Enhancing Text Summarization Capability of Lightweight Models through Dynamic Direct Preference Optimization(DPO) Mechanism

Anonymous ACL submission

Abstract

The abstractive summarization is a natural language processing(NLP) task that involves generating concise summaries of longer documents while preserving key information. Currently, state-of-art summarization methods are dominated by large language models (LLMs), their strong understandings, and generalizations have reshaped summarization research. Unlike those works, we focus on developing a light yet efficient abstractive summarizer targeting for edge-device applications. The primary challenge lies in the limited context understanding and paraphrasing abilities of lightweight models, constrained by their smaller capacity and vocabulary size. To address this, we introduce a novel framework integrating an online feedback mechanism. This system incorporates improvement suggestions to dynamically adjust the model's outputs, enhancing its learning capabilities. Our approach achieves state-of-the-art (SOTA) results on CNN/DailyMail and XSum, outperforming backbones by 19.3% and 12.9%, respectively.

1 Introduction

Abstractive summarization, which produces succinct, novel summaries, has surpassed extractive techniques by enabling more human-like outputs. This shift is largely attributed to advancements in Large Language Models (LLMs), such as GPT-4 (OpenAI et al., 2024) and LLaMA (Touvron et al., 2023). However, these models are often computationally expensive, exceeding the capabilities of edge hardware thus limiting their deployment in resource-constrained environments. (Tan et al., 2024) utilized quantization techniques to balance the model size and performance while (Ge et al., 2022) focusing on cost-effective parameterization methods for edge-device deployment. To address the challenges of deploying summarization models on resource-constrained devices, a common strategy involves using lightweight models. However,

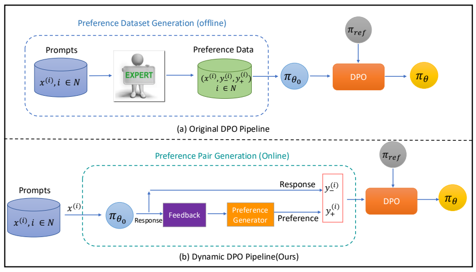

Figure 1: **DPO Pipeline Comparison.** The traditional DPO pipeline shows it upper part (a) that requires to build a preference data forehand, while our DPO pipeline showing in the bottom (b) targeting for dynamic preference pair generation to avoid such data preparation. $x^{(i)}, y_-^{(i)}, y_+^{(i)}$ represent prompt/input, dislike response, preference response, respectively, where $i \in N$, N is number of samples. $\pi_{\theta_0}, \pi_{ref}, \pi_\theta$ represent base model, reference model, and aligned model.

these models often struggle with complex linguistic patterns due to their limited capacity, leading to suboptimal performance when trained via direct supervised learning. Several studies (Jung et al., 2024; Jiang et al., 2024; Pham et al., 2023; Xu et al., 2023) have focused on enhancing lightweight models through advanced knowledge distillation techniques, leveraging the generalization power of LLMs. Despite these efforts, distilled models may still fail to retain critical long-range dependencies and contextual nuances, resulting in generic or factually inconsistent summaries. To mitigate this, researchers have turned to Reinforcement Learning (RL) with human feedback (Paulus et al., 2017; Stiennon et al., 2020), enabling models to make sequence-level decisions that improve coherence and relevance. Additionally, Direct Preference Optimization (DPO) (Choi et al., 2024) has emerged as a cost-effective alternative, bypassing the need for dense reward signals and human feedback.

To enhance the learning and generation capabilities of lightweight models for on-device applications, we introduce a novel framework integrating Direct Preference Optimization (DPO). Traditional

DPO training, as depicted in Figure 1 (a), relies on a preference dataset requiring extensive annotations for each input. This approach has limitations: 1) generating multiple preference per input increases labeling costs; 2) Preferences are pre-generated and fixed, becoming outdated as the model evolves, reducing adaptability. To address these challenges, we propose an online feedback mechanism with dynamic DPO training strategy (see Figure 1 (b)): 1) an automatic scoring module evaluates current responses $y_-^{(i)}$ in real-time; 2) a LLM generates updated preference $y_+^{(i)}$ conditioned on the scoring feedback, enabling dynamic adjustment as the model improves.This approach reduces annotation requirements and ensures preferences remain relevant, aligning with the model's latest state. Further details are provided in the subsequent sections.

In summary, our proposed method has following contributions:

1. We propose a novel framework that unifies the text summary generation, scoring mechanism, feedback mechanism, and summary re-generation into a one-stage learning process, significantly enhance the generation and learning capability for lightweight models.

2. To our best knoweledge, this is the first work to introduce dynamic DPO training concept where the preference response is adaptively generated according to the real-time feedback, that significantly improving the model performance by enforcing the generated response to timely align with its preference.

2 Related Work

Abstractive summarization, which requires generating novel sentences, was initially tackled using encoder-decoder models with attention (Nallapati et al., 2016a). The pointer-generator network (See et al., 2017) improved factual consistency by enabling token copying from the source. Subsequent models like BART (Lewis et al., 2020) and PEGASUS (Zhang et al., 2020) demonstrated strong performance across benchmarks such as CNN/DailyMail and XSum. From 2022 onward, research emphasized LLM-based summarization. ChatGPT and GPT-3.5 were applied using prompt-based methods, as seen in SummIt (Zhang et al., 2023), where iterative summarization improved coherence. Zhang et al. benchmarked several LLMs on summarization tasks (Zhang et al., 2024), showing that while LLMs are fluent, they often lack factual accuracy. To improve factuality, methods such as textual entailment reward modeling during RLHF (Roit and Reichart, 2023) and structured preference learning via DPO (Rafailov et al., 2023) have emerged. Recent research has also focused on summarization in specialized domains. Balde et al. addressed biomedical summarization with a vocabulary-controlled model (Balde et al., 2024), while Zaman et al. proposed SATSUM for scientific texts (Zaman et al., 2024).

3 Approach

The overview is described in Figure 2. The proposed framework integrates generation, scoring, feedback, and re-generation into a unified one-stage learning process, where an automatic scoring mechanism evaluates generated summaries based on customer-defined criteria, leveraging a Large Language Model (LLM)-based judgement process for precise assessments. Additionally, an online feedback mechanism is designed to provide real-time feedback on current summaries according to the improvement suggestions from scoring module, and dynamically adjust preferences as the model updates. This mechanism incorporates an expert preference generator for creating adaptive preferences and a reward system to optimize the model in alignment with these preferences.

3.1 Automatic Scoring Mechanism

This mechanism is designed to identify areas for improvement in the current response relative to the input text. The output, in the form of improvement suggestions, is fed into the online feedback system to generate dynamic preferences. To ensure precision, we replace generic prompts with a detailed scoring rubric that guides the LLM-based judger to evaluate the response progressively. This structured approach ensures that the feedback accurately reflects the quality of the current response, enabling more targeted and effective model updates. As demonstrated in Figure 2 bottom left, *"whether the given summary refers..., whether the given summary is..."* are example rubrics introduced in system prompt to specify the role and task for this LLM judger.

3.2 Online Feedback Mechanism

In this mechanism, an LLM-based preference generator creates an improved summary (preference response) based on improvement suggestions from the scoring module. However, dynamically updating preferences over time introduces challenges: 1)

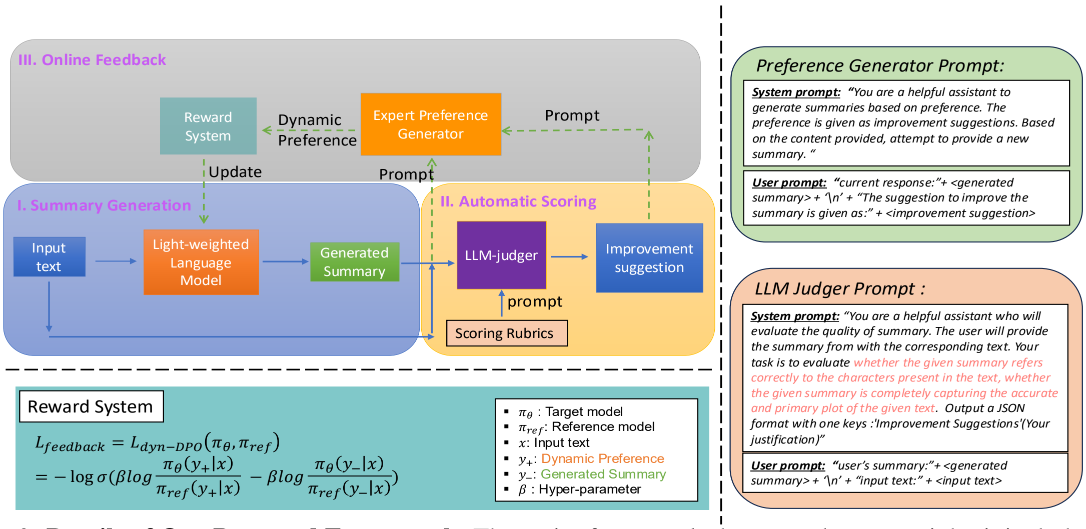

Figure 2: **Details of Our Proposed Framework.** The entire framework shows on the upper right, it includes three modules: summary generation, automatic scoring and online feedback mechanism. The summarizer is the target light-weighted language model which is compatible to the edge-device applications, while our judger and preference generator are LLM-based models. On the bottom left, it demonstrates how is reward system incorporated with dynamic preference. In addition, the prompt template of both LLM judger and preference generator are presented on the right side.

Oscillations: the model may struggle to converge due to frequent preference changes; 2) Lack of Long-Term Consistency: older feedback might be overlooked, hindering the model's ability to learn consistent behaviors. To address these issues, as illustrated in Figure 2 (upper right), the preference generator is guided by three inputs: 1) Current Response: Ensures alignment with the model's latest output; 2) Original Input Text: maintains focus on the primary content, reducing the risk of training oscillations; 3) Improvement Suggestions: Provides specific guidance to address the current response's shortcomings. For instance, if the scoring module identifies missing information, the preference generator incorporates these details into the updated preference, ensuring precision and relevance. This approach balances dynamic updates with long-term consistency, enabling stable and effective learning.

Dynamic DPO Training We utilize DPO training loss in the reward procedure, defined in the below:

$$L_{dynDPO} = -log\sigma(\beta log \frac{\pi_\theta(y_+^{(i)}|x^{(i)})}{\pi_{ref}(y_+^{(i)}|x^{(i)})} - \beta log \frac{\pi_\theta(y_-^{(i)}|x^{(i)})}{\pi_{ref}(y_-^{(i)}|x^{(i)})}) \quad (1)$$

here, we considee all preference pairs $(y_+^{(i)}, y_-^{(i)}), i \in N$ are dynamically generated. Specifically, preference responses are produced by expert preference generator while dis-preferred is the current response. Besides, π_{ref} is initialized as the same as π_θ, but keep frozen during the training. β controls the amount of divergence from π_{ref} and we use 0.5 for following experiments.

4 Experiments

To demonstrate the effectiveness of our method, we evaluated on two popular datasets:CNN-DailyMaill(CNNDM) v3.0.0 (Nallapati et al., 2016b) and XSUM (Narayan et al., 2018). The detailed descriptions of CNNDM and XSUM can be found in A.1. In later experiments, we firstly conduct conduct ablation studies to analyze the contribution of each key component and then compare our method against state-of-arts (SOTA) using ROUGE scores.

4.1 Experiment setup

Training In our experiments, we use Qwen-2.5 (Team, 2024) as both the LLM-based Judger and Preference Generator, guided by distinct prompts (examples in Figure 2). For CNNDM and XSum, we adopt BART-large fine-tuned on these datasets as backbones. As per Equation 1, π_{ref} is initialized to match the backbone but remains frozen during training, while π_θ is updated. We set $\beta = 0.5$ across all experiments and train models using 4 NVIDIA RTX A6000 GPUs. **Inference** During the inference, automatic scoring and feedback mechanism are all eliminated so that only target model(lightweight model) will be adopted. Then, followed by previous works, we use the ROUGE-

F1 which measures the overlap of n-grams between generated summaries and the reference summary.

4.2 Ablation Study

In this section, we conduct various experiments on CNNDM dataset to validate the effectiveness of each components in our framework.

DPO Training Strategy We compare our dynamic DPO training with traditional DPO. For traditional DPO, we pre-collected preference pairs using Qwen2.5-72B and trained the model with Equation 1. In contrast, our dynamic DPO generates preference pairs on-the-fly: the preferred response is created by an expert preference generator, while the dis-preferred response is the current model output. As shown in Table 1, our method outperforms both traditional DPO and baselines across all metrics, achieving significant ROUGE score improvements, demonstrating its effectiveness for lightweight summarization models.

	CNN/DailyMail		
Method	R-1	R-2	R-L
Bart$_{large}$(baseline)(Lewis et al., 2019)	44.0	21.1	40.6
Original DPO	**45.4**	**24.4**	40.0
Ours-dynDPO	**48.1**	**25.8**	**45.5**

Table 1: **Comparison on DPO Training Method.** Note that, preference dataset are prepared by using Qwen2.5-72B forehand for original DPO training.

Backbone Variations To show the flexibility and efficiency of our method for lightweight model training, we utilize various backbones (e.g., BART [1] and T5 [2]), which fine-tuned on CNNDM. Besides, for BART-base, we fine-tuned the original checkpoint ourselves official checkpoint on CNNDM is not provided. As shown in Table 2, our method consistently improves performance across backbones, enabling smaller models (e.g., T5-small) to surpass larger counterparts (e.g., T5-base) in key metrics. This highlights its scalability and effectiveness for edge device applications.

4.3 Compare Against SOTA

Table 3 presents an in-depth analysis, illustrating that our approach outperforms various baseline methods across multiple datasets, underscoring its robustness and effectiveness. The core of DPO lies in aligning the model with human preferences, making the quality of preference responses pivotal. To thoroughly evaluate our method, we employed two LLM scales: Qwen2.5-7B and Qwen2.5-72B (the largest feasible due to computational limits). Identical prompt templates were used for both models to minimize variability and ensure fair comparisons. Results indicate that with Qwen2.5-7B, our method achieves second-best performance, while switching to Qwen2.5-72B yields state-of-the-art (SOTA) results. Notably, the result emphasis the contributions of our approach that integration of real-time feedback enhances the model's adaptive alignment with preferences, enabling timely adjustments and superior performance, further demonstrating the scalability of our approach.

[1] https://huggingface.co/facebook/bart-large-cnn
[2] https://huggingface.co/google-t5

	CNN/DailyMail			
Backbone	Model	R-1	R-2	R-L
Bart$_{Large}$(406M)(Lewis et al., 2019)	Baseline	44.0	21.1	40.6
	Ours-dynDPO	**46.3**	**27.1**	**41.4**
Bart$_{base}$(139M)(Lewis et al., 2019)	Baseline*	40.2	18.2	32.7
	Ours-dynDPO	**43.9**	**24.9**	**38.2**
T5$_{base}$(220M)(Raffel et al., 2020)	Baseline	42.0	20.3	39.4
	Ours-dynDPO	**44.3**	**25.1**	**41.3**
T5$_{small}$(60M)(Raffel et al., 2020)	Baseline	41.2	19.6	38.1
	Ours-dynDPO	**43.5**	**22.5**	**40.0**

Table 2: **Comparison on Various Backbone.** Note, '*' indicates that we fine-tuned the model on CNNDM by our own to obtain the results, then utilized the checkpoint as the backbone for our framework.

	CNN/DailyMail			XSUM		
	R-1	R-2	R-L	R-1	R-2	R-L
T5$_{large}$ (Raffel et al., 2020)	42.4	20.8	39.9	40.1	17.2	32.3
BART$_{large}$ (Lewis et al., 2019)	44.0	21.1	40.6	45.4	22.3	37.3
PEGASUS (Zhang et al., 2020)	44.2	21.6	41.3	46.7	24.4	38.9
GSum (Dou et al., 2021)	45.5	22.3	42.1	45.1	21.5	36.6
SimCLS (Liu and Liu, 2021)	45.6	21.9	41.0	46.6	24.2	39.1
SeqCo (Xu et al., 2022)	45.0	21.8	41.8	45.6	22.4	37.0
TriSum-J (Jiang et al., 2024)	45.9	22.8	42.3	47.4	24.8	39.4
GECSum (Xie et al., 2024)	48.4	24.4	45.1	48.9	25.9	41.5
Ours-dynDPO(w.Qwen2.5-7B)	48.1	**25.8**	**45.5**	47.8	25.4	39.0
Ours-dynDPO(w.Qwen2.5-72B)	**51.0**	**27.5**	**47.5**	**50.0**	**26.2**	**42.2**

Table 3: **Evaluations on CNNDM and XSUM dataset.** 'w.Qwen2.5-7B' indicates our LLM-judger and preference generator are utilized 7B LLM during the training while 'w.Qwen2.5-72B' means we adopt 72B LLM. Besides, all results are obtained by utilizing $BART_{large}$ as backbone which is bold in gray.

5 Conclusion

We introduce a novel framework integrating an automatic scoring module and online feedback mechanism to enhance lightweight models for edge devices. By dynamically updating preferences, our method ensures training flexibility and scalability while eliminating the need for extensive annotations. Experiments demonstrate its superiority over existing methods, showcasing its effectiveness. Future work could explore generating stylish summaries aligned with personal preferences and addressing ethical constraints in this process.

Limitations

One of the limitations of proposed method is that it requires two LLMs, LLM-judger and preference generator, in the framework at the same time, so it would limit the usage for constrained computation resources. On the other hand, dynamic DPO training causes more training hours since both LLM-judger and preference generator need to generate related responses, therefore, additional prompts are needed to instruct the LLM to preciesely generate responses within reasonable length.

References

Mariane Balde, Alice Coucke, and Yannick Estève. 2024. Medvoc: Controlled vocabulary generation for biomedical summarization. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL)*.

Jaepill Choi, Kyubyung Chae, Jiwoo Song, Yohan Jo, and Taesup Kim. 2024. Model-based preference optimization in abstractive summarization without human feedback. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 18837–18851, Miami, Florida, USA. Association for Computational Linguistics.

Zi-Yi Dou, Pengfei Liu, Hiroaki Hayashi, Zhengbao Jiang, and Graham Neubig. 2021. GSum: A general framework for guided neural abstractive summarization. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4830–4842, Online. Association for Computational Linguistics.

Tao Ge, Si-Qing Chen, and Furu Wei. 2022. EdgeFormer: A parameter-efficient transformer for on-device seq2seq generation. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 10786–10798, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Pengcheng Jiang, Cao Xiao, Zifeng Wang, Parminder Bhatia, Jimeng Sun, and Jiawei Han. 2024. TriSum: Learning summarization ability from large language models with structured rationale. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 2805–2819, Mexico City, Mexico. Association for Computational Linguistics.

Jaehun Jung, Peter West, Liwei Jiang, Faeze Brahman, Ximing Lu, Jillian Fisher, Taylor Sorensen, and Yejin Choi. 2024. Impossible distillation for paraphrasing and summarization: How to make high-quality lemonade out of small, low-quality model. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 4439–4454, Mexico City, Mexico. Association for Computational Linguistics.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2019. BART: denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *CoRR*, abs/1910.13461.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. Bart: Denoising sequence-to-sequence pre-training for natural language generation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)*.

Yixin Liu and Pengfei Liu. 2021. SimCLS: A simple framework for contrastive learning of abstractive summarization. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 1065–1072, Online. Association for Computational Linguistics.

Ramesh Nallapati, Bowen Zhou, Cicero dos Santos, Çağlar Gülçehre, and Bing Xiang. 2016a. Abstractive text summarization using sequence-to-sequence rnns and beyond. In *Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning (CoNLL)*.

Ramesh Nallapati, Bowen Zhou, Cicero dos Santos, Çağlar Gulçehre, and Bing Xiang. 2016b. Abstractive text summarization using sequence-to-sequence RNNs and beyond. In *Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning*, pages 280–290, Berlin, Germany. Association for Computational Linguistics.

Shashi Narayan, Shay B. Cohen, and Mirella Lapata. 2018. Don‘t give me the details, just the summary! topic-aware convolutional neural networks for extreme summarization. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 1797–1807, Brussels, Belgium. Association for Computational Linguistics.

OpenAI, Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, Red Avila, Igor Babuschkin, Suchir Balaji, Valerie Balcom, Paul Baltescu, Haiming Bao, Mohammad Bavarian, Jeff Belgum, and 262 others. 2024. Gpt-4 technical report. *Preprint*, arXiv:2303.08774.

Romain Paulus, Caiming Xiong, and Richard Socher. 2017. A deep reinforced model for abstractive summarization. *CoRR*, abs/1705.04304.

Minh-Quang Pham, Sathish Indurthi, Shamil Chollampatt, and Marco Turchi. 2023. Select, prompt, filter: Distilling large language models for summarizing conversations. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 12257–12265, Singapore. Association for Computational Linguistics.

Rohan Rafailov, Eric Mitchell, Alex Tamkin, Stefano Ermon, and Chelsea Finn. 2023. Direct preference optimization: Your language model is secretly a reward model. In *Proceedings of the 11th International Conference on Learning Representations (ICLR)*.

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, 21(140):1–67.

Paul Roit and Roi Reichart. 2023. Factuality-enhanced reinforcement learning for summarization. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL)*.

Abigail See, Peter J. Liu, and Christopher D. Manning. 2017. Get to the point: Summarization with pointer-generator networks. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL)*.

Nisan Stiennon, Long Ouyang, Jeff Wu, Daniel M. Ziegler, Ryan Lowe, Chelsea Voss, Alec Radford, Dario Amodei, and Paul F. Christiano. 2020. Learning to summarize from human feedback. *CoRR*, abs/2009.01325.

Fuwen Tan, Royson Lee, Łukasz Dudziak, Shell Xu Hu, Sourav Bhattacharya, Timothy Hospedales, Georgios Tzimiropoulos, and Brais Martinez. 2024. Mobilequant: Mobile-friendly quantization for on-device language models. *Preprint*, arXiv:2408.13933.

Qwen Team. 2024. Qwen2.5: A party of foundation models.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. 2023. Llama: Open and efficient foundation language models. *Preprint*, arXiv:2302.13971.

Jiawen Xie, Shaoting Zhang, and Xiaofan Zhang. 2024. GECSum: Generative evaluation-driven sequence level contrastive learning for abstractive summarization. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 7581–7595, Torino, Italia. ELRA and ICCL.

Shusheng Xu, Xingxing Zhang, Yi Wu, and Furu Wei. 2022. Sequence level contrastive learning for text summarization. *Preprint*, arXiv:2109.03481.

Yichong Xu, Ruochen Xu, Dan Iter, Yang Liu, Shuohang Wang, Chenguang Zhu, and Michael Zeng. 2023. InheritSumm: A general, versatile and compact summarizer by distilling from GPT. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 13879–13892, Singapore. Association for Computational Linguistics.

Shafiq Joty Zaman, Zhiwei Jin, and Nazneen Fatema Rajani. 2024. Satsum: Scalable abstractive summarization for scientific literature. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL)*.

Jingqing Zhang, Yao Zhao, Mohammad Saleh, and Peter J. Liu. 2020. Pegasus: Pre-training with extracted gap-sentences for abstractive summarization. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*.

Rui Zhang, Zhijing Jin, and Danqi Chen. 2024. Benchmarking large language models for summarization tasks. arXiv preprint arXiv:2402.12345.

Yixuan Zhang, Esin Durmus, and Claire Cardie. 2023. Summit: Iterative summarization with chatgpt. arXiv preprint arXiv:2305.14427.

A Appendix

A.1 Dataset Descriptions

CNNDM a large-scale benchmark for abstractive text summarization, featuring news articles paired with multi-sentence summaries. It contains approximately 300,000 articles with an average article length of 781 words and summary length of 56 words. The dataset is split into 287,227 training, 13,368 validation, and 11,490 test samples, following standard evaluation protocols for summarization tasks.

XSUM dataset is specifically designed for single-sentence summarization, where each summary concisely captures the core point of the source article. It comprises 226,711 BBC news articles, divided into 204,045 training, 11,332 validation, and 11,334 test samples. On average, articles contain 431 words, while summaries are 23 words long, emphasizing the dataset's focus on high compression and precision.