

RTSUM: Relation Triple-based Interpretable Summarization with Multi-level Saliency Visualization

Seonglae Cho, Myungha Jang, Jinyoung Yeo, Dongha Lee*

Yonsei University, Republic of Korea

{sungle3737, jinyeo, donalee}@yonsei.ac.kr, myunghajang@gmail.com

Abstract

In this paper, we present RTSUM, an unsupervised summarization framework that utilizes relation triples as the basic unit for summarization. Given an input document, RTSUM first selects salient relation triples via multi-level saliency scoring and then generates a concise summary from the selected relation triples by using a text-to-text language model. On the basis of RTSUM, we also develop a web demo for an interpretable summarizing tool, providing fine-grained interpretations with the output summary. With support for customization options, our tool visualizes the saliency for textual units at three distinct levels: sentences, relation triples, and phrases. The code¹ and video² are publicly available.

1 Introduction

Text summarization has emerged as a critical tool in the era of information overload, enabling users to quickly understand the essence of long text. Among various summarization techniques, abstractive summarization has gained significant attention due to its ability to generate fluent and concise summaries that capture the main ideas of the source text (Nallapati et al., 2016; See et al., 2017; Tan et al., 2017; Cohan et al., 2018; Xu et al., 2020a; Koh et al., 2022). Nevertheless, despite their advantages in flexibility and reduced redundancy compared to extractive methods, abstractive methods inherently lack interpretability. That is, the absence of a direct link to the source text can make it difficult to trace back the source of information, which makes the summary lack interpretability.

Interpretability in summarization is important to provide users a way to cross-check that the generated summary is factually consistent, and to provide more context to dive into if one wants to know

more about the summarized content. To generate an interpretable summary, extractive summarization techniques can offer advantages. As they directly extract sentences from the text, the sentences themselves serve as the source of information (Xu et al., 2020b; Padmakumar and He, 2021). However, a significant drawback of many extractive methods lies in their sentence-level operation, which limits their ability to extract fine-grained key information (Zheng and Lapata, 2019; Liu et al., 2021). In many cases, a single sentence describes multiple diverse pieces of information that should be treated as distinct facts for summarization. By selecting entire sentences, these methods may include unnecessary or redundant information in the summary, reducing both its efficiency and readability.

To enhance the interpretability of the summarization process by incorporating fine-grained key information, our focus lies on leveraging *relation triples* as the basic unit for summarization. A relation triple in the form of (subject, predicate, object) concisely describes a single piece of information corresponding to its relation (i.e., predicate) between two entities (i.e., subject and object), and it can be effectively identified from a source document by using open information extraction (OpenIE) systems (Angeli et al., 2015; Mausam, 2016).

Using relation triples, our main idea embodies *selection-and-sentencification*, which achieves a combination of extractive and abstractive summarization methods. Specifically, we first select only a few relation triples according to their importance – *saliency* – within the document for summarization, and then reassemble the selected relation triples into the final output summary. This two-step approach enhances the interpretability of the summarization by providing clear explanations for the saliency scores of relation triples and their contributions to the final summary. This clarity allows users to understand the crucial elements driving the summarization process effectively.

*Corresponding author

¹<https://github.com/seonglae/RTSum>

²<https://youtu.be/sFR00xfqvVM>

Formally, we present an unsupervised Relation Triple-based Summarization framework, named **RTSUM**. For relation triple selection, **RTSUM** identifies heterogeneous textual information units with various granularity, which are (1) sentences, (2) relation triples, and (3) phrases, to utilize their own salience all together. Under the principle that more salient textual units are much more relevant to other units semantically and lexically, it models the multi-level salience from the three distinct textual units. Then, it selects the K most salient relation triples based on the multi-level salience scores. For relation triple sentencification, **RTSUM** employs a neural text-to-text architecture as a *relation combiner* to transform the relation triples into the summary sentences. The relation combiner is effectively optimized in a self-supervised manner by using source sentences (sampled from training documents) and their relation triples (extracted from the sampled sentences) as targets and inputs, respectively, while not requiring any reference summaries of its training documents.

Building upon the **RTSUM** framework, we develop an online demo to showcase an interpretable text summarization tool. Given an input document, our tool generates a concise summary, while simultaneously offering fine-grained interpretations by visually depicting the multi-level salience of textual units within the source document. For clarity in visualization, the tool highlights text spans (i.e., textual units) based on their salience score, with numerical ranks provided as annotations. Furthermore, our tool offers customization options, allowing users to personalize the visualization according to their preferences and specific purposes.

Our multi-level salience visualization empowers users to easily identify the textual units that mostly influence the final summary; it also provides valuable insights into the salient semantic structure of the document at a glance, enhancing users’ overall understanding of the summarization process.

2 Preliminary

Textual Information Units. We utilize three different types of textual information units with various granularity: sentences, relation triples,³ and phrases. All sentences and relations are extracted from a source document by using open information extraction (OpenIE) systems. Among several

³For brevity, we use the terms “relation triples” and “relations” interchangeably, in the rest of this paper.

implementations, we employ OpenIE 5⁴ (Mausam, 2016) released by UW and IIT Delhi. Similarly, all noun and verb phrases in the document are identified based on POS labels tagged by the Spacy library. Table 1 shows an example of the three textual information units in a single sentence.

Sentence	Hugh Laurie joins the cast and Julia Louis-Dreyfus is now the president of the United States on HBO’s hit comedy.
Relation	(S: Hugh Laurie, P: joins, O: the cast) (S: Julia Louis-Dreyfus, P: is, O: now the president of the United States on HBO’s hit comedy)
Phrase	Hugh Laurie, joins, cast, Julia Louis-Dreyfus, is, president, United States, HBO’s, hit comedy

Table 1: An example of textual information units.

Relation Triples. Each relation triple, denoted by $r = (\text{sub}, \text{pred}, \text{obj})$, represents a relation (i.e., predicate) between two text spans (i.e., subject and object), and it corresponds to a single piece of information in terms of the relation. The three components are described in natural language, and this allows us to treat them as the sequence of tokens in the vocabulary, similar to sentences and phrases. Thus, we consider the concatenated text of its subject, predicate, and object as the textual description of a relation triple, i.e., $\text{desc}(r) = [\text{sub}||\text{pred}||\text{obj}]$.

Problem Definition. Given a source document D and its information units, including sentences $\mathcal{S} = \{s_1, \dots, s_{N_s}\}$, relation triples $\mathcal{R} = \{r_1, \dots, r_{N_r}\}$, and phrases $\mathcal{P} = \{p_1, \dots, p_{N_p}\}$, the goal of our relation triple-based summarization task is (1) to select the relation triples based on their salience within the document, and (2) to generate a concise summary from the selected salient relation triples. In this paper, we mainly focus on the unsupervised setting where the annotated text-summary pairs are not available for training a summarization model, since such reference summaries are usually noisy, expensive to acquire, and hard to scale.

3 **RTSUM: Relation Triple-based Summarization Framework**

Our summarization framework, named **RTSUM**, consists of the two steps: (1) *information selection* for identifying the salient relation triples based on multi-level salience from various textual information units (Section 3.1), and (2) *information sentencification* for combining the selected relation triples

⁴<https://github.com/dair-iitd/OpenIE-standalone>

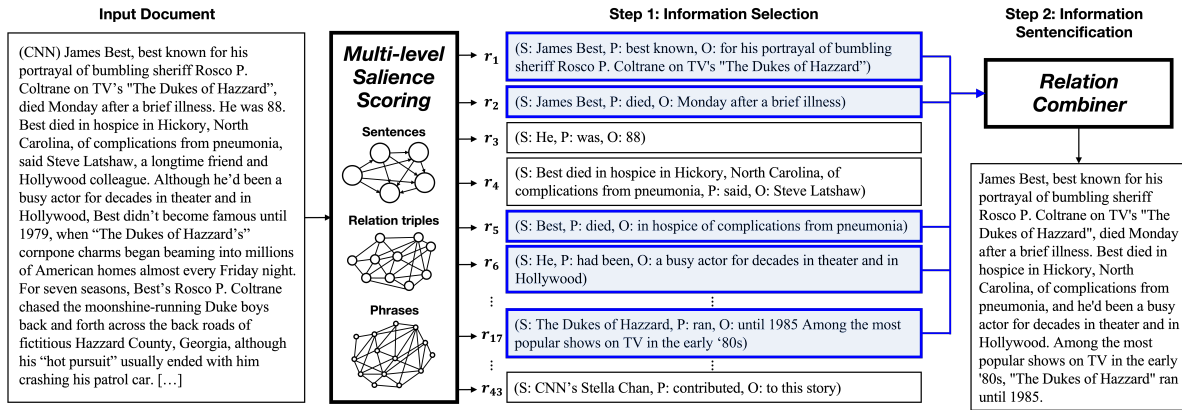


Figure 1: The overall process of the RTSUM framework. RTSUM selects salient relation triples and then generates the plausible sentences from the selected relation triples.

into plausible sentences with the help of a neural text generator (Section 3.2). Figure 1 illustrates the overall process of our RTSUM framework.

3.1 Information Selection

For the selection of relation triples that would be included in the summary, RTSUM models the multi-level saliency for each relation triple by leveraging heterogeneous textual information units. Specifically, it figures out how significant a relation triple is within the document from the perspective of (1) the sentence that the relation triple is extracted from, (2) the relation triple itself, and (3) the phrases that the relation triple contains. In the end, RTSUM selects the most K salient relation triples based on their multi-level saliency scores.⁵

3.1.1 Sentence-level Saliency Score

The sentence-level saliency considers the significance of the sentence that each relation triple is extracted from. Following previous studies (Zheng and Lapata, 2019; Liu et al., 2021), we infer the sentence-level saliency by utilizing sentence order (i.e., a preceding sentence is more likely to contain salient information) and semantic similarity (i.e., the sentence that is more semantically relevant to other sentences is likely to contain salient information). Thus, we construct a sentence-level text graph $\mathcal{G}^s = (\mathcal{S}, E^s)$ with a directed edge from a former sentence node s_i to a latter sentence node s_j , and the edge weight E_{ij}^s is the semantic similarity between the two sentences.

$$E_{ij}^s = \begin{cases} \text{sim}(s_i, s_j) & \text{if } s_i \text{ precedes } s_j \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

⁵The selection (or ranking) strategy based on multi-level saliency can be implemented in various ways (Section 3.1.4).

$\text{sim}(s_i, s_j)$ is defined by the cosine similarity between two sentence representations from a sentence encoder, specifically fine-tuned for the semantic textual similarity (STS) task (Gao et al., 2021).

From the sentence-level text graph, the sentence-level saliency is defined by the degree-based centrality (Zheng and Lapata, 2019). In other words, this centrality is equivalent to the sum of semantic similarities with all of its subsequent sentences.

$$S^s(s_i) = \sum_{s_j \in \mathcal{S}} E_{ij}^s. \quad (2)$$

3.1.2 Relation-level Saliency Score

The relation-level saliency focuses on the meaning of each relation itself in that the semantic similarity among the relation descriptions implies the saliency; that is, a relation description that is more relevant to other relation descriptions is more likely to contain salient information. In this sense, we build a relation-level text graph $\mathcal{G}^r = (\mathcal{R}, E^r)$, whose nodes represent the relation triple and the undirected edge has the weight of the semantic similarity between relation descriptions. Similar to Equation (1), the cosine similarity between relation representations is calculated, and the saliency score is also modeled as the degree-based centrality.

$$E_{ij}^r = \text{sim}(\text{desc}(r_i), \text{desc}(r_j)), \\ S^r(r_i) = \sum_{r_j \in \mathcal{R}} E_{ij}^r. \quad (3)$$

Note that the sequential order of relations is not clearly presented unlike the sentences, because multiple relations are extracted from the same sentence.

3.1.3 Phrase-level Saliency Score

The phrase-level saliency measures the saliency of phrases included in each relation triple, and it

captures the phrase frequency and co-occurrence in the document. As presented in previous work on keyphrase extraction (Mihalcea and Tarau, 2004; Bougouin et al., 2013), we build a phrase-level text graph $\mathcal{G}^p = (\mathcal{P}, E^p)$ whose nodes are the noun and verb phrases extracted from the source document based on POS tags (e.g., Noun, Proper Noun, and Verb). The undirected edges model the weight as how many times two phrases locally co-occur (i.e., within a sliding window) in the sentences \mathcal{S} , and then run the TextRank (Mihalcea and Tarau, 2004) on the graph to compute salience of phrase nodes.

$$E_{ij}^p = \text{co-occur}(p_i, p_j; \mathcal{S}), \quad (4)$$

$$S^p(p_i) = (1 - d) + d \cdot \sum_{p_j \in \mathcal{P}} \frac{E_{ji}^p}{\sum_{p_k \in \mathcal{P}} E_{jk}^p} S^p(p_j),$$

where $d \in [0, 1]$ is the damping factor that indicates the transition probability from one node to another random node. Starting from initial values of S^p usually set to 1.0 for all the nodes, the final salience of each phrase is obtained through iterative computation of Equation (4) until convergence.

3.1.4 Salient Relation Triple Selection

The remaining challenge here is to select relation triples by integrating multi-level salience scores. To this end, RTSUM first identifies the textual information units relevant to each relation triple r_i , including its source sentence s_j and its phrases $p_k (\in \mathcal{P}_{r_i})$, and then transforms their salience scores for the relation triple by $S^s(r_i) := S^s(s_j)$, $S^r(r_i) := S^r(r_i)$, and $S^p(r_i) := 1/|\mathcal{P}_{r_i}| \cdot \sum_{p_k \in \mathcal{P}_{r_i}} S^p(p_k)$.

The most straightforward strategy to select a small number of salient relation triples is to calculate the final score of each relation triple based on *weighted summation* of its three distinct scores and to select the top- K relation triples:

$$S(r_i) = \alpha \cdot S^s(r_i) + S^r(r_i) + \beta \cdot S^p(r_i). \quad (5)$$

Another selection strategy is to adopt *cascade filtering* that excludes less salient relation triples by using the sentence-level, relation-level, and phrase-level salience in a serial order (i.e., $S^s \rightarrow S^r \rightarrow S^p$). The key principle of this filtering process is to keep only the relation triples extracted from the key sentences, and among them, to selectively collect the relation triples that are semantically relevant to the others, and finally, to exclude the ones that do not include many salient phrases.

3.2 Information Sentencification

For the generation of sentences from the selected relation triples (i.e., sentencification), RTSUM builds a *relation combiner* based on a pretrained text-to-text language model, such as BART (Lewis et al., 2020) and T5 (Raffel et al., 2020). Using the relation combiner, RTSUM can perform the abstractive summarization by sentencifying the selected relation triples.

Relation Combiner Training. The relation combiner is effectively optimized with the self-supervised objective for the sentencification task. To be specific, we collect training pairs of (relation triples, sentences) by randomly sampling a couple of sentences from source documents and extracting the relation triples from the sentences. Then, we train the relation combiner based on Maximum Likelihood Estimation (MLE) to generate the sentences by taking the concatenated text of all the extracted relation triples. As a result, it is expected to learn how to introduce linking words, place each component in order considering their relation, and remove duplicated phrases or entities, for plausible sentence generation. To eliminate redundancy, we apply a lightweight string similarity algorithm, Gestalt pattern matching (Ratcliff and Metzner, 1988), as a filter before merging relation triples.

Training Pair Filtering. Despite the benefits of self-supervised training, the relation combiner still has a risk of introducing information that is not presented in a source document (i.e., extrinsic hallucinations) or factual errors against the document (i.e., intrinsic hallucinations) into its output summary. Since the extracted relation triples are not guaranteed to perfectly cover all the information of their source sentences, some training pairs might guide the combiner to generate missing information that does not exist in the input relation triples. To alleviate these hallucinations, we selectively collect the training pairs whose extracted relation triples fully cover the content of the source sentences. Precisely, the pairs of (relation triples, sentences) are excluded from the training set, in case that some of the semantic tokens (i.e., nouns, proper nouns, and verbs) in the source sentences do not appear in the extracted relation triples.

4 Demo: Interpretable Summarizing Tool

Based upon our RTSUM framework, we build an interpretable summarizing tool that provides not only the final summary of an input document but

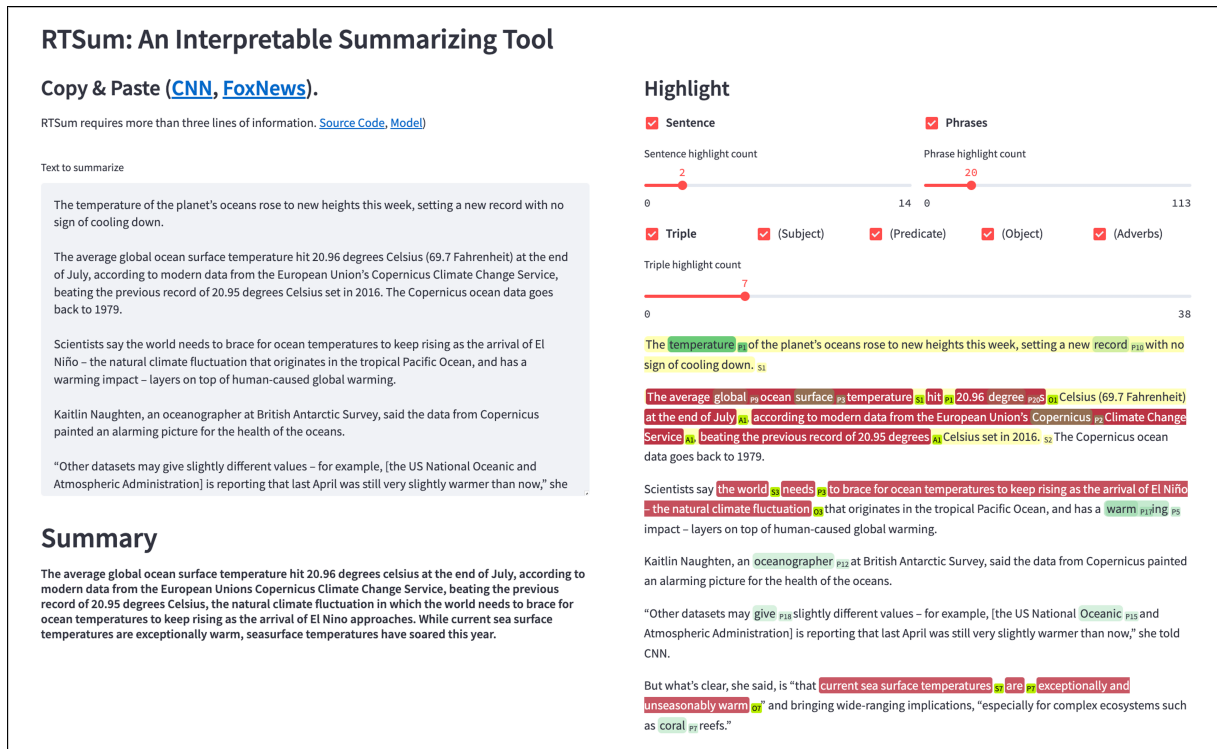


Figure 2: Our interpretable summarizing tool features multi-level salience visualization. Sentences, relation triples, and phrases with a high score are highlighted in yellow, red, and green, respectively. The saliency of each unit is denoted by its opacity. Within each triple, the subjects, predicates, objects, and adverbs are distinguished.

also its fine-grained interpretations.

4.1 Multi-level Salience Visualization

For interpretation of final summaries, our tool provides salience visualization for textual information units with different granularity (Figure 2). It highlights the text spans that correspond to each information unit according to its salience score. For enhanced insights, the salience rank is explicitly annotated next to each span, providing users with a clear understanding of the relative importance of each information unit within the document. This feature allows users to grasp the significance of the textual content and gain a more nuanced and detailed understanding of the document's key points.

Users can personalize the salience visualization based on their unique preferences and specific needs, including customization options as follows:

- **Type of textual units:** Users have the flexibility to choose whether to highlight each type of textual unit. They can opt to further dissect the highlight for a relation triple, differentiating its subject, predicate, and object components.
- **Number of textual units:** Users can manually adjust the number of highlighted instances for

each type of textual unit.

4.2 Implementation Details

Text graph construction. For more reliable summarization, our RTSUM implementation filters out less confident relation triples among the ones extracted from the OpenIE 5 system; only the relation triples of which confidence is larger than 0.7 is considered as valid units. To construct sentence-level and relation-level text graphs, RTSUM utilizes General Text Embeddings (GTE)⁶ as the sentence encoder, which is trained on a large-scale corpus of relevance text pairs covering a wide range of domains and scenarios. Cosine similarity between two sentence (or relation description) embeddings is used for the edge weight in the graphs.

Relation triple selection. RTSUM in our tool simply ranks relation triples by their final salience scores, which are calculated by summing three distinct salience scores (i.e., S^s , S^r , S^p) with the same weight, and then chooses top- K ones. The number of relation triples to be selected is set to $K = 3$.

Relation combiner training. To build a relation combiner, we fine-tune BART⁷ (Lewis et al.,

⁶<https://huggingface.co/thenlper/gte-large>

⁷<https://huggingface.co/facebook/bart-base>

2020) to generate source sentences from the relation triples extracted from the sentences. We use a text corpus in the news domain, CNN/DM (Nallapati et al., 2016) which contains 287,113 news articles available for training. To reduce the risk of hallucination, we filter out the cases that the amount of information in an input text (i.e., a set of relation triples) is shorter than that in an output text (i.e., sentences), as explained in Section 3.2. In addition, we use three special tokens, <subject>, <predicate>, and <object>, to separate three components of each relation triple in an input text, which effectively provides structured information about each triple to the model.

Relation combiner alternatives. While our summarizing tool provides the fine-tuned text-to-text language model as a default relation combiner, it also provides an option to employ instruction-following language models, such as InstructGPT (Ouyang et al., 2022) and ChatGPT. These models can reconstruct plausible sentences from a set of relation triples, when being asked with a proper prompt written in natural language; they can be beneficial in that domain-specific or task-specific fine-tuning process is not required.

5 Related Work

5.1 Unsupervised Extractive Summarization

The most popular approach to unsupervised extractive summarization is to identify key sentences by using a text graph that represents the semantic (or lexical) relationship among text units in a source document. TextRank (Mihalcea and Tarau, 2004) is the first work to adopt a graph-based ranking algorithm (Brin and Page, 1998) to calculate the centrality of sentences in the graph, whose node represents each sentence and edge is modeled as the similarity between two sentences. Several variants of TextRank have been implemented by utilizing symbolic sentence representations (e.g., TF-IDF) (Barrios et al., 2016) or distributed sentence representations (e.g., skip-thoughts) (Kiros et al., 2015) for computing the sentence similarity.

Most recent studies have employed pretrained language models (PLMs), such as BERT (Devlin et al., 2019), to effectively model the salience of each sentence. Zheng and Lapata (2019); Liu et al. (2021) used the degree-based node centrality of the position-augmented sentence graph where the sentence similarity is calculated by PLMs, and Padmakumar and He (2021) defined the selection cri-

terion by using PLM-based pointwise mutual information. Xu et al. (2020b) considered the sentence-level self-attention score as the salience, after optimizing PLMs via masked sentence prediction. Nevertheless, all of them regard a sentence as the basic unit for summarization, so they cannot exclude unnecessary information from each selected sentence.

5.2 Unsupervised Abstractive Summarization

To train a neural model for abstractive summarization without using human-annotated text-summary pairs, most existing methods have adopted the auto-encoding architecture whose encoder compresses a source text into a readable summary (i.e., a few sentences) and decoder reconstructs the original text from the summary (Wang and Lee, 2018; Baziotis et al., 2019; Chu and Liu, 2019). Another line of research has focused on zero-shot abstractive summarization, which takes advantage of large-scale PLMs trained on massive text corpora. Their models are optimized with a self-supervised objective (e.g., gap sentence generation) (Raffel et al., 2020; Zhang et al., 2020) or heuristically-generated references (e.g., lead bias) (Yang et al., 2020; Fang et al., 2022). However, the well-known caveat of abstractive summarization is poor interpretability, which is also related to the hallucination problem; their output summaries mostly contain factual errors or misinformation against the source document (Kryscinski et al., 2020; Maynez et al., 2020).

6 Conclusion

In this paper, we introduce a summarization framework, called RTSUM, which leverages relation triples as the basic units for summarization. Building upon this framework, we have developed a web demo for an interpretable summarizing tool that effectively visualizes the salience of textual units at three distinct levels. Through our multi-level salience visualization, users can easily identify textual units impacting the summary and gain insights into the document’s salient semantic structure.

Our RTSUM framework and its user-friendly tool can effectively capture the essence of a document while maintaining interpretability. The fusion of extractive and abstractive approaches, coupled with intuitive multi-level visualization, holds promise for applications requiring succinct, accurate, and interpretable summaries.

7 Limitations

Our study has the following limitations: Firstly, compared to single-step summarization approaches, our framework is relatively slower due to its multi-step process. Secondly, the current implementation relies on English-specific tools for sentence splitting and relation extraction, limiting its applicability to only English inputs. Lastly, while our research focuses on summarizing news articles effectively, the robustness and performance of our approach on longer or differently formatted text genres, such as books or research papers, has not been comprehensively evaluated.

Acknowledgements

This work was supported by the IITP grant funded by the Korea government (MSIT) (No.2020-0-01361) and the NRF grant funded by the Korea government (MSIT) (No. RS-2023-00244689).

References

- Gabor Angeli, Melvin Jose Johnson Premkumar, and Christopher D. Manning. 2015. [Leveraging linguistic structure for open domain information extraction](#). In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 344–354, Beijing, China. Association for Computational Linguistics.
- Federico Barrios, Federico López, Luis Argerich, and Rosa Wachenchauzer. 2016. [Variations of the similarity function of textrank for automated summarization](#). *CoRR*, abs/1602.03606.
- Christos Baziotis, Ion Androutsopoulos, Ioannis Konstas, and Alexandros Potamianos. 2019. [SEQ³: Differentiable sequence-to-sequence-to-sequence autoencoder for unsupervised abstractive sentence compression](#). 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 673–681, Minneapolis, Minnesota. Association for Computational Linguistics.
- Adrien Bougouin, Florian Boudin, and Béatrice Daille. 2013. [TopicRank: Graph-based topic ranking for keyphrase extraction](#). In *Proceedings of the Sixth International Joint Conference on Natural Language Processing*, pages 543–551, Nagoya, Japan. Asian Federation of Natural Language Processing.
- Sergey Brin and Lawrence Page. 1998. The anatomy of a large-scale hypertextual web search engine. In *Proceedings of the Seventh International Conference on World Wide Web 7*, page 107–117, NLD. Elsevier Science Publishers B. V.
- Eric Chu and Peter J. Liu. 2019. [Meansum: A neural model for unsupervised multi-document abstractive summarization](#). In *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, volume 97 of *Proceedings of Machine Learning Research*, pages 1223–1232. PMLR.
- Arman Cohan, Franck Dernoncourt, Doo Soon Kim, Trung Bui, Seokhwan Kim, Walter Chang, and Nazli Goharian. 2018. [A discourse-aware attention model for abstractive summarization of long documents](#). In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, pages 615–621, New Orleans, Louisiana. Association for Computational Linguistics.
- 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.
- Yuwei Fang, Shuohang Wang, Yichong Xu, Ruochen Xu, Siqi Sun, Chenguang Zhu, and Michael Zeng. 2022. [Leveraging knowledge in multilingual commonsense reasoning](#). In *Findings of the Association for Computational Linguistics: ACL 2022*, pages 3237–3246, Dublin, Ireland. Association for Computational Linguistics.
- Tianyu Gao, Xingcheng Yao, and Danqi Chen. 2021. [SimCSE: Simple contrastive learning of sentence embeddings](#). In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 6894–6910, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.
- Ryan Kiros, Yukun Zhu, Russ R Salakhutdinov, Richard Zemel, Raquel Urtasun, Antonio Torralba, and Sanja Fidler. 2015. [Skip-thought vectors](#). In *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc.
- Huan Yee Koh, Jiaxin Ju, He Zhang, Ming Liu, and Shirui Pan. 2022. [How far are we from robust long abstractive summarization?](#) In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 2682–2698, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.
- Wojciech Kryscinski, Bryan McCann, Caiming Xiong, and Richard Socher. 2020. [Evaluating the factual consistency of abstractive text summarization](#). In

- Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 9332–9346, Online. Association for Computational Linguistics.
- 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.
- Jingzhou Liu, Dominic J. D. Hughes, and Yiming Yang. 2021. [Unsupervised extractive text summarization with distance-augmented sentence graphs](#). In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '21*, page 2313–2317, New York, NY, USA. Association for Computing Machinery.
- Mausam Mausam. 2016. Open information extraction systems and downstream applications. In *Proceedings of the twenty-fifth international joint conference on artificial intelligence*, pages 4074–4077.
- Joshua Maynez, Shashi Narayan, Bernd Bohnet, and Ryan McDonald. 2020. [On faithfulness and factuality in abstractive summarization](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 1906–1919, Online. Association for Computational Linguistics.
- Rada Mihalcea and Paul Tarau. 2004. [TextRank: Bringing order into text](#). In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, pages 404–411, Barcelona, Spain. Association for Computational Linguistics.
- Ramesh Nallapati, Bowen Zhou, Caglar Gulcehre, Bing Xiang, et al. 2016. [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.
- Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. 2022. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35:27730–27744.
- Vishakh Padmakumar and He He. 2021. Unsupervised extractive summarization using pointwise mutual information. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 2505–2512.
- 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*, 21:140:1–140:67.
- John W. Ratcliff and David E. Metzener. 1988. Pattern matching: The gestalt approach. In *Dr. Dobb's Journal of Software Tools*, volume 13, pages "46, 47, 59–51, 68–72". j-DDJ.
- 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 (Volume 1: Long Papers)*, pages 1073–1083, Vancouver, Canada. Association for Computational Linguistics.
- Jiwei Tan, Xiaojun Wan, and Jianguo Xiao. 2017. [Abstractive document summarization with a graph-based attentional neural model](#). In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1171–1181, Vancouver, Canada. Association for Computational Linguistics.
- Yaoshian Wang and Hung-Yi Lee. 2018. [Learning to encode text as human-readable summaries using generative adversarial networks](#). In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 4187–4195, Brussels, Belgium. Association for Computational Linguistics.
- Jiacheng Xu, Shrey Desai, and Greg Durrett. 2020a. [Understanding neural abstractive summarization models via uncertainty](#). In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 6275–6281, Online. Association for Computational Linguistics.
- Shusheng Xu, Xingxing Zhang, Yi Wu, Furu Wei, and Ming Zhou. 2020b. [Unsupervised extractive summarization by pre-training hierarchical transformers](#). In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 1784–1795, Online. Association for Computational Linguistics.
- Ziyi Yang, Chenguang Zhu, Robert Gmyr, Michael Zeng, Xuedong Huang, and Eric Darve. 2020. [TED: A pretrained unsupervised summarization model with theme modeling and denoising](#). In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 1865–1874, Online. Association for Computational Linguistics.
- 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'20*. JMLR.org.
- Hao Zheng and Mirella Lapata. 2019. [Sentence centrality revisited for unsupervised summarization](#). In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence, Italy. Association for Computational Linguistics.