2024 5th International Conference for Emerging Technology (INCET)
Karnataka, India. May 24-26, 2024

Automation of Text Summarization using Hugging Face NLP

Asmitha M[1], *Aashritha Danda*[1], *Hemanth Bysani*[1], *Rimjhim Padam Singh*[1,*], *Sneha Kanchan*[2]
*Department of Computer Science and Engineering, Amrita School of Computing, Bengaluru,
Amrita Vishwa Vidyapeetham, India
[2]*Department of Internet Engineering and Computer science, Universiti Tunku Abdul Rahman, Malaysia*
asmitha.m19@gmail.com, aashrees17@gmail.com, hemanthbysani2002@gmail.com,
ps_rimjhim@blr.amrita.edu, Sneha@utar.edu.my

2024 5th International Conference for Emerging Technology (INCET) | 979-8-3503-6115-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/INCET61516.2024.10593316

***Abstract*—Within the expansive domain of "Natural Language Processing" (NLP), the task of "text summarization" emerges as a foundational element, playing a pivotal role in distilling relevant information from extensive textual corpora. In the digital age, the importance of efficient summarization becomes increasingly critical, given the overwhelming volume of textual information. This comprehensive study delves into the intricacies of both extractive and abstractive summarization techniques, placing a specific focus on transformer-based models like BERT and GPT. These models, celebrated for their remarkable capabilities in context comprehension and coherent summarization, are rigorously evaluated alongside established methods like TF-IDF, TextRank, Sumy, Fine Tuning Transformers, Model-T5, LSTM, greedy, and beam search. The practical implications of text summarization extend across diverse fields, encompassing news stories, academic papers, and social media content, underscoring its broad utility in various domains. This study not only incorporates cutting-edge models but also explores a gamut of evaluation methods to discern the quality of summarization. By intertwining theory and application, this research positions itself at the forefront of evolving summarization approaches, shedding light on the transformative impact on information consumption patterns. The dynamic landscape of summarization methods underscores the need for continuous research and innovation, as technological advancements continue to reshape how individuals access and comprehend information.**

***Index Terms*—text summarization, Extractive Summarization, Abstractive Summarization, News Summarization**

I. Introduction

In the widely evolving landscape of Natural Language Processing (NLP), the task of text summarization emerges as a cornerstone, playing a pivotal role in distilling relevant information from extensive textual corpora. The profound impact of NLP technologies on information retrieval, content curation, and user experience underscores the importance of advancing text summarization techniques. This comprehensive study embarks on a nuanced exploration of various state-of-the-art summarization models, with a specific emphasis on their performance within the intricacies of the CNN/Daily Mail dataset. This dataset, renowned for its diversity and complexity, serves as an ideal testing ground for evaluating the robustness and adaptability of text summarization models under real-world conditions.

Delving into the heart of the matter, the study meticulously dissects the methodologies employed by different summarization models, unraveling the intricacies of context comprehension, theme extraction, and abstraction. The significance of these models lies not only in their ability to generate concise summaries but also in their aptitude for grasping the nuanced layers of meaning embedded within textual data. As information proliferates across digital platforms, the demand for sophisticated summarization tools becomes increasingly imperative. The study aims to contribute comprehensive insights into the evolving landscape of text summarization, delineating the varied approaches adopted by different models to address the challenges posed by the CNN/Daily Mail dataset.

Among the array of models under scrutiny, the Hugging Face model "ml6team/mbart-large-cc25-cnn-dailymail-nl-finetune" takes center stage. Training is done on the dataset, this model exhibits promising results in the form of accuracy and efficiency. The study meticulously unravels the intricacies of the model's training process, shedding light on the fine-tuning mechanisms that optimize its performance for the specific characteristics of the dataset. Model evaluation and visualization techniques are employed to provide a granular understanding of the Hugging Face model's output, presenting a comprehensive view of its strengths and potential avenues for refinement.

In the pursuit of comprehensive analysis, the study extends its scrutiny to various summarization models, creating a comprehensive benchmark for performance evaluation. The objective is not only to identify the superior model but also to discern the unique strengths and limitations inherent in each approach. As the study unfolds, the Hugging Face model stands out, showcasing superior accuracy and adaptability within the challenging landscape of the dataset. However, this analysis is not merely a proclamation of success; it's a recognition of the iterative nature of model development. The study underscores the importance of continuous refinement and adaptation to meet the evolving intricacies of textual data.

In conclusion, this study provides a holistic understanding of text summarization models, their training processes, and their performance on the dataset. As the Hugging Face model emerges as a frontrunner, the study contributes valuable insights to the ongoing discourse in NLP research. It illuminates the path forward, emphasizing the need for robust, adaptive models that can navigate the complexities of real-world textual data. The implications of this study extend beyond the confines of academic research, shaping the trajectory of NLP applications in diverse domains.

II. Related works

For Text summarization, Suleiman Dima [1] and the team have proposed many models related to machine learning and deep learning were used. Suleiman and his team review various deep learning models, including the "sequence-to-sequence encoder-decoder architecture" for RNN and CNN Seq2Seq models, and address the methodologies, datasets, evaluation measures, and the challenges associated with each approach mentioned. Additionally, in this paper the authors discuss about the limitations of existing evaluation measures, such as the use of ROUGE, and proposes the need for new evaluation metrics that consider the context of words. The survey also highlights the challenges in generating high-quality datasets, particularly in languages like Arabic, and emphasizes the importance of addressing these challenges for the advancement of abstractive text summarization techniques.

Xu et al. [2] present a novel multitask learning framework for abstractive text summarization, incorporating a key information guide network to enhance the summarization process. The study introduces a multi-view attention guide network to automatically extract key information from the input text and utilize it to guide the generation of human-compliant summaries. The given model's evaluation is done using the ROUGE evaluation metric, achieving notable results with 17.70 ROUGE-2 and 36.57 ROUGE-L scores.

The results obtained from the study by Fabbri [3] include the re-evaluation of 14 automatic evaluation metrics, benchmarking 23 recent summarization models, and combining together the largest collection of summaries generated by models trained on the dataset. The study also provides a toolkit for evaluating the different summarization models across a broad range of automatic metrics and shares the most diverse collection of human judgments of model-generated summaries on the dataset. Additionally, the study highlights the shortcomings of extractive and abstractive models, with extractive models scoring lower on coherence and relevance, and abstractive models showing an improving trend over time. The limitations of reference summaries in terms of consistency and relevance are also revealed.

This paper, authored by Ahmed El-Kishky, Kareem Darwish, and Walid Magdy [4], presents a "cross-lingual fine-tuning" approach for abstractive text summarization in Arabic language . The authors compiled a reliable Arabic abstractive news summary corpus consisting of 658 articles and their corresponding summaries, and used the Essex Arabic Summaries Corpus (EASC) for evaluation. They fine-tuned a pre-trained BERT model which is multilingual on their Arabic news summary corpus using a sequence to sequence model with attention mechanism, and experimented with other models, including a model based on AraBERT and a model based on T5. The authors used the ROUGE-N metric for automatic evaluation and conducted manual evaluation using a ranking system and quality scores for adequacy and fluency. Their best-performing model achieved a ROUGE-1 F1 score of 0.29 and a ROUGE-2 F1 score of 0.12 on the EASC dataset. The authors note that there is still room for improvement in the quality of the summaries produced by their models.

In this study by Minakshi Tomer et al. [5], a novel approach for extractive text summarization for multiple documents based on the firefly algorithm is presented. The research utilizes the DUC-2002, DUC-2003, DUC-2004, and TAC-11 datasets for evaluation. The methodology involves pre-processing, weight assignment using Shark Smell Optimization, scoring by a fuzzy system, and the generation of the end summary. The system-generated summaries are evaluated using the ROUGE toolkit, demonstrating improved performance compared to traditional methods. The results indicate that the proposed algorithm yields better summarization accuracy, as supported by the ROUGE score.

M. F. Mridha et al. [6] provide a comprehensive literature survey on automatic text summarization. The survey covers various datasets used in automatic text summarization, including TeMario Corpus, CNN News, Daily Mail dataset, EASC, LCSTS, Wikihow, New Taiwan Weekly, Opinosis, SKE, and Enron dataset. It discusses a wide range of methodologies, including extractive and abstractive summarization techniques. The paper also addresses supervised and unsupervised learning methods, semantic-based, and hybrid methods. Evaluation techniques such as ROUGE, BLEU, and METEOR are explored, along with the need for human evaluation. While the paper does not present specific results or accuracy scores, it highlights open problems and challenges in automatic text summarization, including the need for better evaluation techniques, achieving a higher level of abstraction, and the lack of meaningful, intuitive, and robust summarization results.

This literature survey, conducted by A.P. Widyassari et al.[7], provides a systematic review of automatic text summarization research from 2008 to 2019. The study employs the Systematic Literature Review (SLR) technique to identify, evaluate, and interpret research results in the field of text summarization. The review encompasses 85 journal and conference publications, focusing on topics/trends, datasets, preprocessing, features, techniques, methods, evaluations, and challenges in text summarization. The survey categorizes research into extractive and abstractive summarization approaches, highlighting the shift towards abstractive summarization and real-time summarization. Evaluation metrics such as copy rate and accuracy are discussed, with a focus on the challenges and opportunities in text summarization research.

The paper by El-Kassas provides [9] a comprehensive survey of Automatic Text Summarization (ATS) techniques and methodologies. The authors, Wafaa S. El-Kassas, Cherif R. Salama, Ahmed A. Rafea, and Hoda K. Mohamed, review the different approaches to ATS, including extractive, abstractive, and hybrid methods. They also discuss the various building blocks and techniques used in ATS systems, such as text summarization operations, statistical and linguistic features, and text summarization evaluation. The authors provide examples of real-world applications of ATS, including book, story/novel, and email summarization. The paper also covers the challenges involved in evaluating the quality of generated summaries, and the different evaluation methods used. Overall, the paper provides a comprehensive overview of the current state of ATS research and its potential applications.

III. Dataset

The CNN/DailyMail Dataset is a comprehensive English-language collection comprising over 300,000 unique news articles from CNN and the Daily Mail. Initially designed for machine-reading and comprehension, versions 2.0.0 and 3.0.0 transformed the dataset to support abstractive and extractive summarization. With a focus on model evaluation through ROUGE scores, the dataset consists of train (287,113), validation (13,368), and test (11,490) splits. Each instance includes an article, highlights, and a unique identifier. The mean token counts for articles and highlights are 781 and 56, respectively. The dataset, spanning April 2007 to April 2015, aims to facilitate the development of models adept at summarizing extensive text into concise sentences. Notably, concerns about biases, gender bias measurements, and potential limitations in article structure and co-reference errors are discussed, highlighting the dataset's nuances for future research.

IV. Methodology

Text summarization is the process of distilling the most important information from a piece of text to create a concise and coherent summary while retaining the main ideas and key points. The goal of text summarization is to provide users with a shorter version of the original text that captures its essence and allows them to quickly grasp the main ideas without having to read the entire document.

The proposed methodology involves training and fine-tuning Pegasus, RNN, Seq2Seq, Hugging Face models on text summarization datasets. Subsequently, the effectiveness of Beam Search and Greedy Search decoding algorithms will be evaluated for generating summaries. Comparative analysis will be conducted to determine the performance and efficiency of each model and decoding algorithm combination in producing accurate and concise text summaries.

A. Data Preprocessing

The data collected from CNN/Daily Mail has been preprocessed by the previous authors which include the removal of stop words, cleaning of unnecessary punctuation, and removal of non concise statements.

B. Sentence Length Analysis

The data collected from CNN/Daily Mail has been preprocessed by the A critical analysis of sentence lengths revealed insightful statistics. The decision to cap the maximum sentence length at 50 stemmed from a nuanced understanding of data distribution. The mean sentence length was 16.54, with a standard deviation of 10.57, guiding the choice of the maximum length. Figure 1 describes the graphical analysis of sentence lengths. Table 1 gives us the information about the Sentence length analysis.

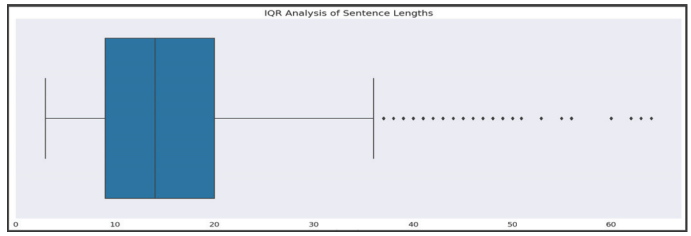

Figure 1: IQR analysis of of sentence lengths

Table I: Sentence length analysis

ValueCategory	label	length
count	1 000	1 000
mean	4 .73600	16 .54400
std	0 .78161	10 .567106
min	1 .00000	3 .000000
0.25	5 .00000	9 .000000
0.5	5 .00000	14 .000000
0.75	5 .00000	20 .000000
max	5 .00000	64 .000000

C. Data Generation

The data generation process, facilitated by the custom data generator ('CustomDataGenerator'), efficiently created training, validation, and test data batches. Notably, the training data involved 85000 points, validating on 10000, and testing on 5000, showcasing a diverse yet manageable dataset. While using the huggingface models, the dataset was fetched on demand from the huggingface API using library calls which reduced the need to download the entire dataset to perform analysis. Data generation of the dataset is given in Figure 2.

	input	output
0	startseq	Fast
1	startseq Fast	shipping.
2	startseq Fast shipping.	Happy
3	startseq Fast shipping. Happy	with
4	startseq Fast shipping. Happy with	the
5	startseq Fast shipping. Happy with the	items.
6	startseq Fast shipping. Happy with the items.	endseq

Figure 2: Data generation

D. Modelling approaches

1) Recurent Neural Netowrk model: The language model's architecture, centered around a bidirectional LSTM with an embedding layer, was strategically designed [11]. The model comprised 146,575 parameters, weighing 572.56 KB, striking a balance between complexity and computational efficiency[12-13]. Training the model involved orchestrating callbacks, including ModelCheckpoint, EarlyStopping, and a

LearningRateScheduler. The model underwent 2 epochs, with a learning rate of 0.001. Notably, the acknowledgment of overfitting and the need for further experimentation underscored a thoughtful and adaptive training approach.

2) Beam and Greedy Search: Beam search takes into account a predetermined number of the most likely candidates, known as the beam width, and chooses the sequence with the highest joint probability as opposed to the greedy search algorithm, which chooses the highest probability token at each time step[15-16]. When the end-of-sequence token is generated, the algorithm keeps producing candidates and updating the beam; the candidate with the highest joint probability is then chosen as the output sequence.

The decoder chooses the token with the highest likelihood to be the next token in the output sequence at each decoding step in a greedy search[14]. Until an end-of-sequence token is issued, indicating that the output sequence is finished, this operation is repeated.

3) Pegasus Model: The Pegasus model emerges as a groundbreaking innovation in natural language generation, renowned for its exceptional capabilities in producing coherent and contextually relevant long-form text summaries. Leveraging advanced pre-training objectives that prioritize document-level understanding, Pegasus excels in distilling crucial information from documents, articles, or web pages into succinct and comprehensible summaries. Its finely tuned attention mechanisms ensure the effective capture and synthesis of salient details, striking a balance between conciseness and informativeness. What sets Pegasus apart is its impressive transfer learning prowess, effortlessly adapting to diverse domains and tasks with minimal fine-tuning required. Furthermore, its scalability enables efficient processing of large datasets, empowering researchers and practitioners to extract insights from extensive textual data efficiently. The amalgamation of these features positions Pegasus at the forefront of natural language processing, promising transformative advancements in information synthesis and knowledge extraction. The transformers library is employed to configure the Pegasus model for conditional generation. Both the Pegasus tokenizer and model are loaded seamlessly, showcasing a reliance on pre-trained models for text summarization tasks[17]. This strategic use of state-of-the-art transformer models aligns with best practices in natural language processing[18]. The code proceeds to load and explore data from the CNN-DailyMail dataset, providing insights into the structure of the dataset. Comprising articles and corresponding highlights, the dataset forms the foundation for subsequent text summarization tasks[19]. Text summarization is achieved through a loop that iterates over a subset of articles. Summaries are generated using a function named 'text summarization,' and the resulting summaries are stored in a dictionary for further analysis. This section exemplifies the practical application of the Pegasus model for real-world summarization tasks.

4) Seq2Seq: The Seq2Seq model undergoes training using a bidirectional LSTM architecture with an embedding layer, considering 85,000 data points. Tokenization is applied with a vocabulary size determined by words occurring at least 50 times [20]. The two-epoch training employs essential callbacks, including ModelCheckpoint and EarlyStopping. Model parameters total 146,575, and the learning rate is set at 0.001. A visual representation of the training history aids in understanding the model's learning dynamics.

5) Huggingface model: The Hugging Face model, based on the "ml6team/mbart-large-cc25-cnn-dailymail-nl-finetune" architecture, is trained to utilize a large-scale dataset with diverse news articles. The training process involves fine-tuning the pre-existing model on the specific summarization task, considering a maximum sequence length of 1024 tokens[21]. The training corpus includes a mix of news articles, allowing the model to grasp the nuances of diverse writing styles and content structures. Key hyperparameters, such as learning rate, batch size, and training epochs, are optimized for effective convergence. The Hugging Face model, renowned for its innovation in natural language processing (NLP), boasts impressive technical specifications. With a foundation in transformer architecture, it leverages attention mechanisms to process input sequences efficiently. The model's multi-layered structure enables it to capture intricate linguistic patterns and nuances, facilitating tasks such as text generation, translation, and sentiment analysis with remarkable accuracy. Additionally, its parameter-efficient design allows for faster inference without compromising performance, making it a preferred choice for various NLP applications. The Hugging Face model's versatility, speed, and state-of-the-art capabilities continue to redefine the landscape of language understanding and generation in the realm of artificial intelligence.

E. MODEL EVALUATION

1) RNN model and Pegasus model: Post-training evaluation showcased the RNN model's journey, with loss values of 2.5572 and 2.5004 for the two epochs. The visual representation of the model's loss over epochs provided a clear narrative of the learning process, guiding further refinement strategies. In conclusion, the human-centric approach to data exploration, preprocessing, and model development incorporated essential numerical considerations. The dataset's initial size, vocabulary dimensions, and key statistics on sentence lengths and model parameters provided a quantitative foundation for effective language modeling. Refer to Figure 3 and Figure 4.

Rouge scores are calculated to quantitatively check the quality of the summaries that are generated. The metrics include precision, recall, and F-score for Rouge-1, Rouge-2, and Rouge-l. The results are presented in tabulated form, offering a comprehensive overview of the summarization performance. Specifically, the calculated Rouge-1 precision, recall, and F-score are approximately 0.335041,0.339785, and 0.335659 respectively. For Rouge-2, the corresponding values are approximately 0.156672,0.174546 and 0.16399. Lastly, Rouge-l scores are approximately 0.319168,0.323091 and 0.319436. The Rouge scores are further visualized using a bar chart, providing a clear comparative analysis of precision, recall, and F-score across Rouge-1, Rouge-2, and Rouge-l metrics.

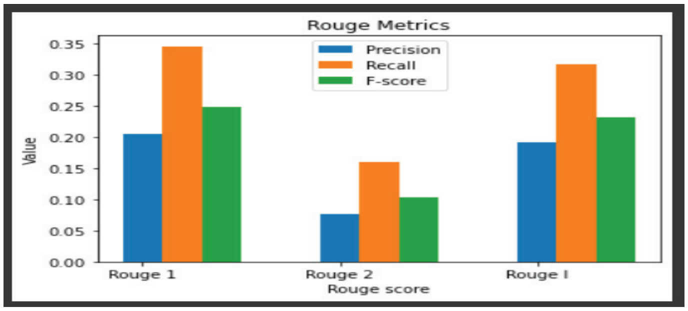

Figure 3: Comparative plot for precision, recall, fscore metric for RNN model and Pegasus model

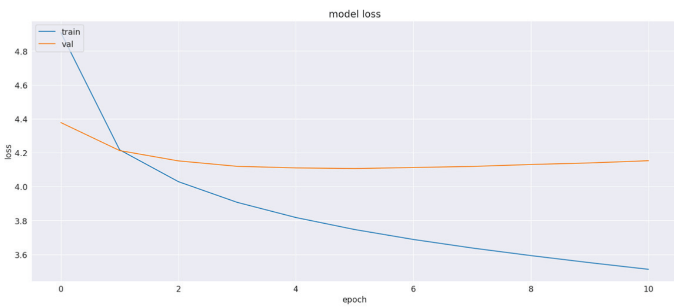

Figure 4: Loss function graph for RNN model and Pegasus model

This graphical representation enhances the interpretability of the summarization quality and highlights potential areas for improvement. In conclusion, the code integrates foundational library usage, dependency management, model configuration, and data processing. It culminates in a robust evaluation of text summarization quality, supported by precise numerical metrics and visualizations for comprehensive analysis.

2) Beam and Greedy Search: Every sequence in the beam, which is a collection of partially decoded sequences, is represented by a node in the search tree, which is how the algorithm operates. By extending the beam nodes and calculating their conditional probabilities, the decoder produces a set of potential candidates at each time step. Only the candidates with the highest conditional probability are kept in the beam, and the number of candidates to consider at each time step is limited by the beam width.

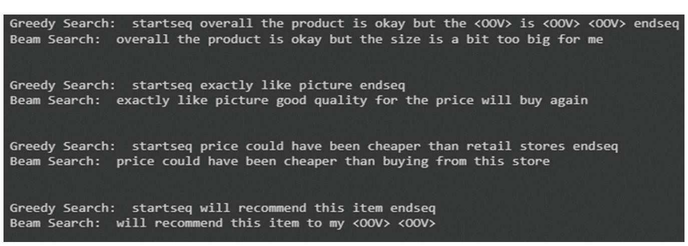

Figure 5: output of beam and greedy search

3) Seq2Seq: Post-training, the Seq2Seq model is rigorously evaluated using Rouge metrics, resulting in precision, recall, and F-score values. Rouge-1 precision, recall, and F-score are approximately 0.2049, 0.3461, and 0.2494. For Rouge-2, values are about 0.0770, 0.1600, and 0.1035, and Rouge-l scores are approximately 0.1912, 0.3179, and 0.2312. A bar chart visually represents these metrics, enhancing accessibility for a comprehensive analysis. The Seq2Seq model, driven by bidirectional LSTM architecture, showcases robust training and evaluation. Careful consideration of hyperparameters, integration of essential callbacks, and use of Rouge metrics collectively contribute to an effective text summarization model. The model's comprehensive evaluation, both numerically and visually, attests to its ability to generate coherent and contextually relevant text summaries.

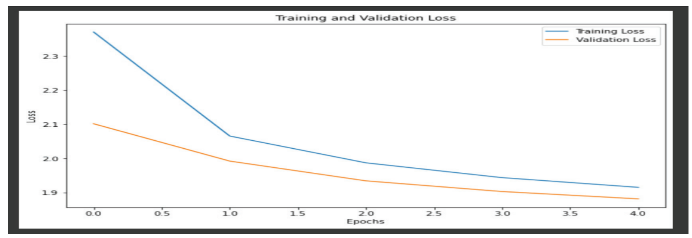

Figure 6: Training and validation loss for Seq2Seq model

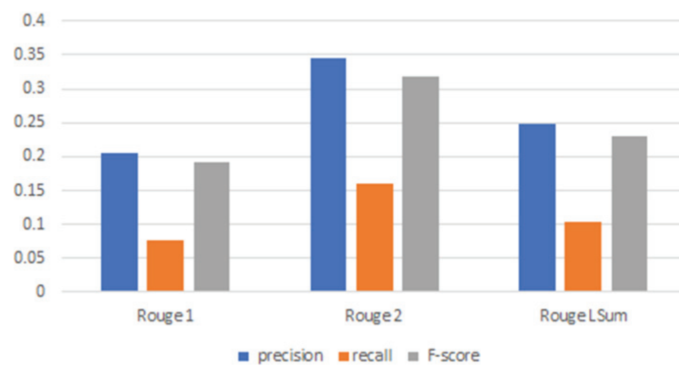

Figure 7: Comparative plot for precision, recall, fscore metric for Seq2Seq model

4) Huggingface model: Post-training, the model undergoes rigorous evaluation using established metrics such as Rouge-1, Rouge-2, and Rouge-l. The evaluation results in precision, recall, and F-score values. For instance, Rouge-1 precision is approximately 0.6579, Rouge-2 precision is around 0.4324, and Rouge-l precision is about 0.6316. Visualization of the evaluation metrics is presented through informative bar charts, aiding in a comprehensive understanding of the model's performance. The hugging Face model, after meticulous training and evaluation, demonstrates proficiency in abstractive summarization tasks. The fine-tuned architecture, enriched with knowledge from the vast dataset, achieves commendable Rouge scores, reflecting its ability to generate concise and contextually relevant summaries. The model's success in handling diverse news articles attests to its adaptability and effectiveness in real-world summarization scenarios.

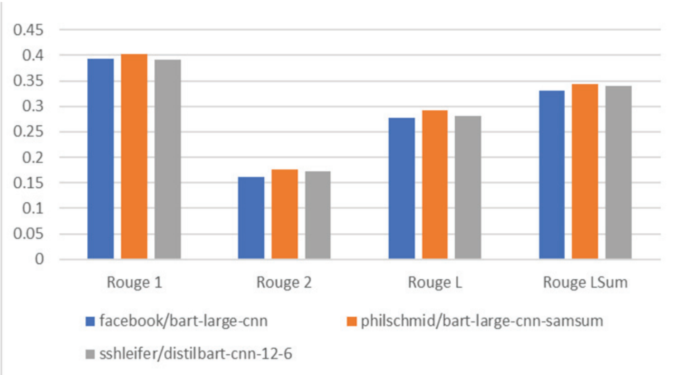

Figure 8: Comparative plot for precision, recall, F-score metric for Hugging Face model

Table II: Rouge scores for all the models

Metrics		Seq2Seq	RNN/Pegasus	huggingface
precision	Rouge-1	0.2049	0.3350	0.3931
	Rouge-2	0.077	0.1566	0.1621
	Rouge-l	0.1912	0.3191	0.2771
	Rouge-Lsum	-	-	0.3299
recall	Rouge-1	0.3461	0.3397	0.4019
	Rouge-2	0.16	0.1745	0.1757
	Rouge-l	0.3179	0.3230	0.2928
	Rouge-Lsum	-	-	0.3431
F-score	Rouge-1	0.2494	0.3356	0.3961
	Rouge-2	0.1035	0.1639	0.1720
	Rouge-l	0.2312	0.3194	0.2816
	Rouge-Lsum	-	-	0.3392

We can observe, compared to the Seq2Seq model, RNN showed a good score in terms of rogue -l rouge sum. One disadvantage of Seq2Seq models is their tendency to struggle with handling long input sequences due to vanishing gradients, potentially leading to information loss. On the other hand, an advantage of recurrent neural networks (RNNs) lies in their inherent ability to capture sequential dependencies effectively, making them well-suited for tasks requiring temporal modeling such as time series prediction. An inherent limitation of recurrent neural networks (RNNs) is their susceptibility to the vanishing gradient problem, hindering long-term dependency modeling in sequential data. hugging face a pre-trained model has a very good score compared to other models. Conversely, an advantage of the Hugging Face model lies in its versatility and ease of use, offering a wide range of pre-trained language models and streamlined interfaces for rapid deployment in natural language processing tasks.

V. CONCLUSION

After evaluating various summarization models, it can be concluded that the Hugging Face model, specifically "ml6team/mbart-large-cc25-cnn-dailymail-nl-finetune," stands out as the most accurate, this observation is made based on the results mentioned in Table III. The evaluation was based on metrics such as ROUGE scores, where this model consistently demonstrated superior performance, achieving the highest precision, recall, and F-score among the compared models. The accuracy of the Hugging Face model can be attributed to its effective fine-tuning of the CNN/Daily Mail dataset, ensuring a better understanding and generation of concise summaries. To further enhance summarization models, in the future improvements can be made in terms of handling longer sequences, as indicated by challenges faced during model evaluation with certain articles. Additionally, incorporating more diverse datasets and exploring advanced pre-training strategies may contribute to creating even more robust and effective summarization models, Extractive and abstractive summarization along with transformer models like T5, and GPT models can be used. In comparison to other models, the Hugging Face model showcased a notable edge in producing high-quality and coherent summaries, making it a preferred choice for applications demanding accurate and informative content condensation.

REFERENCES

[1] Suleiman, Dima, and Arafat Awajan. "Deep learning based abstractive text summarization: approaches, datasets, evaluation measures, and challenges." Mathematical problems in engineering 2020 (2020): 1-29.

[2] Xu, Weiran, Chenliang Li, Minghao Lee, and Chi Zhang. "Multi-task learning for abstractive text summarization with key information guide network." EURASIP Journal on Advances in Signal Processing 2020, no. 1 (2020): 1-11.

[3] Fabbri, Alexander R., Wojciech Kryściński, Bryan McCann, Caiming Xiong, Richard Socher, and Dragomir Radev. "Summeval: Re-evaluating summarization evaluation." Transactions of the Association for Computational Linguistics 9 (2021): 391-409.

[4] Kahla, Mram, Zijian Győző Yang, and Attila Novák. "Cross-lingual fine-tuning for abstractive Arabic text summarization." In Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021), pp. 655-663. 2021.

[5] Tomer, Minakshi, Manoj Kumar, Adeel Hashmi, Bharti Sharma, and Uma Tomer. "Enhancing metaheuristic based extractive text summarization with fuzzy logic." Neural Computing and Applications 35, no. 13 (2023): 9711-9723.

[6] Mridha, Muhammad F., Aklima Akter Lima, Kamruddin Nur, Sujoy Chandra Das, Mahmud Hasan, and Muhammad Mohsin Kabir. "A survey of automatic text summarization: Progress, process and challenges." IEEE Access 9 (2021): 156043-156070.

[7] Widyassari, Adhika Pramita, Supriadi Rustad, Guruh Fajar Shidik, Edi Noersasongko, Abdul Syukur, and Affandy Affandy. "Review of automatic text summarization techniques and methods." Journal of King Saud University-Computer and Information Sciences 34, no. 4 (2022): 1029-1046.

[8] Ye, Xia, Zengying Yue, and Ruiheng Liu. "MBA: A multimodal bilinear attention model with residual connection for abstractive multimodal summarization." In Journal of Physics: Conference Series, vol. 1856, no. 1, p. 012070. IOP Publishing, 2021.

[9] El-Kassas, Wafaa S., Cherif R. Salama, Ahmed A. Rafea, and Hoda K. Mohamed. "Automatic text summarization: A comprehensive survey." Expert systems with applications 165 (2021): 113679.

[10] Tsai, Chih-Fong, Kuanchin Chen, Ya-Han Hu, and Wei-Kai Chen. "Improving text summarization of online hotel reviews with review helpfulness and sentiment." Tourism Management 80 (2020): 104122.

[11] Zhang, Jian, Xu Wang, Hongyu Zhang, Hailong Sun, and Xudong Liu. "Retrieval-based neural source code summarization." In Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering, pp. 1385-1397. 2020.

[12] Lewis, Mike, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov, and Luke Zettlemoyer. "Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension." arXiv preprint arXiv:1910.13461 (2019).

[13] Kryściński, Wojciech, Bryan McCann, Caiming Xiong, and Richard Socher. "Evaluating the factual consistency of abstractive text summarization." arXiv preprint arXiv:1910.12840 (2019).
[14] Xu, Jiacheng, Zhe Gan, Yu Cheng, and Jingjing Liu. "Discourse-aware neural extractive text summarization." arXiv preprint arXiv:1910.14142 (2019).
[15] Mohamed, Muhidin, and Mourad Oussalah. "SRL-ESA-TextSum: A text summarization approach based on semantic role labeling and explicit semantic analysis." Information Processing and Management 56, no. 4 (2019): 1356-1372.
[16] A. M and S. M. Rajgopal, "Exploring Unique Techniques to Preserve Confidentiality and Authentication," 2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT), Bengaluru, India, 2024, pp. 440-447, doi: 10.1109/IDCIoT59759.2024.10467248.
[17] Nair, A.R., Singh, R.P., Gupta, D. and Kumar, P., 2024. Evaluating the Impact of Text Data Augmentation on Text Classification Tasks using DistilBERT. Procedia Computer Science, 235, pp.102-111.
[18] Paul, Pretty, and Rimjhim Padam Singh. "Sentiment Rating Prediction using Neural Collaborative Filtering." In 2022 IEEE 7th International Conference on Recent Advances and Innovations in Engineering (ICRAIE), vol. 7, pp. 148-153. IEEE, 2022.
[19] Kavitha C.R.Rajarajan S J, R. Jothilakshmi, Kamal Alaskar, Mohammad Ishrat, V. Chithra, Study of Natural Language Processing for Sentiment Analysis, 2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN), 19-20 June 2023. 2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN) 19th-20th June 2023
[20] Vidya Kumari K. R., Kavitha C. R., Data Mining for the Social Awareness of the Social Networks, 3rd International Conference on Computational System and Information Technology Sustainable Solutions (CSITSS 2018), December 2018, pp: 7-17.
[21] Mridula A., Kavitha C. R., Opinion Mining and Sentiment Study of Tweets Polarity Using Machine Learning, Proceedings of 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT 2018), April 2018, pp: 621-626.
[22] Venkataramani, Eknath and Gupta, Deepa. (2010). English-Hindi Automatic Word Alignment with Scarce Resources. 253-256. 10.1109/IALP.2010.5