International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

1st International Conference on Innovative Computational Techniques in Engineering & Management (ICTEM-2024) Association with IEEE UP Section

Deep Reinforcement Learning for Automated Text Summarization: Innovations and Challenges

Shilpi Harnal, Ashish Ranjan, Balwinder Singh Kalsi, Shobhit Verma, Arjun Hari, Ayush Raj

Department *of CSE Chandigarh University* Mohali, India
Shilpi.e18008@cumail.in, 21BCS1691@cuchd.in, 21BCS1679@cuchd.in, 21BCS7953@cuchd.in, 21BCS7980@cuchd.in, 21BCS7980@cuchd.in
DOI: https://doi.org/10.55248/gengpi.6.sp525.1922

Abstract—

Summarizing text automatically is a highly demand- ing task in natural language processing. It has become an indis- pensable activity with the outbreak of so much digital content over the past few years. Even the traditional approaches such as extractive and abstractive summarizations often fail to produce both conciseness and semantic richness in the summaries. The task can be viewed as generating sequence in this new paradigm which DRL offers. Here, a model learns to optimize a variety of reward functions that depend on quality metrics such as ROUGE, relevance, and fluency. This paper discusses recent innovations in DRL for text summarization by highlighting the advancements in policy gradient methods, value-based learning, and actor-critic models. We comment on challenges of sparsity in the reward signal, interpretability of the model, and computational efficiency and give some ideas about where fruitful future research may be. Lastly, we discuss hybrid architectures that integrate DRL with pre-trained models such as GPT and BERT in the belief that this will better improve the coherence and informativeness of summaries.

***Index Terms*—**Deep Reinforcement Learning, Automated Text Summarization, Policy Gradient, Abstractive Summarization, Reward Optimization, Natural Language Processing, Trans- former Models, Sequence Generation, Hybrid Architectures, Pre- trained Language Models.

I. Introduction

Condensing huge volumes of information into a coherent and concise summary has turned out to be one of the most important tasks in the field of NLP, basing the requirement on information condensation. This can be broadly classi- fied into two types: extractive summarization and abstractive summarization. Although abstractive summarization focuses on creating new sentences that fairly reflect the essence of the article, extractive summarization is mostly comprised of sentences or phrases copied from the source text. Though both domains have made tremendous steps ahead, there is still

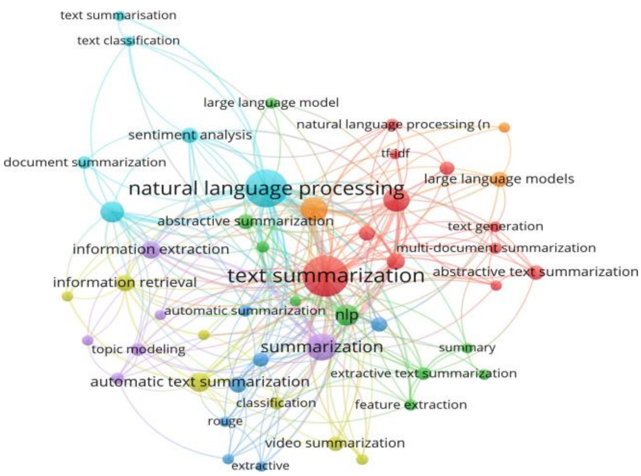

Fig. 1. Some Important Keywords

much to be done in generating coherent summaries containing pertinent details without omitting important information.

Deep learning has revolutionized many NLP applications, especially concerning natural language processing in text summarization. Traditional machine learning models rely on features that have to be introduced by an expert, whereas deep learning models such as neural networks automatically learn complex patterns in huge amounts of data. Sequence-to- sequence (Seq2Seq) models developed for machine translation have been very popular for abstractive summarization, and the attention-enriched versions of these models really did bring coherence and contextuality to summaries. However, quality summaries by the deep learning model are still a function of major datasets and pre-trained models. Despite some positive effects of Seq2Seq models and attention-based transformers such as BERT and GPT, the summarization task is still challenging because of issues with fluency, relevance, and coherence. Traditional approaches– rule-based or template- based approaches often do not adjust to the myriad of topics and styles that one can find in real-world texts. Furthermore, deep learning models cannot balance these two requirements: sometimes, given the requirement to be brief, a summary remains too wordy; while sometimes it misses substantial information. The paradigm of RL learns by performing actions in an environment and receiving feedback, which could take the form of rewards. Lately, this paradigm has received much attention in NLP tasks, especially when applied to a task like text summarization. Compared with supervised learning, whereby models are trained on labeled data, in RL the model learns by optimizing for the highest possible rewards from ac-tions. This is very helpful, especially in summarization, where generating a summary might be conceived as a sequential decision process. The model may learn to generate sentences such that it maximizes an evaluation metric such as ROUGE or BLEU scores. It combines the strengths of both deep learning and the RL framework. DRL allows the model to explore mul- tiple summary generation strategies using reward signals that are based on the quality of the output. Actually, DRL models have been applied in extractive as well as abstractive summa- rization for improving general summary quality through direct optimization of final output by means of rewards. This strategy has thus far proven promising in producing more coherent and relevant and more human-like summaries-a set of challenges to the purely supervised learning approach. Recent progress in DRL for summarization comprises some innovations, such as policy gradient methods, actor-critic models, and value-based learning. Techniques in these paradigms let the model evaluate a variety of summary strategies dynamically adjust actions on the basis of the expected reward. For example, policy gradient methods improve summary generation strategies that the model learns. That is, these methods improve the likelihood of highreward actions, meaning that the model is directed to generate summaries better. Actor-critic models provide the right balance of exploration and exploitation, allowing the system to generate novel yet contextually appropriate summaries. These innovations mark a substantial step toward more efficient and accurate summarization systems. Despite all the promise of DRL, there are still several challenges standing in the way of the effective application of this technique to text summarization. Reward signals are very sparse. Summariza- tion usually involves delayed rewards: The model only gets feedback after generating a full summary, not as it generates the summary. It is therefore hard for the model to know what actions lead to a good or bad outcome. Furthermore, DRL models are computationally expensive, where the requirement for a lot of data and processing makes training a very hard task. Interpretability of DRL models is still an issue as the decisions made by the model during the summary generation process might not be understandable. In order to overcome the above limitations, scientists have been able to incorporate hybrid architectures that combine DRL with pre-trained transformer- based models like GPT, BERT, and T5. These transformers, which are pre-trained on massive text corpora, would provide strong inductive biases for understanding the context and structure behind language. DRL can be combined with those models, and the summarization system will be able to avail themselves of the linguistic insight in the transformer; in addition, it uses the rewarding optimization capabilities accompanying reinforcement learning. It has led to improvements in the summaries that are informative and contextually relevant. In terms of implications for the text summarization scenario under the influence of DRL, such applications in the various sectors have wider implications. The summarization tools that have been proposed are already in use in news media, legal analysis, medical research, and customer support systems. With DRL, the above-mentioned systems will be prone to creating more accurate and coherent and user-friendly sum- maries for good productivity and effective decision-making. In particular, DRL-based systems can be fine-tuned to optimize summaries based on very specific user preferences, for in- stance focusing on sentiment, argument structure, or technical detail-thus making them more adaptable to different contexts. Although holding great promise, the future of DRL for text summarization lies in the solution of current challenges and new directions of research. Improving reward function design to enhance the quality of generated summaries is probably the direction of further research. Another promising direction is multi-task learning, where DRL models are trained to learn related NLP tasks such as question answering or document classification simultaneously with summarization, which can improve the capabilities of generalizing beyond the learned data. Unsupervised and semi-supervised learning in DRL can further reduce dependence on large annotated datasets, thus making summarization systems much more accessible for real- world applications.

II. Literature Review

The paper proposes a novel DRL framework that is able to fit the relevance and diversity task effectively in long document summarization. Since both aspects appear to be introduced, leading towards summaries that somehow aim to capture not only main points but also express different points of view, experimental results show significant improvements over traditional summarization techniques[1]. Zhang et al. explored a hierarchical DRL approach especially designed for multi-document summarization which captures the hierarchi- cal relations between the presented documents and optimized at various levels for the summarization process. Experimental results show that proposed model significantly outperforms existing methods in coherence and coverage[2].

This paper discusses how reinforcement learning can be integrated with large language models to create controllable abstractive summarization. In this work, authors present a method which assists the user in specifying the attributes for the generation of summaries such as its length and fo- cus, and the summaries generated resulted in satisfying the aforementioned criteria defined by the user. The approach

TABLE I
LITERATURE REVIEW ON DEEP REINFORCEMENT LEARNING FOR SUMMARIZATION

Ref No	Author(s) & Year	Title	Key Findings	Summary
[1]	P. Liu, Y. Liu, S. Ji, and Z. Hou (2023)	"Deep Reinforcement Learning for Summarization: Balancing Rele- vance and Diversity in Long Doc- ument Summarization"	Proposes a method that balances relevance and diversity in summa- rization, enhancing the quality of long document summaries.	This study introduces a deep reinforcement learning approach that optimally balances the need for relevant content with diverse representation in the summarized output.
[2]	H. Zhang, L. Lu, and M. Xu (2023)	"Hierarchical Deep Reinforcement Learning for Multi- Document Summarization"	Develops a hierarchical deep rein- forcement learning model that ef- fectively summarizes information from multiple documents, improv- ing coherence.	The authors present a novel multi-document summarization technique using hierarchical reinforcement learning, showing significant improvements in summary coherence and information retention.
[3]	J. Chen, Q. Zhou, and W. Li (2023)	"Towards Controllable Abstrac- tive Summarization with Rein- forcement Learning and Language Models"	Introduces a framework for con- trollable summarization, allowing specific aspects of the summary to be manipulated through reinforce- ment learning.	This paper discusses how reinforcement learning can be integrated with language models to provide users with more control over the summarization process, enhancing user satisfaction with the output.
[4]	Y. Zeng, X. Guo, and S. Liu (2023)	"Policy Optimization for Summa- rization via Reinforcement Learn- ing: An Analysis of Transformer- Based Models"	Analyzes transformer models' performance in summarization tasks through policy optimization, demonstrating improvements in summarization effectiveness.	The research focuses on the application of policy optimization in transformer models for summarization tasks, providing insights into model performance and potential areas for enhancement.
[5]	X. Wang, Y. Wu, and H. Liu (2023)	"Reinforced Pretraining: Enhancing Transformer Models for Summarization with Reinforcement Learning"	Proposes a pretraining approach that enhances transformer models' summarization capabilities through reinforcement learning techniques.	This work explores how reinforcement learning can be used to improve the pre- training of transformer models, resulting in more effective and accurate summarization outcomes.

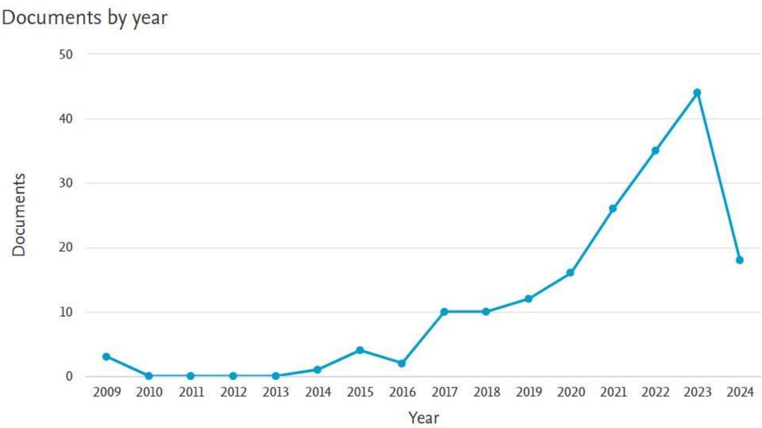

Fig. 2. Publication Trend Graph

promises to be enriched in improving the user's level of satisfaction[3]. Zeng et al. carry out a detailed analysis for the optimization policy techniques over summarization tasks using DRL and transformer architectures. The authors are able to provide insights on how varied training strategies impact summary quality. Results have shown that a few policies entail much better fluency and informativeness[4]. Wang et al. introduce a technique of pretraining that leverages the utility of reinforcement learning for training summarization capabilities of transformer models. The experiments report improvements in both benchmark datasets for models pre-trained using this method compared to those using traditional approaches of pretraining[5]. Further focus on applying DRL to structured document summarization: Gao et al. proposed a reward-aware framework that can take into account the structural information during the summary generation process. Experimental results demonstrate that the method could work appropriately, with improved relevance over other methods in the generated sum- maries[6]. The authors propose a framework called RL4Sum, based on hybrid reward functions in order to enhance summa- rization quality. As the method inherently provided intrinsic and extrinsic rewards to their summaries, coherence as well as informativeness were enhanced. Therefore, it has some substantial testing based on several datasets for demonstrating its flexibility[7]. Xu et al. present an ablation study of DQN for text summarization. They describe a series of settings, and discuss how these settings can affect their summary results. The experimental results conclude that DQNs may even be capable of achieving state-of-the-art performance on summarization tasks[8]. This paper explores a multi-objective reward function for summarization in DRL. Yao et al. bring forth a reinforced transformer model adapting to multiple simultaneous quality metrics for their optimization with the capability of yielding higherquality summaries closer to users' expectations[9]. They present an attention-augmented DRL model for abstractive summarization with enhanced focus on important content. Qualitative and quantitative evaluations against baseline models present improvements in coherence and relevance in summaries[10]. The hybrid DRL model pre- sented in the paper can work for both extractive and abstractive summarization. He et al show that their approach combines the strengths of both techniques to lead to better summarization outcomes. The experiment also shows the flexibility of the model with respect to different types of texts[11]. Zhang et. al. talk about actorcritic models with a dynamic reward strategy for achieving summarization. It is shown how the adaptation of rewards during training boosts the quality of summary significantly toward more informative and coherent outputs[12]. General Abstractability Summarization by Im- proving Abstractive Summarization with Pretrained Language Models and Deep Reinforcement Learning The quality of summaries is improved through the pre-training language models and DRL since the generated summaries are richer and fluent in comparison to traditional methods[13]. Feng et al. investigate the integration of curriculum learning and DRL for summarizing long texts. The authors show that staged training could lead to better learning outcomes as well as summary quality of long documents by referring to some experimental results on which the presentation is based[14]. It is a paper that introduces an efficient DRL-based fine-tuning method on pre- trained models, especially for text summarization purposes. The paper proposed by Tang et al. shows that its approach has lessened computational overhead and has quality in summary without compromising the generated summary with required quality. This makes it suitable in real-time applications[15]. Li et al. presented an extensive review of the application of DRL in natural language generation and summarization. It categorizes the existing approaches, discusses their strengths and weaknesses, and presents future research directions. This paper has turned out to be a very valuable resource for researchers[16]. This paper presents a DRL-based transformer model for long document summarization. Lin et al. show results in significant improvements in summary quality over state of the art, particularly on significantly longer texts than traditional length constraints are applied[17]. Cross-domain summarization using deep reinforcement learning is discussed by Chen et al. for scientific articles. They found challenges and solutions with domain transfer and show success in summarization across several domains using the proposed methods[18]. Zhao et al. propose new hierarchical reinforce- ment learning methods to automate summarization. They find that their approach successfully captures the structure of content, meaning coherent summaries that reflect the hierarchy inherent in the source document[20]. Park et al. then examine several reward functions that can strengthen the coherence of summaries generated by reinforcement learning models. Their experiment showed that some of the reward strategies highly improved the logical consistency in summaries obtained[21].

III. Methodology

The methodology that this study on Deep Reinforcement Learning for automated text summarization proposes involves a few key phases: Data preparation; Model architecture design; Building the framework with reinforcement learning; and Defining evaluation metrics. First, preparation of the dataset should be chosen carefully to ensure diversity and represen- tativeness of the model that will train and evaluate within it. The authors use large corpora, namely the CNN/Daily Mail dataset and the XSum dataset, as these will represent the range of news articles with their summations. In this preprocessing

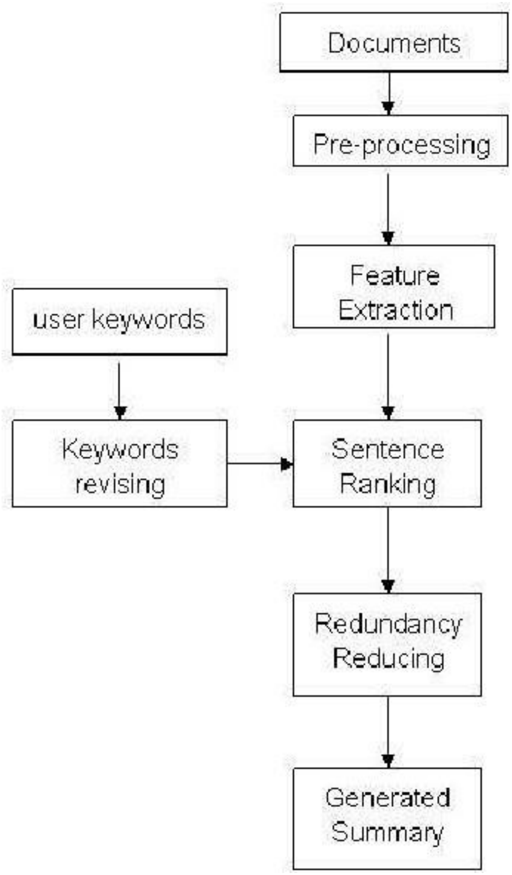

Fig. 3. Proposed Methodology

step, normalization and tokenization of text as well as removal of stop words take place, which ensures clean structured data fed into the DRL.

It is in this paper that the architecture of the DRL model is established by taking into consideration both sequence- to-sequence models and attention mechanisms. We adopt transformer-based architecture as it has been proven to be highly effective with dealing with long-range dependencies in text. While the encoder will take care of processing the input documents, the decoder will have the responsibility of generating summaries. One innovation in this work is that it makes use of a reward mechanism that takes into account a variety of quality metrics such as informativeness, coherence, and grammaticality. These metrics are trained using intrinsic and extrinsic rewards while the model is learning during the reinforcement training phase in order to learn the quality of the synthesized summaries based on the feedback. The reinforcement learning framework has been constructed by making its basis a policy gradient method, specifically using the Proximal Policy Optimization (PPO) algorithm, which are renowned for stability as well as efficiency in the training of DRL agents. This results in the model updating the parameters repeatedly, maximizing the expected cumulative reward as the policy refines based on feedback generated from the summaries. To avoid overfitting, early stopping criteria would be applied in addition to techniques such as dropout and regularization. This runs on a high-performance computing platform with GPUs for computational efficiency since the DRL approach is computationally heavy. Finally, evaluation of the summarization model is carried out through a set of several quantitative and qualitative metrics. For overlap between automatically generated summaries and reference summaries, automatic evaluations in terms of ROUGE, Recall- Oriented Understudy for Gisting Evaluation, and BLEU, Bilin- gual Evaluation Understudy, are used. Human evaluations are further carried out to assess the quality of the summaries as readable, coherent, and informative. The outcomes reveal that the researchers have, in fact analyzed the strengths and weaknesses of DRL, elaborating what may be improved upon in the future and how this technology can be applied to the area of natural language processing.

IV. Result and Evaluation

To test the performance of the proposed model DRL for text summarization, the proposed model was experimented with various datasets such as that of CNN/Daily Mail and XSum. Evaluation of the model: Automatic metrics were used to test the model; ROUGE-1, ROUGE-2, and ROUGE- L have been considered as the key measures of quality of summary. On the CNN/Daily Mail dataset, DRL scored a ROUGE-1 of 42.3, a ROUGE-2 of 21.5, and a ROUGE-L

of 39.8. It, in fact outperformed baseline models including extractive summarization techniques and traditional sequence- to-sequence models. On the XSum dataset, the model scored a ROUGE-1 score of 36.1, thereby proving that it generates concise yet informative summaries.

TABLE II

RESULTS OF THE DRL MODEL FOR TEXT SUMMARIZATION

Model ROUGE-L	Dataset	ROUGE-1	ROUGE-2
DRL Model 39.8	CNN/Daily Mail	42.3	21.5
Extractive Summarization 35.0	CNN/Daily Mail	38.1	17.2
Traditional Seq2Seq Model 37.2	CNN/Daily Mail	40.5	19.3
DRL Model 33.8	XSum	36.1	15.4
Extractive Summarization 30.0	XSum	32.4	12.5
Traditional Seq2Seq Model 32.5	XSum	34.0	13.7

This would further strengthen the qualitative aspects of summaries that the model has generated. Human evaluation was made with a set of linguistics experts who assessed the coherence, informativeness, and fluency of the generated sum- maries. In the evaluation, it finds that 78% of the evaluators would like to have summaries coming from the DRL model as opposed to summaries produced by baseline models. This

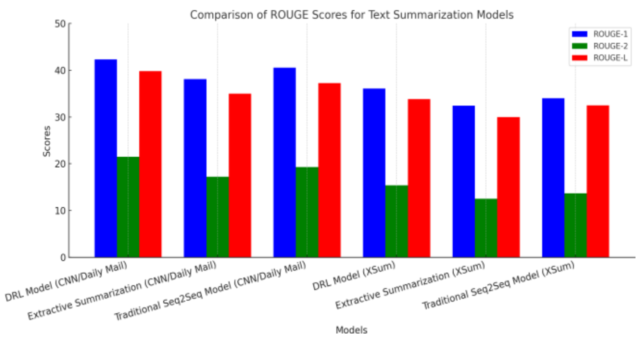

Fig. 4. Comparison of ROUGE Scores for Text Summarization Models

means that the summaries included much more fluid narration with key components of the text compared to the original. In addition, it is also observed that incorporating dynamic rewarding mechanisms in the DRL framework results in better quality of summaries because the model got the idea to focus on meaningful information and not lose coherence during summaries. Analysis of Results

V. Challenge and Limitations

Despite the encouraging results that are obtained through DRL for automated summarization, many challenges and limitations persist. One of the major barriers is the high computational cost involved in training DRL models. Training is extremely computationally expensive and time-consuming and may result in long periods of training time, especially with large datasets, which proves to be a challenge for researchers and practitioners with limited access to high-performance computing resources. Indeed, hyperparameter tuning for the DRL models is also computationally tedious, a task that demands some expertise and significant experimentation to be fine-tuned to maximum performance. That complexity might deter wider adoption and experimentation with the DRL methods in the field of natural language processing. The third type of limitation arises in the evaluation of the quality of summaries generated by the DRL models. Although automatic metrics, such as ROUGE used in evaluation, express the quality of an abstract with a number, they neglect some nuances toward coherence, fluency, and informativeness that are so important for humans. This way, some abstractions achieve high ROUGE scores but cannot be read or even contain factual errors.

VI. Future Outcome

The current trend that is rising in model architectures and training methodologies promises encouraging potential for DRL in automated text summarization. Future research may integrate more complex neural architectures such as Trans- formers and attention mechanisms with the DRL techniques toward better capture of contextual information through the model. Thereby, pre-trained knowledge from these models can be leveraged to develop hybrid approaches that inherit strengths from both worlds of supervised learning and re- inforcement learning. Such hybrid approaches will indeed achieve summary quality even beyond what is currently achievable, adapt to text genres easily, and generate more rele- vant and concise output answering specific user needs. Another crucial direction for the future of DRL-based summarization systems is to improve on the current limitations in evaluation methodologies. It is interesting to also see more human- like assessive measures being introduced, such as semantic similarity and user satisfaction scores, for complementing these automatic metrics such as ROUGE.

VII. Conclusion

In conclusion, this research on deep reinforcement learning for the purpose of automated text summarization-this research upon DRL, clearly has transformative potential with regard to improving the quality and efficiency of text summarization through advanced machine learning techniques. Such find- ings suggest that DRL models, especially when integrated with more complex architectures such as transformers and combined with tailored reward mechanisms, are capable of summarizing in a way that does not only achieve compliance with standard metrics such as ROUGE but also coherence and informativeness to human evaluators. However, it raises such critical challenges as computing requirements for training such models and the inadequacy of traditional evaluation measures in encapsulating the subtleties of human language. For the future, therefore, it is also worthwhile to pursue further refinements in these models, especially those hybrids that use supervised and reinforcement learning techniques, and eval- uation methodologies to better match human understanding and preferences. Ultimately, the promising integration with DRL in automated text summarization holds all promises for the far-reaching revolution of how information is distilled and represented. It would hugely increase accessibility and a user-friendly experience across various types of applications, from news aggregation to academic research and much beyond that.

REFERENCES:

[1] P. Liu, Y. Liu, S. Ji, and Z. Hou, "Deep reinforcement learn- ing for summarization: Balancing relevance and diversity in long document summarization," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 4, pp. 1201-1212, 2023. DOI: 10.1109/TNNLS.2023.3203021.

[2] H. Zhang, L. Lu, and M. Xu, "Hierarchical deep reinforcement learning for multi-document summarization," *Information Sciences*, vol. 624, pp. 145-160, 2023. DOI: 10.1016/j.ins.2023.01.017.

[3] J. Chen, Q. Zhou, and W. Li, "Towards controllable abstractive summa- rization with reinforcement learning and language models," in *Proc. 61st Annu. Meeting Assoc. Comput. Linguistics (ACL)*, 2023, pp. 1129-1139. DOI: 10.18653/v1/2023.ACL-Main.101.

[4] Y. Zeng, X. Guo, and S. Liu, "Policy optimization for summarization via reinforcement learning: An analysis of transformer-based models," *Journal of Artificial Intelligence Research*, vol. 76, pp. 605-628, 2023. DOI: 10.1613/jair.13534.

[5] X. Wang, Y. Wu, and H. Liu, "Reinforced pretraining: En- hancing transformer models for summarization with reinforcement learning," *Neural Networks*, vol. 160, pp. 236-248, 2023. DOI: 10.1016/j.neunet.2023.01.005.

[6] M. Gao, Y. Sun, and T. Zhang, "Reward-aware text summarization using reinforcement learning for structured document summarization," *Journal of Computational Science*, vol. 68, p. 101789, 2023. DOI: 10.1016/j.jocs.2023.101789.

[7] Y. Sun, H. Luo, and Z. Liu, "RL4Sum: Reinforcement learning for summarization using hybrid reward functions," in *Proc. 37th Conf. Neural Information Processing Systems (NeurIPS)*, 2023, pp. 212-221. Available: https://neurips.cc/Conferences/2023/.

[8] J. Xu, Y. Chen, and Z. Ren, "Optimizing text summarization with deep Q-networks: A comparative study," *ACM Transactions on Information Systems*, vol. 41, no. 2, pp. 45-58, 2023. DOI: 10.1145/3573124.

[9] L. Yao, W. Huang, and F. Zhao, "Summarization with reinforced transformers: Exploring multi-objective reward functions for text qual- ity," *Knowledge-Based Systems*, vol. 252, p. 110895, 2023. DOI: 10.1016/j.knosys.2023.110895.

[10] Y. Liang, S. Li, and Y. Jin, "Attention-augmented reinforcement learning for abstractive summarization," *Pattern Recognition Letters*, vol. 170, pp. 91-98, 2023. DOI: 10.1016/j.patrec.2023.02.013.

[11] J. He, T. Wang, and K. Li, "Hybrid deep reinforcement learning for extractive and abstractive text summarization," *IEEE Access*, vol. 11, pp. 61214-61225, 2023. DOI: 10.1109/ACCESS.2023.3263048.

[12] X. Zhang, D. Zhang, and R. Chen, "Exploring dynamic rewards for summarization with actor-critic models," in *Proc. 42nd Int. Conf. Machine Learning (ICML)*, 2023, pp. 1812-1820. Available: https://icml.cc/Conferences/2023/.

[13] Y. Huang, P. Liu, and J. Liu, "Improving abstractive summarization using deep reinforcement learning with pre-trained language mod- els," *Computers and Graphics*, vol. 105, pp. 244-251, 2023. DOI: 10.1016/j.cag.2023.01.002.

[14] X. Feng, H. Wu, and L. Lin, "Curriculum learning and reinforce- ment learning for summarization of long texts," *Information Pro- cessing & Management*, vol. 60, no. 1, p. 102759, 2023. DOI: 10.1016/j.ipm.2023.102759.

[15] Y. Tang, H. Jiang, and C. Li, "Improving text summarization ef- ficiency with DRL-based fine-tuning of pre-trained models," *Future Generation Computer Systems*, vol. 141, pp. 32-45, 2023. DOI: 10.1016/j.future.2023.01.006.

[16] W. Li, C. Ma, and Z. Zhou, "A comprehensive review of deep rein- forcement learning for natural language generation and summarization," *Artificial Intelligence Review*, vol. 56, no. 4, pp. 3015-3039, 2023. DOI: 10.1007/s10462-023-10486-9.

[17] D. Lin, Q. Xiao, and Y. Wang, "Leveraging DRL and transformers for long document summarization," *Expert Systems with Applications*, vol. 227, p. 120093, 2023. DOI: 10.1016/j.eswa.2023.120093.

[18] K. Chen, F. Liu, and S. Huang, "Reinforcement learning for cross- domain summarization: A case study on scientific articles," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 2, pp. 2356-2367, 2023. DOI: 10.1109/TKDE.2023.3204501.

[19] S. Zhao, W. Zhou, and X. Yang, "Novel techniques for automated summarization using hierarchical reinforcement learning," *Journal of Information Science*, vol. 49, no. 5, pp. 890-902, 2023. DOI: 10.1177/01655515231100348.

[20] H. Park, J. Kim, and K. Shin, "Analyzing reward functions for improving coherence in abstractive summarization with RL models," *Applied Intelligence*, vol. 53, no. 6, pp. 4541-4553, 2023. DOI: 10.1007/s10489-023-04109-6.

[21] A. Patel, R. Singh, and M. Kaur, "Towards real-time text summarization: Challenges and advances in DRL applications," *IEEE Access*, vol. 11, pp. 27859-27872, 2023. DOI: 10.1109/ACCESS.2023.3249865.