

Headline Token-based Discriminative Learning for Subheading Generation in News Article

Joonwon Jang

Sejong University / Republic of Korea
joonwon.lainshower@gmail.com

Misuk Kim*

Sejong University / Republic of Korea
mskim.sju@gmail.com

Abstract

The news subheading summarizes an article’s contents in several sentences to support the headline limited to solely conveying the main contents. So, it is necessary to generate compelling news subheadings in consideration of the structural characteristics of the news. In this paper, we propose a subheading generation model using topical headline information. We introduce a discriminative learning method that utilizes the prediction result of masked headline tokens. Experiments show that the proposed model is effective and outperforms the comparative models on three news datasets written in two languages. We also show that our model performs robustly on a small dataset and various masking ratios. Qualitative analysis and human evaluations also show that the overall quality of generated subheadings improved over the comparative models.

1 Introduction

The news headline summarizes the article to grab the attention and interest of the readers (Dor, 2003; Ifantidou, 2009; Ecker et al., 2014). However, the headline is written in a brief form of short sentences with topic-related phrases (Yamada et al., 2021), making it hard for users to grasp the entire content of the news article from the headline alone. To tackle this problem, some news vendors provide a *subheading*, usually located right below the headline, to convey its main content within several sentences. This component can provide a core and informative summary of a news article that cannot be conveyed by the headline alone. Mainly, subheadings are written by professional news writers with concise content that corresponds to the main body of the news.

Recently, Hasan et al. (2021) released XLSum, a multilingual news summary dataset, referring to subheading as a summary of the article. Therefore, generating subheadings can be considered an abstractive summarization problem that needs to cap-

ture the topical knowledge from the body of the article. The main approach is to add auxiliary signals to make the model aware of topical knowledge. Dou et al. (2021) and Aralikkatte et al. (2021) add an external guidance signal by lexical similarity between input text and summary. Yamada et al. (2021) extracts the context word sequences from the reference to reflect some important phrases from the article. Although these external auxiliary sources provide diverse topical signals, they are cost-intensive to heuristically manipulate and have limitations in guiding the overall topical information of the article. Other approaches incorporate contrastive learning into sequence-to-sequence (seq2seq) model, allowing the model to learn topical representation of the input text (Lee et al., 2021; Liu et al., 2021; Wu et al., 2020). They explicitly constructs positive or negative inputs to introduce contrastive loss as an augmentation of MLE training.

In this work, we propose a novel framework for generating compelling news subheadings by discriminating whether each token in the reconstructed headline is the same as the token in the original headline. Unlike previous approaches that use heuristically extracted topical information or positive and negative pairs, we utilize *headline* that fundamentally implies the topic of the entire article. We make full use of this indispensable object as a guide signal through token-based discriminative learning. We conducted comparative experiments on three datasets written in English or Korean to evaluate the performance of our model and verified our model through additional qualitative analysis with human evaluations.

2 Datasets

We used one English news summarization dataset and two Korean news summarization datasets.

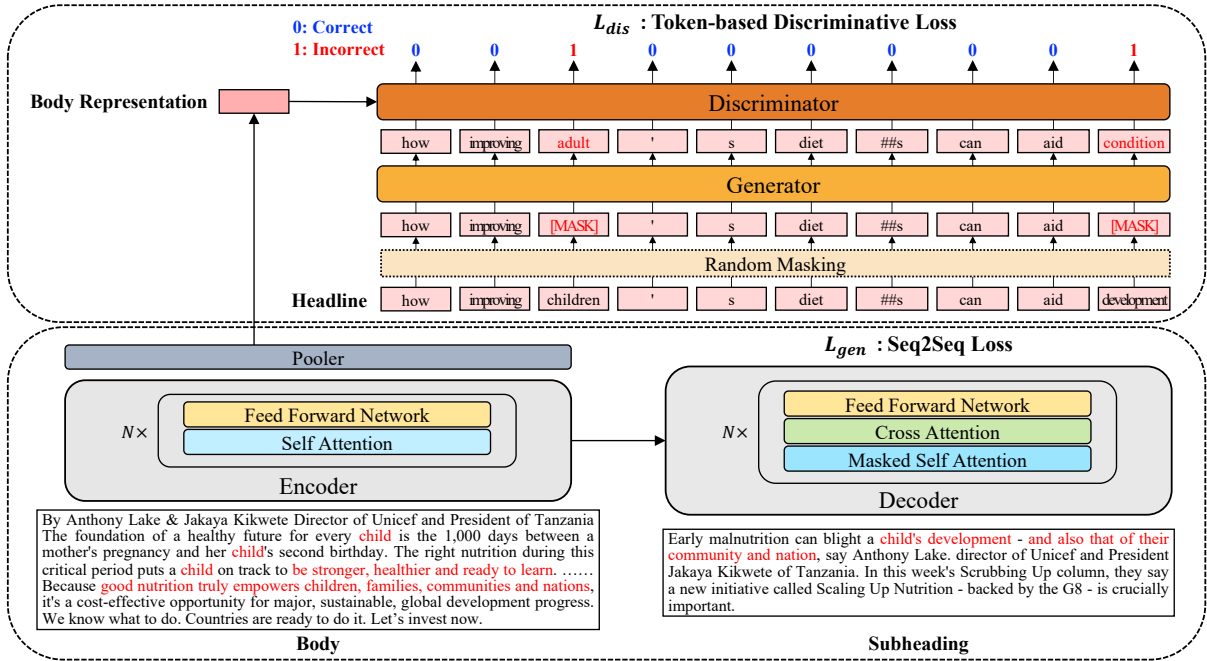


Figure 1: Framework of the proposed model. The lower part of the figure represents subheading generation, and the upper part of the figure represents token-based discriminative learning. Prediction results of discriminator encourage encoder to focus on the topical information.

2.1 XLSum

XLSum (Hasan et al., 2021) is a highly abstract multilingual news summarization dataset containing online articles crawled from the British Broadcasting Corporation (BBC). They regard a bold paragraph containing one or two sentences at the beginning of each article as a summary, a subheading. We use the English and Korean versions with train set, valid set, and test set pairs of $\{306522, 11535, 11535\}$ and $\{4407, 550, 550\}$, respectively.

2.2 YonhapNews

However, the size of the training dataset in XLSum’s Korean version is insufficient for fine-tuning. We construct a dataset from YonhapNews, one of the most reliable news outlets in South Korea, to evaluate the performance of the model with a sizable Korean dataset. In YonhapNews, the subheading is located right below the headline to condense the body in abstractly, like the BBC. Train set, valid set, and test sets in YonhapNews are $\{208750, 26094, 26094\}$ and will be released for academic use.¹

¹<https://github.com/Lainshower/Subheading-Gen>

3 Proposed Method

As shown in Figure 1, our proposed model consists of subheading generation (bottom) and token-based discriminative learning (top) parts. The loss occurring in each part is defined as L_{gen} and L_{dis} , and the model is trained by minimizing the following loss:

$$L = L_{gen} + \lambda \cdot L_{dis}, \quad (1)$$

where λ is a weighted hyperparameter for the two losses.

3.1 Subheading Generation

We use BART (Lewis et al., 2020) as our seq2seq model, where the encoder takes a body B as an input, generates an input representation of the body, and passes it to the decoder, which outputs the subheading \hat{S} . The loss of subheading generation is as follows:

$$L_{gen} = - \sum_{i=1}^M \log(p(s_i | s_{1:i-1}, B; \theta)), \quad (2)$$

where a body $B = (b_1, b_2, \dots, b_N)$ and its subheading $S = (s_1, s_2, \dots, s_M)$ consists of N token vectors b_i and M token vectors s_j , respectively, and the seq2seq model is optimized to learn the θ parameters to minimize the negative log-likelihood.

Dataset	Model	Size	Rouge-1	Rouge-2	Rouge-L	BERTScore
XLSum-ENG	BART	139M	36.52	15.33	30.49	77.68
	MT5	582M	37.81	14.59	28.30	74.48
	T5	220M	37.87	15.97	29.09	75.96
	Pegasus	571M	39.57	16.63	32.45	77.93
	T5 w/multi	247M	35.37	14.71	29.84	75.74
	Ours	282M	39.84	18.07	33.77	78.96
XLSum-KOR	BART	124M	25.23	15.91	23.37	74.60
	MT5	582M	29.16	14.03	25.40	69.94
	T5	247M	13.99	4.10	13.48	67.92
	T5 w/multi	247M	17.55	7.03	16.60	71.58
	Ours	269M	27.41	17.56	25.48	75.58
YonhapNews	BART	124M	21.41	11.00	19.21	72.19
	MT5	582M	25.24	11.28	21.13	70.70
	T5	247M	20.10	8.70	17.96	71.67
	T5 w/multi	247M	22.21	9.91	19.91	71.21
	Ours	269M	24.15	13.21	21.93	73.08

Table 1: Subheading generation performance for each dataset. Size represents the number of parameters in each model. Our model outperforms the comparative models except for Rouge-1 score for XLSum-KOR and YonhapNews.

3.2 Token-based Discriminative Learning

Inspiring by [Chuang et al. \(2022\)](#), we inject topical knowledge of the article by discriminating whether tokens in the reconstructed headline are the same as the original. As described in Appendix A, headline has high lexical similarity with subheading compared to other objects in the article. Therefore, headline reconstruction helps model to aware of topical information which is inherent in the headline. The loss of token-based discriminative learning is as follows:

$$L_{dis} = \sum_{i=1}^L \left(-\mathbb{1}(h'_i = h_i) \log D(H'|B, H) - \mathbb{1}(h'_i \neq h_i) \log(1 - D(H'|B, H)) \right), \quad (3)$$

where a headline $H = (h_1, h_2, \dots, h_L)$ consist of L tokens and the reconstructed headline H' is $H' = G(H_{masked})$ where G is the generator and a masked headline H_{masked} is obtained with random mask $M = [m_1, m_2, \dots, m_L], m_t \in [0, 1], H_{masked} = H \cdot M$. Using the compressed body representation of the encoder, the discriminator D predicts whether the tokens in the reconstructed headline H' are the same as the original headline H . As shown in Figure 1, the generator predicts the masked headline into “*how improving adult’s diet can aid condition*”. Using the encoded body representation and partially mispredicted headline, the model trains the incorrectly predicted “*adult*”, and “*condition*” and the correctly predicted rest of the tokens, respectively. Back-propagated gradients of the discriminator D cause the encoder to include the topical information of the article in the body representation by classifying whether the tokens

in the reconstructed headline H' come from the original headline or not.

Model	Generated Subheading
BART	The UK’s oil and gas industry generated negative tax receipts in 2015-16, according to HM Revenue and Customs (HMRC).
MT5	Have led to a fall in tax receipts from UK oil and gas production, according to HM Revenue and Customs (HRMC)
T5	revenues have fallen to their lowest level since records began in the 1960s, according to new figures from HM Revenue and Customs (HMRC)
Pegasus	tax receipts from oil and gas production in the UK have fallen to their lowest level, according to HM Revenue and Customs (HMRC)
Ours	Revenues from the North Sea oil and gas industry have fallen to their lowest level since records began, according to HM Revenue and Customs (HMRC).

Table 2: Example of generated subheading for each model. The original headline is “*North Sea receipts hit record low*” and the reference subheading is “*The UK government has incurred a loss from North Sea oil and gas production for the first time since records began nearly 50 years ago*”. The body of the article can be found in the XLSum-ENG test set with the corresponding id= ‘uk-scotland-scotland-business-36388621’.

4 Experiments

4.1 Experimental Setting

We use pretrained models BART and ELECTRA ([Clark et al., 2020](#)). Unlike [CLS] representation of BERT ([Devlin et al., 2019](#)), BART doesn’t have a special input representation token. As such, we use an average pooler to compress the output of the encoder and freeze the generator to keep generating noise headline for token-based discriminative learning. Optimal parameters

were obtained in the search spaces with learning rate $\{1e-5, 2e-5, 3e-5, 4e-5\}$, masking ratio $\{0.1, 0.2, 0.3, 0.4, 0.5\}$, and lambda $\{0.1, 0.01\}$.

4.2 Comparative Models

BART, T5 (Raffel et al., 2020), and MT5, (Xue et al., 2021) were used as comparative models in all datasets. Also, Pegasus (Zhang et al., 2020) for English were used as comparative models. For a fair comparison, we use the concatenated body with a headline in the input of the comparative models.

4.3 Experimental Results

Table 1 shows the results of subheading generation performance for each dataset. Model performance was evaluated using Rouge (Lin, 2004) and BERTScore (Zhang et al., 2019). In XLSum-ENG, our model outperforms all comparative models. In particular, our model performs better than MT5 or Pegasus, which have more than double the model size. In other words, token-based discriminative learning can improve generation performance more efficiently than simple concatenation. Our model outperforms the comparative models in all other metrics except the Rouge-1 score for the Korean language datasets. MT5 records the highest Rouge-1 score on both Korean datasets. However, because Korean is decomposed into many sub-words due to its morphological richness, it is not suitable to evaluate performance with Rouge-1 score alone. In particular, in terms of BERTScore, our model scored 5.64% and 2.32% higher than MT5 in XLSum-KOR and YonhapNews, respectively. This indicates that our model can generate semantically relevant subheadings. Moreover, good performance on small datasets (i.e., XLSum-KOR) demonstrate the robustness of our model.

Table 2 shows an example of the generated subheadings for each model. We can see that our model utilizes “North Sea” and “record low” from the headline to better condense topical information in the article. Additional qualitative results are described in Appendix C.

	XLSum-ENG	XLSum-KOR	YonhapNews
BART	2.53 (0.64)	2.00 (0.38)	3.53 (0.92)
MT5	2.71 (0.61)	2.07 (0.59)	2.60 (0.51)
T5	2.64 (0.74)	2.20 (0.41)	2.73 (0.46)
Pegasus	3.64 (0.50)	-	-
Ours	4.21 (0.70)	3.67 (0.49)	3.40 (0.74)

Table 3: The average score of human evaluation for XLSum-ENG, XLSum-KOR, and YonhapNews. Numbers in parentheses indicate the standard deviations.

4.4 Human Evaluations

We conduct human evaluations to verify whether the subheadings of the proposed method are more topically relevant than the baselines. Three samples were randomly selected from each test dataset, and subheadings generated along with their corresponding headlines and body were shown to the workers and evaluated on a five-point Likert scale. Table 3 shows that our model generates topic-relevant subheadings better than the baselines on two datasets, and is particularly robust on a small dataset (i.e., XLSum-KOR). In the case of YonhapNews, BART showed the highest score, but the independent t-test showed that the average difference between Ours and BART was insignificant ($p>0.663$).

4.5 Comparison with Multi-task Learning

To verify whether our model effectively learns topical information from the headlines, we conduct additional experiments with a multi-task learning. Different prefixes were used to know the model what the current training task is. One task maps news body text to subheading, and the other maps news body text to headline. We experiment with T5 because it has less discrepancy with our pre-training objectives. T5 w/multi rows in Table 1 show the multi-task learning results, demonstrating that our method is more effective in learning headline information.

4.6 Ablation Studies

We perform ablation studies in terms of masking ratio to analyze the effectiveness of token-based discriminative learning. Figure 2 shows the results of the Rouge-2 score and BERTScore according to the masking ratio for each dataset. We also plot the performance of the T5 with similar parameter sizes to ours. Our model outperforms T5 in all masking ratio ranges. This indicates that our model is not significantly sensitive to masking ratio. In particular, the large performance difference of XLSum-KOR demonstrates the robustness of our model on the small dataset. The original headline is completely incorrectly reconstructed if the masking ratio exceeds 0.4, limiting ability of the model to learn crucial topical information from the headline considering the token length of the headline. However, for small masking ratios such as 0.1, the generator can completely reconstruct the original headline, but it is limited in maximizing the benefits of token-based discriminative learning. Headline to-

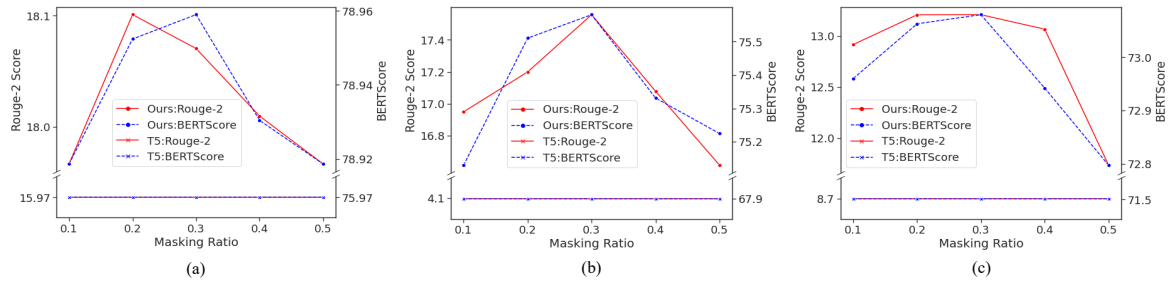


Figure 2: Rouge-2 score and BERTScore according to masking ratio for (a) XLSum-ENG, (b) XLSum-KOR, and (c) YonhapNews. The performance of the T5 in two Korean datasets is plotted with a ‘x’ marker. The masking ratio between 0.2 and 0.3 shows the best performance in all datasets.

ken length distribution is described in Appendix B. Therefore, it is recommended to set the masking ratio between 0.2 and 0.3 in order to utilize the token-based discriminative learning for subheading generation entirely.

5 Conclusions

In this paper, we propose a novel model for generating a subheading for news article. Along with token-based discriminative learning, our model can effectively utilize topical information from a headline that is essential in articles and does not require additional manipulated information. Experiments on three datasets written in two different languages show the effectiveness of the proposed model. Also, qualitative results and human evaluation show that the overall quality of generated subheadings is improved compared to comparative models. We expect that our model will be extended in future research to an abstractive summarization task that include both a headline and a body text, such as legal texts or papers.

Limitations

Our study outperformed all comparative models in generating subheadings through token-based discriminative learning. However, the experiments mainly used limited languages such as English and Korean due to a lack of large-scale multilingual training data and the need for significant GPU resources. We, therefore, encourage further investigations to expand the versatility of the proposed model by utilizing large-scale multilingual language datasets to verify expandable applications in various morphological characteristics.

Ethics Statement

As YonhapNews is one of the most reliable media outlets in South Korea, articles from YonhapNews are published through a rigorous verification process and will be deleted or revised if they contain any form of bias. However, the period of data collected is three years, and there may be past article content that has not been modified by new facts, so we cannot guarantee that all articles in YonhapNews dataset are completely unbiased. Nevertheless, this dataset has sufficient potential to develop into various studies and thus is released for academic uses.

Acknowledgement

This work was supported by Institute of Information & Communications Technology Planning & Evaluation(IITP) (Project Number: 2021-0-00469) and the Technology Innovation Program (or Industrial Strategic Technology Development Program-Knowledge service industry technology development project - service core technology development) funded by the Ministry of Trade, Industry & Energy(MOTIE, Korea) (Project Number: 20018758).

References

- Rahul Aralikkatte, Shashi Narayan, Joshua Maynez, Sascha Rothe, and Ryan McDonald. 2021. [Focus attention: Promoting faithfulness and diversity in 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 1: Long Papers)*, pages 6078–6095, Online. Association for Computational Linguistics.
- Yung-Sung Chuang, Rumen Dangovski, Hongyin Luo, Yang Zhang, Shiyu Chang, Marin Soljagic, Shang-Wen Li, Scott Yih, Yoon Kim, and James Glass. 2022. [DiffCSE: Difference-based contrastive learning for](#)

- sentence embeddings. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4207–4218, Seattle, United States. Association for Computational Linguistics.
- Kevin Clark, Minh-Thang Luong, Quoc V Le, and Christopher D Manning. 2020. Electra: Pre-training text encoders as discriminators rather than generators. *arXiv preprint arXiv:2003.10555*.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.
- Daniel Dor. 2003. On newspaper headlines as relevance optimizers. *Journal of pragmatics*, 35(5):695–721.
- 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.
- Ullrich KH Ecker, Stephan Lewandowsky, Ee Pin Chang, and Rekha Pillai. 2014. The effects of subtle misinformation in news headlines. *Journal of experimental psychology: applied*, 20(4):323.
- Tahmid Hasan, Abhik Bhattacharjee, Md. Saiful Islam, Kazi Mubasshir, Yuan-Fang Li, Yong-Bin Kang, M. Sohel Rahman, and Rifat Shahriyar. 2021. XLsum: Large-scale multilingual abstractive summarization for 44 languages. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 4693–4703, Online. Association for Computational Linguistics.
- Elly Ifantidou. 2009. Newspaper headlines and relevance: Ad hoc concepts in ad hoc contexts. *Journal of Pragmatics*, 41(4):699–720.
- Seanie Lee, Dong Bok Lee, and Sung Ju Hwang. 2021. Contrastive learning with adversarial perturbations for conditional text generation. In *International Conference on Learning Representations*.
- 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, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7871–7880, Online. Association for Computational Linguistics.
- Chin-Yew Lin. 2004. Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, pages 74–81.
- Junpeng Liu, Yanyan Zou, Hainan Zhang, Hongshen Chen, Zhuoye Ding, Caixia Yuan, and Xiaojie Wang. 2021. Topic-aware contrastive learning for abstractive dialogue summarization. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 1229–1243, Punta Cana, Dominican Republic. Association for Computational Linguistics.
- 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.
- Hanlu Wu, Tengfei Ma, Lingfei Wu, Tariro Manyumwa, and Shouling Ji. 2020. Unsupervised reference-free summary quality evaluation via contrastive learning. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 3612–3621, Online. Association for Computational Linguistics.
- Linting Xue, Noah Constant, Adam Roberts, Mihir Kale, Rami Al-Rfou, Aditya Siddhant, Aditya Barua, and Colin Raffel. 2021. mT5: A massively multilingual pre-trained text-to-text transformer. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 483–498, Online. Association for Computational Linguistics.
- Kosuke Yamada, Yuta Hitomi, Hideaki Tamori, Ryohhei Sasano, Naoaki Okazaki, Kentaro Inui, and Koichi Takeda. 2021. Transformer-based lexically constrained headline generation. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 4085–4090, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.
- Jingqing Zhang, Yao Zhao, Mohammad Saleh, and Peter Liu. 2020. Pegasus: Pre-training with extracted gap-sentences for abstractive summarization. In *International Conference on Machine Learning*, pages 11328–11339. PMLR.
- Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q Weinberger, and Yoav Artzi. 2019. Bertscore: Evaluating text generation with bert. In *International Conference on Learning Representations*.

A Similarity between Subheading and Headline

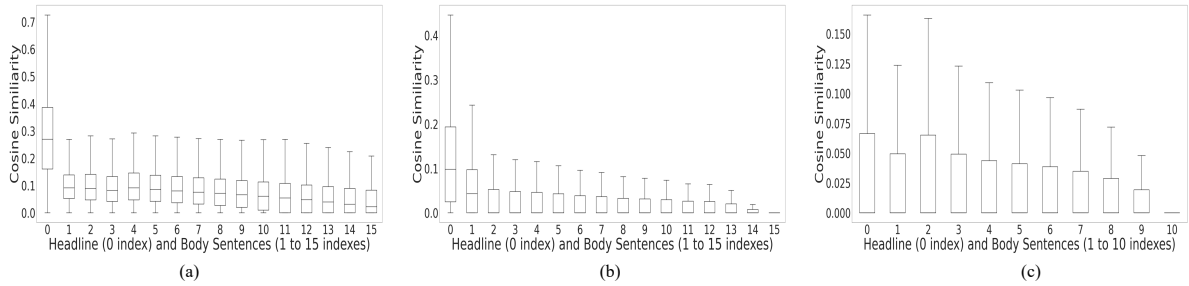


Figure A.1: Cosine similarity between the TF-IDF values of subheading-headline (0 index in X-axis) and between the TF-IDF values of the subheading-body sentences (1 to N indexes in X-axis) of the news article. Each figure (a), (b), and (c) represents for dataset XLSum-ENG, XLSum-KOR, and YonhapNews, respectively.

B Length of Headline

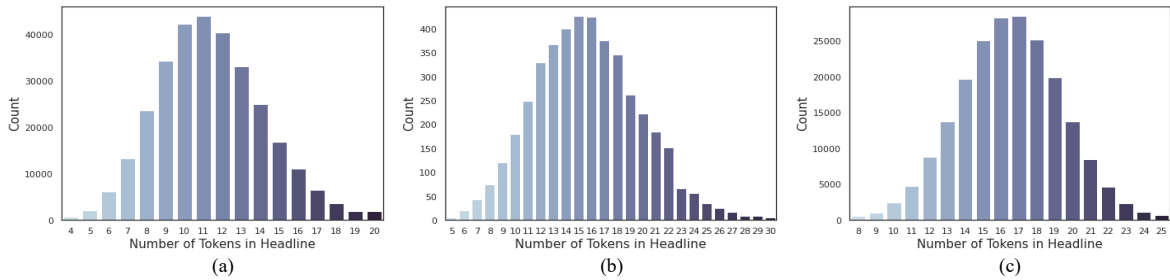


Figure B.1: Distribution of the token length in the headline of (a) XLSum-ENG, (b) XLSum-KOR, and (c) YonhapNews.

C Qualitative Results

Model	Generated Subheading
BART	신종 코로나바이러스 감염증(코로나19) 백신 구매 계약을 맺은 영국 정부가 존슨 총리에게 코백스를 통해 백신을 기부할 것을 촉구했다. (The British government, which signed a contract to purchase a new coronavirus infection (COVID-19) vaccine, urged Prime Minister Johnson to donate the vaccine through COVAX.)
MT5	단체들이 신종 코로나바이러스 감염증(코로나19) 백신을 구매한 후에도 잉여 물량의 대부분을 공유할 것이라고 촉구했다. (Organizations urged that they will share most of the surplus even after purchasing a new coronavirus infection (COVID-19) vaccine.)
T5	신종 코로나바이러스 감염증(코로나19) 백신을 사용할 수 있다고 밝혔다. (It was announced that a new coronavirus infection (COVID-19) vaccine can be used.)
Ours	영국이 신종 코로나바이러스 감염증(코로나19) 백신의 잉여 물량을 저개발국에 기부하겠다고 밝혔다. (The <u>UK</u> has announced that it will <u>donate remaining of COVID-19 vaccine</u> to underdeveloped countries.)

Table C.1: Example of generated subheading for each model in XLSum-KOR. The original headline is “코로나19 백신: 전국민 접종하고도 1억 도즈 남는 영국... ‘남는 백신 기부할 것’ (COVID-19 Vaccine: 100 million doses of UK left after national vaccination... ‘Donate the remaining vaccines.’)” and the reference subheading is “국제 구호 단체들이 보리스 존슨 영국 총리에게 영국이 저개발국가에 기부할 수 있는 코로나19 백신이 얼마나 되는지를 조속히 밝힐 것을 촉구하고 있다. (International aid organizations are urging British Prime Minister Boris Johnson to reveal as soon as possible how many COVID-19 vaccines Britain can donate to underdeveloped countries.)” The body of the article can be found in the test set with corresponding id=‘international-56553770’.

Model	Generated Subheading
BART	11억원 기부...군부대도 방역에 힘 보탤다 (Donating 11 billion won ... The military unit also helped with quarantine)
MT5	잇단 후원...경북도교육청, 장병 130여명 투입해 방역 지원 (Continuous sponsorship ... Gyeongbuk Provincial Office of Education dispatched 130 soldiers to provide quarantine support)
T5	각계서 정금·성금 잇따라 (a series of donations from all walks of life)
Ours	착한 임대인 운동 확산...제201특공여단, 경산역·버스터미널 방역 지원 (Spreading the Good Renters Movement ... The 201st Special Forces Brigade, provide quarantine support to <u>Gyeongsan</u> Station and Bus Terminal)

Table C.2: Example of generated subheading for each model in YonhapNews. The original headline is “‘코로나19 극복 함께해요’...대구·경북에 성금·물품 답지 (‘Let’s overcome COVID-19 together’...Daegu and Gyeongsangbuk-do collect donations to deliver goods)” and the reference subheading is “착한 임대인 운동 공공기관에 확산...군부대는 방역 지원 (The Good Renters Movement Spreads to Public Institutions... Military units provide quarantine support)”. This news article covers the quarantine support from various industry fields, including the rental industry and the military service, in response to the COVID-19 in Daegu and Gyeongsangbuk-do, city and state located in Korea. Subheading generated by our model include “Gyeongsan”, located in Gyeongsang-do, showing that it reflects the locational information of quarantine support that occurs through the headline. The body of the article can be found in ‘<https://www.yna.co.kr/view/AKR2020030914600053>’