanum</i> di Kebun Raya Cibodas, Sumatra Barat, mekar pada Selasa (09/05) pagi waktu setempat.</p>
------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------

The model proposed in this study, mBART50, shows significant results in summarizing Indonesian text. Proven by better evaluation using the ROUGE matrix with values (R1 = 35.9, R2 = 16.4, R3 = 29.9) compared to other models that have been evaluated in previous studies. The mBART model was trained on the XL-Sum ID dataset, which consists of filtered Indonesian news from BBC News. The mBART50 model has a specific approach to text handling in the Indonesian context. Pre-processing that filters out noise and irrelevant words in the documents adds to the quality of the summaries produced. The use of mBART as a multilingual transformer model signifies novelty in this research, while the rough-based evaluation provides high confidence in the validity of the evaluation results. Thus, the mBART50 model not only makes a significant contribution to improving the quality of Indonesian text summarization but also brings novelty by adopting a better approach to the use of language-specific technologies in the text summarization domain.

## CONCLUSION

Based on the research, it can be concluded that the use of mBART in Indonesian text summarization has been explored, resulting in better progress in the development of text summarization models. The evaluation method using ROUGE shows better values with ROUGE-1 of 35.94, ROUGE-2 of 16.43, and ROUGE-L of 29.91. Nevertheless, the performance of the model is still not optimal and has not been able to outperform existing models. Challenges such as performance improvement and resource efficiency remain an important focus for future research. Thus, this research makes an important contribution to the development of Indonesian text summarization and highlights the need for quality and effectiveness improvements in future text summarization methods.

## REFERENCES

- [1] H. Lin and V. Ng, "Abstractive Summarization: A Survey of the State of the Art," *Proc. AAAI Conf. Artif. Intell.*, vol. 33, no. 01, pp. 9815–9822, 2019, doi: 10.1609/aaai.v33i01.33019815.
- [2] W. S. El-Kassas, C. R. Salama, A. A. Rafea, and H. K. Mohamed, "Automatic text summarization: A comprehensive survey," *Expert Syst. Appl.*, vol. 165, p. 113679, 2021, doi: 10.1016/j.eswa.2020.113679.
- [3] T. Hasan *et al.*, "XL-Sum: Large-Scale Multilingual Abstractive Summarization for 44 Languages," *Find. Assoc. Comput. Linguist. ACL-IJCNLP 2021*, pp. 4693–4703, 2021, doi: 10.18653/v1/2021.findings-acl.413.
- [4] Á. Hernández-Castañeda, R. A. García-Hernández, Y. Ledeneva, and C. E. Millán-Hernández, "Language-independent extractive automatic text summarization based on automatic keyword extraction," *Comput. Speech Lang.*, vol. 71, no. May 2021, p. 101267, 2022, doi: 10.1016/j.csl.2021.101267.
- [5] Q. Wang, P. Liu, Z. Zhu, H. Yin, Q. Zhang, and L. Zhang, "A text abstraction summary model based on BERT word embedding and reinforcement learning," *Appl. Sci.*, vol. 9, no. 21, 2019, doi: 10.3390/app9214701.
- [6] P. R. Dedhia, H. P. Pachgade, A. P. Malani, N. Raul, and M. Naik, "Study on Abstractive Text Summarization Techniques," *Int. Conf. Emerg. Trends Inf. Technol. Eng. ic-ETITE 2020*, no. February 2020, 2020, doi: 10.1109/ic-ETITE47903.2020.087.
- [7] 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.
- [8] P. Verma, S. Pal, and H. Om, "A comparative analysis on Hindi and English extractive text summarization," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 18, no. 3, 2019, doi: 10.1145/3308754.
- [9] A. S. Girsang and F. J. Amadeus, "Extractive Text Summarization for Indonesian News Article Using Ant System Algorithm," *J. Adv. Inf. Technol.*, vol. 14, no. 2, pp. 295–301, 2023, doi: 10.12720/jait.14.2.295-301.
- [10] and J. D. H. Vaibhav Gulati, Deepika Kumar, Daniela Elena Popescu, "Extractive Article Summarization Using Integrated TextRank and BM25+ Algorithm Vaibhav," *Electronics*, vol. 12, no. 2, pp. 1–17, 2023, doi: <https://doi.org/10.3390/>.
- [11] W. Etaiwi and A. Awajan, "SemG-TS: Abstractive Arabic Text Summarization Using Semantic Graph Embedding," *Mathematics*, vol. 10, no. 18, 2022, doi: 10.3390/math10183225.
- [12] A. Reda *et al.*, "A Hybrid Arabic Text Summarization Approach based on Transformers," in *2nd International Mobile, Intelligent, and Ubiquitous Computing Conference, MIUCC 2022*, 2022, pp. 56–62. doi: 10.1109/MIUCC55081.2022.9781694.
- [13] Y. Iwasaki, A. Yamashita, Y. Konno, and K. Matsubayashi, "Japanese abstractive text summarization using BERT," *Adv. Sci. Technol. Eng. Syst.*, vol. 5, no. 6, pp. 1674–1682, 2020, doi: 10.25046/AJ0506199.
- [14] P. Kouris, G. Alexandridis, and A. Stafylopatis, "Abstractive text summarization based on deep learning and semantic content generalization," *ACL 2019 - 57th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf.*, pp. 5082–5092, 2020, doi: 10.18653/v1/p19-1501.
- [15] S. Song, H. Huang, and T. Ruan, "Abstractive text summarization using LSTM-CNN based deep learning," *Multimed. Tools Appl.*, vol. 78, no. 1, pp. 857–875, 2019, doi: 10.1007/s11042-018-5749-3.
- [16] R. Rani and D. K. Lobiyal, "An extractive text summarization approach using tagged-LDA based topic modeling," *Multimed. Tools Appl.*, vol. 80, no. 3, pp. 3275–3305, 2021, doi: 10.1007/s11042-020-09549-3.
- [17] R. C. Belwal, S. Rai, and A. Gupta, "A new graph-based extractive text summarization using keywords or topic modeling," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 10, pp. 8975–8990, 2021, doi: 10.1007/s12652-020-02591-x.
- [18] H. Zhang, J. Cai, J. Xu, and J. Wang, "Pretraining-based natural language generation for text summarization," *CoNLL 2019 - 23rd Conf. Comput. Nat. Lang. Learn. Proc. Conf.*, pp. 789–797, 2019, doi: 10.18653/v1/k19-1074.
- [19] S. Lamsiyah, A. El Mahdaouy, S. E. A. Ouatik, and B. Espinasse, "Unsupervised extractive multi-document summarization method based on transfer learning from BERT multi-task fine-tuning," *J. Inf. Sci.*, 2021, doi: 10.1177/0165551521990616.
- [20] M. Lewis *et al.*, "BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension," *Proc. Annu. Meet. Assoc. Comput. Linguist.*, pp. 7871–7880, 2020, doi: 10.18653/v1/2020.acl-main.703.
- [21] D. Uthus, S. Ontañón, J. Ainslie, and M. Guo, "mLongT5: A Multilingual and Efficient Text-To-Text Transformer for Longer Sequences," vol. 5, 2023.
- [22] I. Gusev, "Dataset for Automatic Summarization of Russian News," *Commun. Comput. Inf. Sci.*, vol. 1292 CCIS, pp. 122–134, 2020, doi: 10.1007/978-3-030-59082-6\_9.
- [23] D. Taunk and V. Varma, "Summarizing Indian Languages using Multilingual Transformers based Models," *CEUR Workshop Proc.*, vol. 3395, pp. 435–442, 2022.
- [24] H. Nguyen, L. Phan, J. Anibal, A. Peltekian, and H. Tran, "VieSum: How Robust Are Transformer-based Models on Vietnamese Summarization?," 2021.
- [25] T.-H. Nguyen and T.-N. Do, "Text Summarization on Large-scale Vietnamese Datasets," *J. Inf. Commun. Converg. Eng.*, vol. 20, no. 4, pp. 309–316, 2022, doi: 10.56977/jicce.2022.20.4.309.
- [26] K. Zheng and W. Zheng, "Deep Neural Networks Algorithm for Vietnamese Word Segmentation," *Sci. Program.*, vol. 2022, 2022, doi: 10.1155/2022/8187680.
- [27] Y. Wang, J. Du, J. Kuang, C. Chen, M. Li, and J. Wang, "Two-Scaled Identification of Landscape Character Types and Areas: A Case Study of the Yunnan–Vietnam Railway (Yunnan Section), China," *Sustain.*, vol. 15, no. 7, 2023, doi: 10.3390/su15076173.
- [28] H. Face, "Indonesian News XL-Sum Dataset," 2024.
- [29] C. Y. Lin, "Rouge: A package for automatic evaluation of summaries," *Proc. Work. text Summ. branches out (WAS 2004)*, no. 1, pp. 25–26, 2004.
- [30] J. Zhong and Z. Wang, "MTL-DAS: Automatic Text Summarization for Domain Adaptation,"

- Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/4851828.
- [31] J. D'Silva and U. Sharma, "Automatic text summarization of konkani texts using pre-trained word embeddings and deep learning," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 2, pp. 1990–2000, 2022, doi: 10.11591/ijece.v12i2.pp1990-2000.
- [32] J. Richardson, "SentencePiece: A simple and language independent subword tokenizer and detokenizer for Neural Text Processing," pp. 66–71, 2018.
- [33] Y. Tang *et al.*, "Multilingual Translation with Extensible Multilingual Pretraining and Finetuning," 2020.
- [34] S. Lamsiyah, A. E. Mahdaouy, S. E. A. Ouatik, and B. Espinasse, "Unsupervised extractive multi-document summarization method based on transfer learning from BERT multi-task fine-tuning," *J. Inf. Sci.*, vol. 49, no. 1, pp. 164–182, 2023, doi: 10.1177/0165551521990616.
- [35] A. Ghadimi and H. Beigy, "Hybrid multi-document summarization using pre-trained language models," *Expert Syst. Appl.*, vol. 192, 2022, doi: 10.1016/j.eswa.2021.116292.
- [36] Y. Liu *et al.*, "Multilingual denoising pre-training for neural machine translation," *Trans. Assoc. Comput. Linguist.*, vol. 8, pp. 726–742, 2020, doi: 10.1162/tacl\_a\_00343.
- [37] G. L. A. Babu and S. Badugu, "Deep learning based sequence to sequence model for abstractive telugu text summarization," *Multimed. Tools Appl.*, vol. 82, no. 11, pp. 17075–17096, 2023, doi: 10.1007/s11042-022-14099-x.
- [38] G. Sharma and D. Sharma, "Improving Extractive Text Summarization Performance Using Enhanced Feature Based RBM Method," *Rev. d'Intelligence Artif.*, vol. 36, no. 5, pp. 777–784, 2022, doi: 10.18280/ria.360516.
- [39] S. Ullah and A. B. M. A. Al Islam, "A framework for extractive text summarization using semantic graph based approach," in *6th International Conference on Networking, Systems and Security, NSysS 2019*, 2019, pp. 48–58. doi: 10.1145/3362966.3362971.
- [40] J.-W. He, W.-J. Jiang, G.-B. Chen, Y.-Q. Le, and X.-F. Ding, "Enhancing N-Gram Based Metrics with Semantics for Better Evaluation of Abstractive Text Summarization," *J. Comput. Sci. Technol.*, vol. 37, no. 5, pp. 1118–1133, 2022, doi: 10.1007/s11390-022-2125-6.
- [41] Z. Deng, F. Ma, R. Lan, W. Huang, and X. Luo, "A Two-stage Chinese text summarization algorithm using keyword information and adversarial learning," *Neurocomputing*, vol. 425, pp. 117–126, 2021, doi: 10.1016/j.neucom.2020.02.102.
- [42] L. Xue *et al.*, "mT5: A Massively Multilingual Pre-trained Text-to-Text Transformer," *NAACL-HLT 2021 - 2021 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. Proc. Conf.*, pp. 483–498, 2021, doi: 10.18653/v1/2021.naacl-main.41.
- [43] M. K. Eddine, N. Tomeh, N. Habash, J. Le Roux, and M. Vazirgiannis, "AraBART: a Pretrained Arabic Sequence-to-Sequence Model for Abstractive Summarization," 2022.
- [44] N. Landro, I. Gallo, R. La Grassa, and E. Federici, "Two New Datasets for Italian-Language Abstractive Text Summarization," *Inf.*, vol. 13, no. 5, pp. 1–12, 2022, doi: 10.3390/info13050228.
- [45] A. Bukhtiyarov and I. Gusev, "Advances of Transformer-Based Models for News Headline Generation," *Commun. Comput. Inf. Sci.*, vol. 1292 CCIS, pp. 54–61, 2020, doi: 10.1007/978-3-030-59082-6\_4.