Comparative Analysis of Summarization Methods for Skin Care Product Reviews: A Study on BERT, BART, and T5 Models

Nisrina Arintia Maghfiroh
Faculty of Computer Science
University of Brawijaya
Malang, Indonesia
nisrinaarintiam@student.ub.ac.id

Fitra Abdurrachman Bachtiar
Intelligent Computing and Visualization Laboratory, Faculty of Computer Science
name of organization
University of Brawijaya
Malang, Indonesia
fitra.bachtiar@ub.ac.id

Lailil Muflikhah
Intelligent Computing and Visualization Laboratory, Faculty of Computer Science
University of Brawijaya
Malang, Indonesia
lailil@ub.ac.id

2023 International Conference on Advanced Mechatronics, Intelligent Manufacture and Industrial Automation (ICAMIMIA) | 979-8-3503-0922-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICAMIMIA60881.2023.10427745

***Abstract*— The growth of the skin care industry is driving consumers to look for product review information. The skincare product currently widely used by women and men is facial wash products. On the Female Daily website, the most frequently purchased facial wash products are Cetaphil products. With so many purchases of this product, potential buyers want to try the product by reading the reviews first. However, the large number of reviews is a barrier to accessing relevant information. Therefore, effectively summarizing reviews can solve this problem by presenting key information more concisely and digestibly. This study aims to evaluate the effectiveness of three different summary methods, namely BERT, BART, and T5, in summarizing skin care product reviews. The stages in these three models are explained in detail. Evaluation results using the ROUGE Score method with the Cetaphil facial wash product review dataset show that the BART model consistently has a higher average ROUGE Score than the BERT and T5 models. Specifically, the BART model achieved an average ROUGE-1 of 0.87718, ROUGE-2 of 0.80689, and ROUGE-L of 0.87688, while the BERT model achieved an average ROUGE-1 of 0.84591, ROUGE-2 of 0.77585, and ROUGE- L is 0.84479, and model T5 has an average ROUGE-1 of 0.74995, ROUGE-2 of 0.70048, and ROUGE-L of 0.73023. These results suggest that the generative approach implemented by BART is more effective in addressing the diversity and complexity of skin care product reviews. Therefore, using the BART model is a better choice for producing skin care product review summaries that are more informative and useful for consumers. In the future, these findings may contribute to developing more sophisticated product review summary systems.**

Keywords*— Natural language processing, Text Summarization, Product Reviews, BERT, BART, T5

I. Introduction

The rapid development of the skin care industry has increased the need for in-depth information about skin care products. Consumers tend to look for product reviews before purchasing to understand the experiences of other users and the results that can be achieved [1]. However, along with the increasing popularity of skin care products, the number of reviews spread across various online platforms is also increasing. The main challenge in accessing skin care product reviews is the large volume and variety. These long and varied reviews can hinder consumers from accessing relevant information. Therefore, summarizing reviews becomes an effective solution to solve this problem. Review summaries allow key information from those reviews to be presented more concisely and easily digestible [2].

This research [3] reviews the evolution of Deep Learning in COVID-19 news article summarization. Five NLP transformer models were evaluated, including BERT, GPT-2, XLNet, BART, and T5. BERT emerged as the winner in producing the best summary based on the ROUGE score. Extractive models, especially BERT, outperform abstractive models. Validation results with Word Cloud show that the "CoVShorts" application's summary remains focused on the original article's theme. The following steps include mobile application development, integration of extraction URLs, and exploration of combined NLP models for better summarization.

In product review summarization, extractive methods involve extracting important information from the original text without making significant changes. The methods used include BERT (Bidirectional Encoder Representations from Transformers), BART (Bidirectional and Auto-Regressive Transformers), and T5 (Text-To-Text Transfer Transformer). BERT is known for its ability to deeply understand the context of sentences and words and preserve relevant information from the original text. BART can produce more creative summaries by abstractively rephrasing sentences, while T5 has a versatile approach to summarization [4]. However, these three methods have disadvantages in handling long texts and depend on the quality of the original text and model parameters [5].

The selection of summarization methods in the context of skin care product reviews, namely BERT, BART, and T5, is based on the special advantages of each method. BERT was chosen for its ability to understand language context and retain relevant information in summaries. Meanwhile, BART was chosen for its ability to produce creative summaries and help simplify complex reviews. T5 was chosen for its flexibility in designing language processing tasks. Therefore, this study aims to evaluate the effectiveness of these three methods in producing quality product review summaries.

II. Research Method

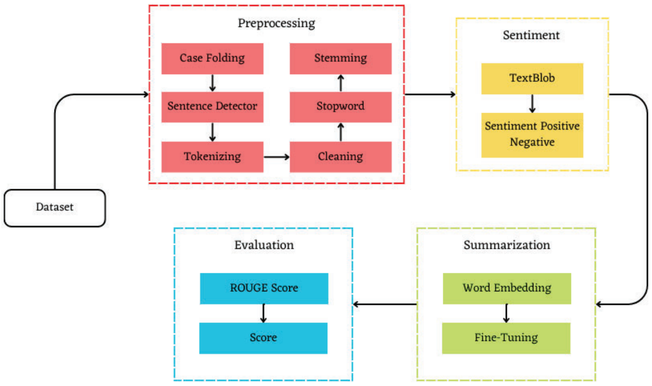

Fig. 1. Research methodology

A. Literature Review

While realizing how important and valuable it is to summarize texts, it has attracted the attention of many researchers over the years. The following subsection will review some of the past literature in this area.

In the research [6], experimental results show that the rapid growth of COVID-19 news articles in the global impact of the SARS-CoV-2 pandemic gave rise to the importance of text summaries. This research uses the latest NLP models, such as BERT, GPT-2, XLNet, BART, and T5, to summarize articles related to COVID-19. The results show that BERT, a transformer autoencoder model, summarizes SARS-CoV-2 news. These findings prompted the creation of a web application, "CoVShorts," which provides summaries of COVID-19 articles, helping the public understand evolving information. The BERT model has the highest ROUGE-2 and ROUGE-L scores, 0.354 and 0.364, followed by GPT-2 and XLNet. The T5 and BART models had the lowest scores.

Research by [4] news articles have a role in conveying important information from around the world to the public. However, news articles are often long and complex, making it difficult for readers to consume information efficiently. Therefore, summaries of news articles are very important. In this context, research has brought forward trained transformer-based models such as BERT, GPT-2, XL Net, BART, and T5 for both extractive and abstractive text summarization tasks. The ROUGE score is used to evaluate the quality of the summaries produced by the models. The research results show that BERT achieved ROUGE-1 of 0.662, ROUGE-2 of 0.510, and ROUGE-L of 0.413, which shows good performance. One of BERT's advantages is its ability to understand language context in both directions, enabling more informative text summarization.

This research [7] compares three text summarization models, namely BERT, BART, and T5, in creating relevant automatic summaries for research papers. After analysis and implementation of these three models, T5 was proven to be the most suitable for the proposed problem formulation. This aims to meet the needs of many researchers, professionals, and students in obtaining significant information from scientific papers with abstracts that may need to be more informative. The results were evaluated using the ROUGE metric, which aligns with previous research expectations and trends. Model T5 obtained ROUGE-1 results of 0.37364, ROUGE-2 of 0.33823, and ROUGE-L of 0.37364. Apart from that, T5 is also integrated with a front-end that users can access to interact with the system.

In a business context [8], online reviews play an important role in purchasing decisions. This paper aims to summarize Amazon product reviews based on non-fake reviews. This research compares constraining machine learning algorithms with SVD dimensionality reduction and text mining approaches for summarization. Models were drilled using a dataset of labeled Amazon reviews, primarily for the product "Samsung Galaxy S20 FE," with 1444 reviews. The results show that the model can classify reviews as fake or not with a training accuracy of 81%. Non-fake reviews were then analyzed using Vader's method using positive and negative sentiment. Review summarization is carried out extractively with the Gensim model. In various situations, Gensim succeeded in producing effective review summaries, especially paragraph-based review summarization, compared to LexRank summarization.

This research [9] aims to summarize online product reviews in Bengali using a Recurrent Neural Network (RNN) model based on Long Short-Term Memory (LSTM) and Sequence-to-Sequence (Seq2Seq). Experimental results show that the proposed model generates frequent summary predictions from the original text with minimal training loss. The Word Mover's Distance (WMD) method is also used to measure the similarity between human and machine summaries. Experimental results show that WMD is more effective than the Jaccard method in assessing text similarity. These findings have important implications for developing text summarization tools in Bengali, especially in the context of online product reviews. They can provide significant benefits to users in decision-making based on product reviews.

Previous studies have used BERT, BART, XLNet, T5, and GPT-2 methods for text summarization. However, research has yet to specifically focus on using these methods in summarizing skin care product reviews. Therefore, this study aims to highlight the performance of the BERT, BART, and T5 methods in summarizing skincare product reviews. This third method was chosen because, after sentiment analysis, it is important to retain relevant sentiment information in the product summary. BERT, BART, and T5 can understand and retain relevant information, including sentiment information, so they are the right choice in the context of this research.

B. Dataset

This dataset consists of 10,000 reviews of Cetaphil skin care products in Indonesian, taken from the Female Daily website. Data collection was carried out using web scraping techniques with the Hypertext Transfer Protocol (HTTP) method, which allows effective retrieval of information related to user reviews. This information is associated with the user's Cetaphil product experience and is stored in the "REVIEW" column.

C. Preprocessing

The initial stage of this research is data preprocessing, which aims to process raw data into data ready to be used [10]. The preprocessing process involves several stages, including case folding to change uppercase letters to lowercase, sentence detection to separate documents into sentences, tokenization to change sentences into groups of words with spaces as separators, cleaning to remove unimportant characters such as punctuation, symbols, emojis, numbers, and irrelevant words (stopwords), as well as stemming to remove affixes from words. This preprocessing stage is important before the data can be processed further [11].

D. Sentiment Analysis

In this research, sentiment was analyzed using TextBlob, a library that can be implemented in both Python 2 and Python 3 to process text data. This library provides an easy-to-use interface for performing various natural language processing (NLP) tasks, including language translation and text classification [12]. TextBlob is also able to calculate sentiment by taking into account polarity and subjectivity, making it possible to identify whether the content is positive or negative, with polarity > 0 considered positive and < 0 considered negative [13]. The results of sentiment analysis will reveal positive and negative sentiments, which are then used to create a summary.

E. Summarization with Models

Extractive summarization involves extracting meaningful information from the original text. BERT models are known for their ability to understand the context of sentences and words deeply, keeping relevant information to produce informative summaries. [14].

Furthermore, the BART (Bidirectional and Auto-Regressive Transformers) method is a model capable of producing extractive summarization. BART has the advantage of producing more creative summaries than BERT because BART's advantage lies in its ability to produce more general summaries but still reflect the essence of the information [4].

On the other hand, the T5 (Text-To-Text Transfer Transformer) method is a model that is also capable of extractive summarization. T5 is known for its versatile approach, where text is also considered input and output in text form. This provides flexibility in designing various natural language processing tasks, including summarization [7].

The three extractive summarization methods have advantages and disadvantages. The main drawbacks include handling long text or less relevant information and dependence on the quality of the original text and model parameters. This research aims to evaluate the effectiveness of these three methods in facing these challenges and produce a quality product review summary.

F. Performance Evaluation of Models

This research uses the ROUGE Score evaluation method, essential in text summarization because it measures the similarity between models and human summaries. ROUGE Score has variations such as ROUGE-1, ROUGE-2, and ROUGE-L. These metrics provide a quantitative summary quality measure and are used to develop summary models. In extractive summarization testing with three transformer models, the ROUGE Score evaluates model performance and enables further improvements. [15].

Precision measures the degree to which words in the system summary appear in references. The higher the precision, the more words are relevant to the reference in the system summary. The formula for calculating precision can be found in Equation 1 [11].

$$Precision = \frac{number\ of\ similar\ words}{total\ words\ in\ system\ summary} \quad (1)$$

Recall is a method for measuring the extent to which the words in the system summary include the words in the reference summary. The higher the recall value, the more words from the reference successfully captured by the system summary. The recall value is calculated using Equation 2 [11].

$$Recall = \frac{number\ of\ similar\ words}{total\ words\ in\ reference\ summary} \quad (2)$$

F1-score combines precision and recall to evaluate the balance between the quality of the system summary and the system's ability to capture relevant information from references. The formula for calculating the F1-score is in Equation 3 [11].

$$F - Score = \frac{2 \times Precision \times Recall}{Precision\ + Recall} \quad (3)$$

The research uses the ROUGE Score method, which checks the similarity between model summaries and human summaries. ROUGE Score has variations such as ROUGE-1 (unigram), ROUGE-2 (bigram), and ROUGE-L (most extended word sequence). This metric measures the quality of summaries produced by text summarization models, considering various aspects of similarity to human references. [15].

III. Result And Discusion

The summarization results using the three transformers models with ROUGE Score evaluation are illustrated in this section.

A. Preprocessing

This preprocessing process involves several important steps. First, the code imports the required libraries, including re (regular expressions), csv (for reading and writing CSV files), WordPunctTokenizer (for word tokenization), and Sastrawi's StemmerFactory (for the stemming process in Indonesian). Next, initial variables are initialized, including the name of the input CSV file, a list of stopwords, and the name of the output file for the preprocessing results.

The first step is “Case Folding,” where the text is converted to lowercase for consistency. Then, "Normalization and Cleaning" cleans up the text by removing special characters and leaving only letters and numbers. The next process is “Stopword Removal,” which removes common words that do not provide additional information. "Stemming using Literature" changes words to their basic form. In addition, the text is also tokenized using WordPunctTokenizer to break it into words. The list of stopwords is read from a file and stored in set form.

Data from the CSV file is read, each review is processed with the preprocessing steps described, and the results are stored in variables. The final result is text data that is clean and ready for further analysis. The results of preprocessing can be seen in Figure 3.

```
                                              REVIEW
0       Udah hampir 5 tahunan pake ini, walau ga konti...
1       Aku cobain cleanser ini karena dapat hadiah da...
2       Ini penyelamat aku waktu bruntusan parah, kare...
...
9997    Awal pake ini saat kondisi wajahku normal-norm...
9998    Sabun cuci muka yg worth to buy, ini buat all ...
9999    salah satu gentle cleanser yang aku suka sih i...
```

Fig. 2. Original Review

```
                     Prepocessing Data
0      udah,5,tahun,pake,ga,kontinyu,kadang,ling...
1      cobain,cleanser,hadiah,sepupu,emang,pakai...
2      selamat,bruntusan,parah,gentle,cocok,kuli...
...
9997   pake,kondisi,wajah,normalnormal,aja,ngera...
9998   sabun,cuci,muka,worth,to,buy,all,skin,typ...
9999   salah,gentle,cleanser,suka,all,one,pakai,...
```

Fig. 3. Preprocessed Review

B. Sentiment Using TextBlob

In this process, we start by setting up the required modules and determining the CSV file used as a data source. Next, we have the sentiment analysis function, which uses the TextBlob algorithm to determine whether the text has a positive or negative sentiment. The preprocessing data is read from a CSV file, and each row is retrieved, the sentences are combined, and the sentiment is analyzed. The sentiment analysis results are stored in a list for further processing and displayed in the output. After analyzing all the data, the results are saved in a new CSV file.

```
Sentiment Review
Sentence 1
Review          : udah,5,tahun,pake,ga,kontinyu,kadang,ling,face,wash,serius,emang,cetaphil,mild
Sentiment       : 1
Label Sentiment : Positive

Sentence 2
Review          : cobain,cleanser,hadiah,sepupu,emang,pakai,produk,merek,kasih,1,cleanser,cobain
Sentiment       : -1
Label Sentiment : Negative
```

Fig. 4. Sentiment Review

C. Models Implementation

The stages in the three models used to summarize skin care product reviews, namely BERT, BART, and T5, have unique characteristics. First, in the BERT model, the initial stage involves tokenization, where the product review text is broken down into smaller tokens, such as words. Then, the BERT model processes each token in parallel and produces a context-rich word vector representation. Furthermore, BERT uses an attention mechanism to understand the relationship between words in the review. The following critical stage is refinement, where the BERT model is designed specifically for the summarization task. This involves training data containing pairs of original text and human summaries. During training, the BERT model learns to produce summaries that reflect the main content of the original text.

Second, the BART model has a unique generative approach. The initial stage involves tokenization and encoding, similar to BERT. However, BART does decoding to produce a summary. This involves a learning process that teaches the model to reassemble and combine sentences from the original text into a suitable summary. BART uses a "denoising autoencoder" approach to optimize its results, meaning the model is taught to understand and produce better text by removing noise and errors in the original text.

Third, the T5 model applies a text-to-text approach. The initial stage is tokenization and encoding, similar to BERT and BART. Then, T5 combines the original text and task labels required for summarization. The T5 model generates summaries by treating original text as input and labels as output tasks. During training, the model learns to generate summaries that match the given task labels.

Third, the T5 model applies a text-to-text approach. The initial stage is tokenization and encoding, similar to BERT and BART. Then, T5 combines the original text and task labels required for summarization. The T5 model generates summaries by treating original text as input and labels as output tasks. During training, the model learns to generate summaries that match the given task labels.

D. Model Performance Evaluation with ROUGE

During the performance evaluation of three models using the ROUGE Score, the first step is to collect evaluation data, including a summary of the models (system summary) and a human summary (reference summary). Then, both summaries are broken down into small tokens, such as words or phrases, through a tokenization process, allowing ROUGE Score to measure their similarity. ROUGE Score then measures the overlap of tokens in the system and reference summaries. For example, ROUGE-1, ROUGE-2, and ROUGE-L measure the similarity of unigrams, bigrams, and the most extended sequence in a summary. ROUGE Score results are analyzed for each metric, including precision, recall, and F-measure calculations. These scores provide an understanding of the extent of each model's ability to produce summaries that approximate human summaries across varying levels of token similarity. Thus, evaluation with the ROUGE-1, ROUGE-2, and ROUGE-L methods provides a comprehensive insight into the performance of the three models in summarizing skin care product reviews.

1) BERT Model: The BERT model testing results can be seen in Figure 5, which displays the ROUGE Score from several reviews. In addition, Figure 6 displays the average ROUGE Score results from the BERT model.

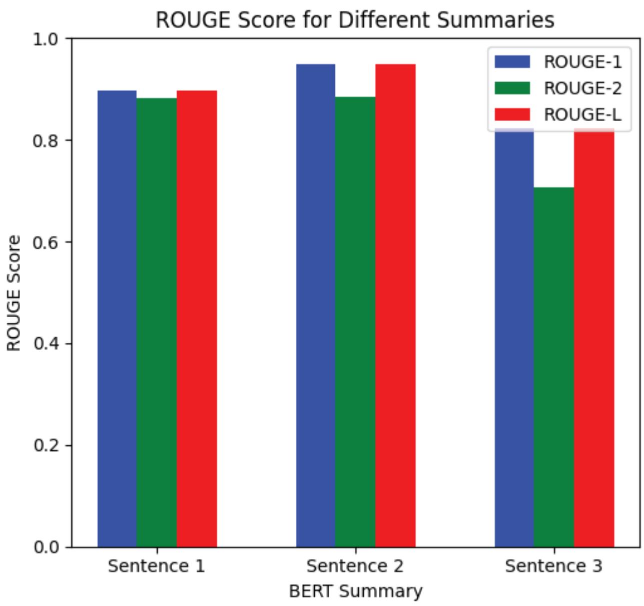

Fig. 5. BERT Model Performance Evaluation

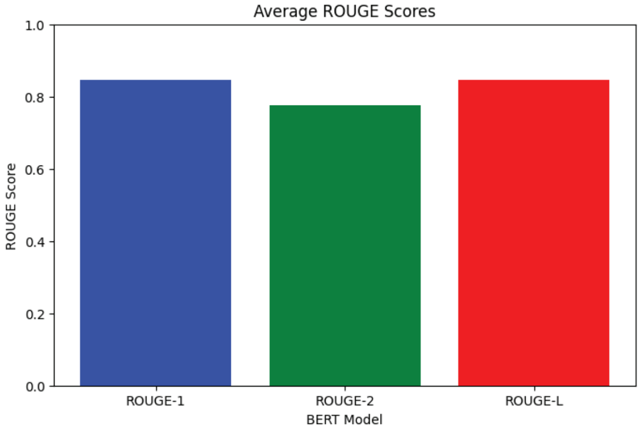

Fig. 6. Average BERT Model Performance Evaluation

Table 1. Rouge Score Results from Bert

Sentence	Rouge-1	Rouge-2	Rouge-L
S1	0.89655	0.88135	0.89655
S2	0.94999	0.88372	0.94999
S3	0.82352	0.70588	0.82352
Average	0.84591	0.77585	0.84479

2) BART Model: The BART model testing results can be seen in Figure 7, which displays the ROUGE Score from several reviews. In addition, Figure 8 displays the average ROUGE Score results from the BART model.

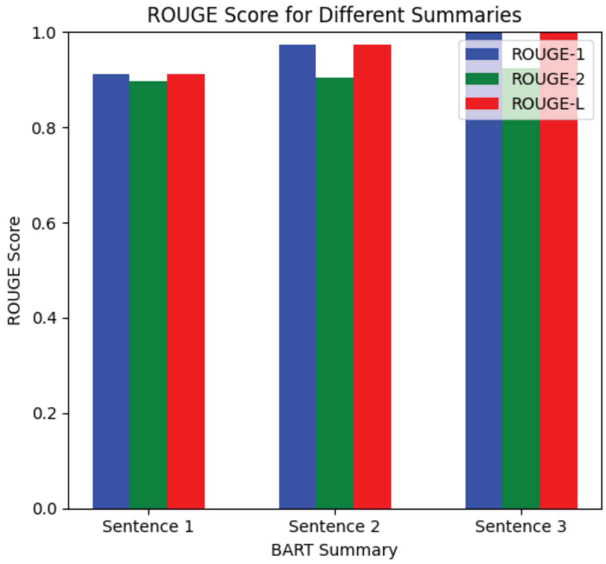

Fig. 7. BART Model Performance Evaluation

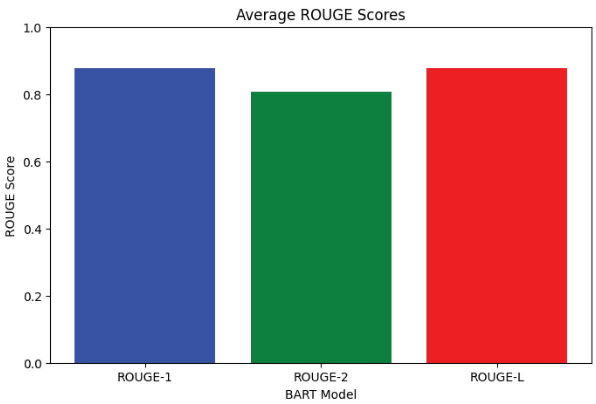

Fig. 8. Average BART Model Performance Evaluation

Table 2. Rouge Score Results from Bart

Sentence	Rouge-1	Rouge-2	Rouge-L
S1	0.91228	0.89655	0.91228
S2	0.97435	0.90476	0.97435
S3	0.99999	0.92307	0.99999
Average	0.87718	0.80689	0.87688

3) T5 Model: The T5 model testing results can be seen in Figure 9, which displays the ROUGE Score from several reviews. In addition, Figure 10 displays the average ROUGE Score results from the T5 model.

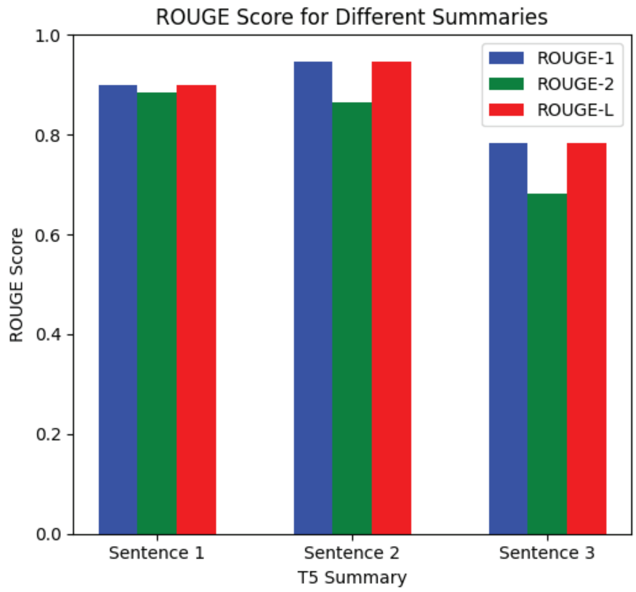

Fig. 9. T5 Model Performance Evaluation

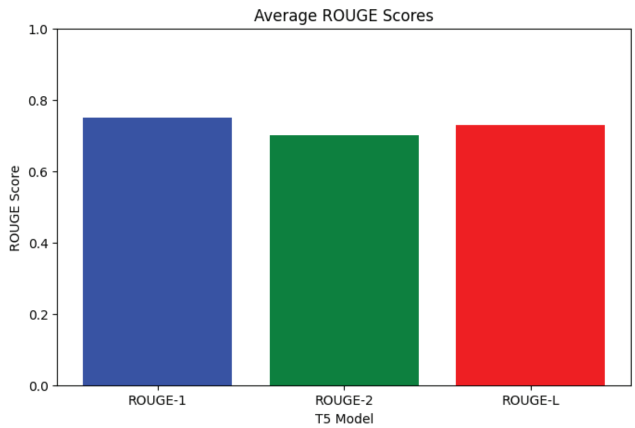

Fig. 10. Average T5 Model Performance Evaluation

Table 3. Rouge Score Results from T5

Sentence	Rouge-1	Rouge-2	Rouge-L
S1	0.89999	0.88524	0.89999
S2	0.94736	0.86363	0.94736
S3	0.78260	0.68085	0.78260
Average	0.74995	0.70048	0.73023

After testing the performance of three models using the ROUGE Score method on a dataset of 10,000 reviews of Cetaphil brand facial cleansing products, the results show that the BART model has a higher average ROUGE Score than the BERT and T5 models. Specifically, the BART model has an average ROUGE-1 of 0.87718, ROUGE-2 of 0.80689, and

ROUGE-L of 0.87688. Meanwhile, the BERT model has an average ROUGE-1 of 0.84591, ROUGE-2 of 0.77585, and ROUGE-L of 0.84479, and the T5 model has an average ROUGE-1 of 0.74995, ROUGE-2 of 0.70048, and ROUGE-L of 0.73023. These results indicate that the BART model is consistently superior in its ability to summarize skincare product reviews compared to the BERT and T5 models.

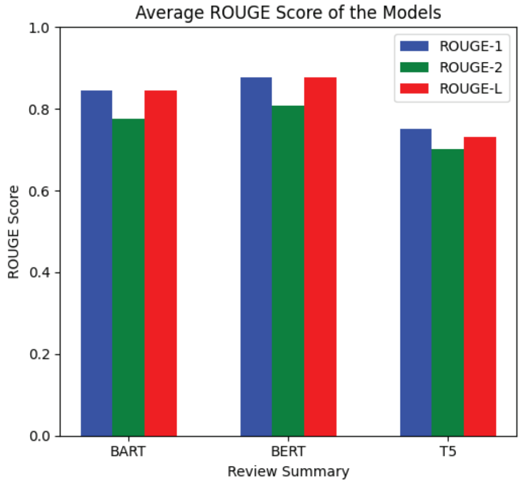

Fig. 11. Overall Average Result of Models

IV. Conclusion

This study evaluates three summarization methods, namely BERT, BART, and T5, in the context of the growth of the skincare industry and the increasing demand for product information. Each method has unique advantages: BERT in language understanding, BART in the creativity of summarizing complex text, and T5 in the flexibility of language processing. Evaluation results using the ROUGE Score method using the Cetaphil brand facial wash product review dataset show that the BART model consistently provides the highest average ROUGE Score, with ROUGE-1 of 0.87718, ROUGE-2 of 0.80689, and ROUGE-L of 0.87688. Compared with BERT and T5, which had lower scores, this confirms that the generative approach used by BART is more effective in addressing the diversity and complexity of skincare product reviews. Therefore, using the BART model is recommended to produce skin care product review summaries that are informative and useful for consumers and potentially contribute to developing more sophisticated product review summary systems in the future.

References

[1] C. Qalbyassalam, R. F. Rachmadi, and A. Kurniawan, “Skincare Recommender System Using Neural Collaborative Filtering with Implicit Rating,” in *Proceeding of the International Conference on Computer Engineering, Network and Intelligent Multimedia, CENIM 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 272–277. doi: 10.1109/CENIM56801.2022.10037471.

[2] R. Noviani, F. Anugrah, T. I. Manangka, and Yuniarty, “The Role of Social Media Marketing and Product Characteristics on Increasing Consumer Satisfaction and Korean Skincare Products Repurchase Intention in Indonesia,” in *ICBIR 2022 - 2022 7th International Conference on Business and Industrial Research, Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 210–215. doi: 10.1109/ICBIR54589.2022.9786467.

[3] 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.

[4] W. H. Ong, K. G. Tay, C. C. Chew, and A. Huong, “A Comparative Study of Extractive Summary Algorithms Using Natural Language Processing,” in *2020 IEEE Student Conference on Research and Development, SCOReD 2020*, Institute of Electrical and Electronics Engineers Inc., Sep. 2020, pp. 406–410. doi: 10.1109/SCOReD50371.2020.9251032.

[5] A. Choudhary, M. Alugubelly, and R. Bhargava, “A Comparative Study on Transformer-based News Summarization,” in *Proceedings - International Conference on Developments in eSystems Engineering, DeSE*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 256–261. doi: 10.1109/DeSE58274.2023.10099798.

[6] 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.

[7] S. Badhe, M. Hasan, V. Rughwani, and R. Koshy, “Synopsis Creation for Research Paper using Text Summarization Models,” in *2023 4th International Conference for Emerging Technology, INCET 2023*, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/INCET57972.2023.10170144.

[8] R. R. Sathiya, L. Monish Raaj, S. Deekshan, P. Arjun Dev, and S. Aakash Muthiah, “Detection and Summarization of Honest Reviews Using Text Mining,” in *8th International Conference on Smart Structures and Systems, ICSSS 2022*, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/ICSSS54381.2022.9782167.

[9] M. J. Islam, K. Ahammad, and M. K. H. Chowdhury, “Summarizing online product reviews in bengali based on similarity using sequence to sequence RNNs,” in *Proceedings of 2020 11th International Conference on Electrical and Computer Engineering, ICECE 2020*, Institute of Electrical and Electronics Engineers Inc., Dec. 2020, pp. 33–36. doi: 10.1109/ICECE51571.2020.9393048.

[10] P. Raundale and H. Shekhar, “Analytical study of Text Summarization Techniques,” in *2021 Asian Conference on Innovation in Technology, ASIANCON 2021*, Institute of Electrical and Electronics Engineers Inc., Aug. 2021. doi: 10.1109/ASIANCON51346.2021.9544804.

[11] N. A. Maghfiroh, G. W. Wicaksono, and C. S. K. Aditya, “Peringkasan Berita Online Corona Virus dengan Metode Lexical Chain dan Word Sense Disambiguation,” *Komputika : Jurnal Sistem Komputer*, vol. 10, no. 2, pp. 145–151, Aug. 2021, doi: 10.34010/komputika.v10i2.4499.

[12] M. S. Meghana, D. Abhijith, S. Aysha, and P. K. Kollu, “Sentiment Analysis on Amazon Product Reviews using LSTM and Naive Bayes,” in *Proceedings - 7th International Conference on Computing Methodologies and Communication, ICCMC 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 626–631. doi: 10.1109/ICCMC56507.2023.10084052.

[13] O. Octaria, D. Manongga, A. Iriani, H. D. Purnomo, and I. Setyawan, “Mining Opinion Based on Tweets about Student Exchange with Tweepy and TextBlob,” in *Proceedings - 2022 9th International Conference on Information Technology, Computer and Electrical Engineering, ICITACEE 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 102–106. doi: 10.1109/ICITACEE55701.2022.9924013.

[14] A. Srikanth, A. S. Umasankar, S. Thanu, and S. J. Nirmala, “Extractive text summarization using dynamic clustering and co-reference on BERT,” in *Proceedings of the 2020 International Conference on Computing, Communication and Security, ICCCS 2020*, Institute of Electrical and Electronics Engineers Inc., Oct. 2020. doi: 10.1109/ICCCS49678.2020.9277220.

[15] M. R. Ramadhan, S. N. Endah, and A. B. J. Mantau, “Implementation of Textrank Algorithm in Product Review Summarization,” in *ICICoS 2020 - Proceeding: 4th International Conference on Informatics and Computational Sciences*, Institute of Electrical and Electronics Engineers Inc., Nov. 2020. doi: 10.1109/ICICoS51170.2020.9299005.