Rewards with Negative Examples for
Reinforced Topic-Focused Abstractive Summarization

Anonymous EMNLP submission

Abstract

We consider the problem of topic-focused abstractive summarization, where the goal is to generate an abstractive summary focused on a particular topic, a phrase of one or multiple words. We hypothesize that the task of generating topic-focused summaries can be improved by showing the model what it must *not* focus on. We introduce a deep reinforcement learning approach to topic-focused abstractive summarization, trained on rewards with a novel negative example baseline. We define the input in this problem as the source text preceded by the topic. We adapt the CNN-Daily Mail and New York Times summarization datasets for this task. We then show through experiments on existing rewards that the use of a negative example baseline can outperform the use of a self-critical baseline, in ROUGE, BERTSCORE, and human evaluation metrics.

1 Introduction

Topic-focused summarization is the task of generating a summary given a source text and a specific query or topic. Approaches to topic-focused summarization include query relevance and importance (Gupta et al., 2007), multi-modality manifold ranking (Wan et al., 2007; Wan, 2008; Wan and Xiao, 2009), and query attention (Nema et al., 2017). The DUC 2005 and 2006 datasets (Dang, 2005, 2006) are examples of datasets that are widely used for this task. These datasets are much smaller than benchmark datasets for generic summarization, resulting in fewer research work to train topic-focused summarization on state-of-the-art systems (Deutsch and Roth, 2019).

In parallel, there has been growing work in recent years on reinforcement learning approaches to (generic) abstractive summarization. Proposed rewards aim to optimize non-differentiable summarization metrics like ROUGE (Lin, 2004) and BERTSCORE (Zhang et al., 2019), or to encourage desirable summary aspects like semantic cohe-

sion (Celikyilmaz et al., 2018) and entity coherence (Sharma et al., 2019). Many reinforced abstractive summarization methods use the self-critical baseline or SCST (Rennie et al., 2017) to cap their rewards. This self-critical baseline is obtained by greedily searching for a sequence that maximizes the likelihood probability of the current model.

In this work, we propose a reinforcement learning-based approach to topic-focused summarization. First, we adapt widely used generic summarization benchmarks to this task, such that we aim to generate only one out of three summary sentences, given a corresponding topic. Then, instead of using the self-critical baseline, we introduce a novel baseline that uses negative examples: a sentence that contains information that the summarization model should *not* focus on.

We run experiments on two existing generic summarization datasets adapted to our task: CNN-Daily Mail (Hermann et al., 2015; Nallapati et al., 2017) and New York Times (Sandhaus, 2008). Our experiments span two existing rewards: the popular ROUGE-L reward and the Distributed Semantic Reward (DSR) of Li et al. (2019), inspired by BERTSCORE (Zhang et al., 2019). Our results show that using our negative example baseline outperforms the self-critical baseline across both datasets and both rewards. We obtain improvements on both datasets in ROUGE and BERTSCORE metrics, and human annotators find that summaries generated with our negative baseline for rewards are generally more relevant to the given topic.

2 Related Work

Topic-Focused Summarization. There are different definitions of topic-focused summarization. The DUC datasets (Dang, 2005, 2006) propose summarization of documents given a question, also called a query. Vanderwende et al. (2007) propose SumFocus, a system for topic-focused multi-

document extractive summarization. SumFocus is comprised of four components: a generic extractive summarization system, a topic-focusing component, sentence simplification, and lexical expansion of topic words.

Deutsch and Roth (2019) define the task of *summary cloze* as the problem of deciding which content to select in topic-focused summarization, given a context (partial summary). They propose a neural model with separate encoders for the topic and the partial summary.

Reinforcement Learning for Summarization. There is a growing body of work that use reinforcement learning (RL) methods to optimize non-differentiable rewards. ROUGE scores remain a popular RL reward. Other rewards include sentence selection to improve ROUGE scores (Chen and Bansal, 2018; Pasunuru and Bansal, 2018), optimizing question answering metrics (Scialom et al., 2019), and adding desirable custom features to generated summaries (Böhm et al., 2019; Sharma et al., 2019).

Li et al. (2019) find that rewards based on BERTSCORE (Zhang et al., 2019) and ROUGE each optimize their own metric, but decrease the other one.

Wang et al. (2018) introduce a topic-aware reinforced summarization model. The authors experiment with generic – not topic-focused – summarization datasets, and infer topics using LDA (Blei et al., 2003). Information about the topics is then infused into the model using a topic-aware attention mechanism and topic embeddings.

Whereas a few (Narayan et al., 2018; Li et al., 2018) use the REINFORCE algorithm (Williams, 1992), many RL-based summarization approaches (Paulus et al., 2018; Pasunuru and Bansal, 2018; Li et al., 2018; Celikyilmaz et al., 2018; Yang et al., 2018; Li et al., 2019) use the self-critical sequence training approach (SCST) (Rennie et al., 2017).

3 Problem Statement

We tackle the task of topic-focused abstractive summarization as the problem of producing an abstractive summary focused on a given topic, a phrase of one or multiple words. The generated summary should include information from the input text that is related to the topic, and exclude all other information. Consequently, different topics with the same input text should yield different summaries.

More formally, given an input text $\mathbf{x}$, a topic $\mathbf{t}$ and a corresponding reference summary $\mathbf{y}$, we aim to maximize the probability that we generate the right summary:

$$p(\mathbf{y}|\mathbf{x}, \mathbf{t}) > p(\mathbf{y}'|\mathbf{x}, \mathbf{t}) \quad (1)$$

for all $\mathbf{y}' \neq \mathbf{y}$.

4 Rewards with a Negative Example

In self-critical sequence learning (Rennie et al., 2017; Wang et al., 2018), the RL loss formula is as follows for our task:

$$\mathcal{L}_{\text{RL}} = -(r(\mathbf{y}^s) - b) \sum_{i=1}^{N} \log P(y_i^s|y_1^s, ..., y_{i-1}^s, \mathbf{x}, \mathbf{t}) \quad (2)$$

where b is the RL baseline, and $b = r(\hat{\mathbf{y}})$ in self-critical sequence learning. $\mathbf{y}^s$ is a sampled summary, and $\hat{\mathbf{y}}$ is a summary obtained greedily by maximizing the probability of the overall sequence.

In our particular task, we propose to use a baseline with a negative example. The intuition is that we encourage the model to generate summaries that are more similar to the reference summary than the negative example. This negative example is an independent sentence from the summary of the corresponding source text, but which does not contain the topic. The negative example acts as a sample of undesirable information, and helps the summarization model learn what kind of information to exclude.

Given a reference summary $\mathbf{y}$, and a negative summary $\bar{\mathbf{y}}$, our RL loss with negative examples is defined as in equation 2, where we define the RL baseline: $b = r(\bar{\mathbf{y}})$.

We apply our method on two popular rewards for summarization: rewards based on the ROUGE and BERTSCORE metrics between the sampled summary and the reference summary.

ROUGE-L reward. Given a sampled summary $\mathbf{y^s}$ and a reference summary $\mathbf{y}$, we define the ROUGE-L reward as follows:

$$r_{\mathbf{R}}(\mathbf{y^s}) = \text{ROUGE}(\mathbf{y^s}, \mathbf{y}) \quad (3)$$

BERTSCORE Reward. We adopt the Distributed Semantic Reward (DSR) definition of Li et al. (2019). This reward measures the semantic similarity with the reference summary. It is defined as follows:

$$r_{\mathbf{s}}(\mathbf{y^s}) = \text{F}_{\text{BERT}}(\mathbf{y^s}, \mathbf{y}) \quad (4)$$

where F_{BERT} is the F1 formula of BERTScore (Zhang et al., 2019). It is defined as:

$$F_{BERT} = 2 * \frac{P_{BERT} * R_{BERT}}{P_{BERT} + R_{BERT}} \quad (5)$$

where the precision P_{BERT} and recall R_{BERT} are defined as follows for a given reference y and candidate y':

$$P_{BERT} = \frac{1}{|y'|} \sum_{\hat{x}_j \in y'} \max_{x_i \in y} \mathbf{x}_i^\top \hat{\mathbf{x}}_j \quad (6)$$

$$R_{BERT} = \frac{1}{|y|} \sum_{x_i \in y} \max_{\hat{x}_j \in y'} \mathbf{x}_i^\top \hat{\mathbf{x}}_j \quad (7)$$

Loss Formula. Whereas Pasunuru and Bansal (2018) train by alternating multiple rewards, Li et al. (2019) propose a single loss formula combining DSR and ROUGE rewards. However, their results show that combining DSR and ROUGE does not yield better results in either ROUGE or F_{BERT} scores, compared to using only one reward at a time, along with the summarization loss. We decide to also use only one reward at a time.

We aim to optimize the following loss function:

$$\mathcal{L} = -(1 - \gamma) * \log p(\mathbf{y}|\mathbf{x}, \mathbf{t}) \\ + \gamma * \mathcal{L}_{RL-N}(\mathbf{y}^s, \mathbf{y}, \bar{\mathbf{y}}) \quad (8)$$

where γ is a hyperparameter, and the first term is the negative log-likelihood loss with the reference summary as the target.

5 Experiments and Results

Datasets. In our experiments, we use two popular summarization benchmarks: the non-anoymized CNN-Daily Mail dataset (Hermann et al., 2015; Nallapati et al., 2017) and the New York Times (NYT) dataset (Sandhaus, 2008).

In the CNN-Daily Mail dataset, summaries are usually 3 independent sentences, corresponding to the "highlights" at the top of the article on the original website. We consider each sentence separately as a reference summary, thereby creating on average 3 datapoints from 1 datapoint in the original dataset.

We filter the large NYT dataset to only get articles with 3-sentence summaries. Similarly to the CNN-Daily Mail dataset, summary sentences are independent, making them fit for topic-focused abstractive summarization.

Dataset	Set	# datapoints
CNN-Daily Mail	Train	748,170
	Dev	20,705
	Test	31,815
NYT	Train	139,497
	Dev	9,911
	Test	16,021

Table 1: Statistics for our focused version of the CNN-Daily Mail and NYT datasets.

We use TopicRank (Bougouin et al., 2013), a keyphrase extraction algorithm, to get the 10 most popular keyphrases of the input text. Out of these 10 keyphrases, we pick the highest-scoring one which only appears in the specific summary sentence $\mathbf{y}_i$ to be the topic of $\mathbf{y}_i$. We consider one of the other reference summary sentences of the same input text as the negative example. Therefore, each datapoint in our version of the datasets contains an input text, a reference summary, a negative summary, and a topic. We make sure that no same input text (article) appears in more than one data split. We show the statistics of the dataset used in Table 1.

Training Details. We set $\gamma = 0.9984$ following Paulus et al. (2018) to balance the magnitude difference. We adopt the BART Large architecture (Lewis et al., 2019) as it set a state of the art in the generic summarization of the CNN-Daily Mail dataset, among other tasks. We use a learning rate of $3e-5$. We start training from the best model trained on the cross-entropy objective only, on our topic-focused summarization datasets.

We prepend the topic and a separator token to the input text of each datapoint. As the average reference summary is shorter (on average 13 tokens), we set the length of generated summaries at test time between 10 and 20 tokens.

Baseline Models. We train four extractive baseline models and three abstractive summarization baseline models.

The four extractive baseline models are common summarization baselines, that are meant to give an idea about the difficulty of the task. The first baseline model is **Lead-1**, which chooses the first sentence from the source article as the generated summary. The second baseline model is **BM25** (Robertson et al., 1995), an IR-based score to rank search results given the query. In our case, BM25 ranks sentences of the source article given

DATASET		CNN-DAILY MAIL				NEW YORK TIMES			
METRIC		R1 F1	R2 F1	RL F1	F_{BERT}	R1 F1	R2 F1	RL F1	F_{BERT}
EXTRACTIVE SUMMARIZATION BASELINE MODELS									
Lead-1		19.94	6.54	16.33	69.08	22.22	9.89	18.44	69.48
BM25		22.92	9.71	19.79	69.95	30.38	15.80	26.45	73.23
SumFocus		32.08	15.82	28.14	73.92	29.00	14.64	25.27	72.86
Oracle-1		47.52	29.85	44.10	79.47	53.34	36.15	49.49	80.64
NEURAL ABSTRACTIVE SUMMARIZATION MODELS									
REWARD	RL b								
None	–	38.88	21.91	35.10	76.71	45.84	30.93	40.83	79.32
ROUGE	Self-critical	39.64	22.78	37.10	74.82	46.52	32.43	42.49	77.69
ROUGE	Negative (Ours)	**39.97**	**23.22**	**37.87**	74.12	**47.02**	**32.87**	**43.12**	78.03
BERTSCORE	Self-critical	39.10	21.83	36.77	77.34	45.67	31.62	41.70	79.93
BERTSCORE	Negative (Ours)	39.38	22.41	37.29	**77.65**	45.51	31.25	41.57	**80.35**

Table 2: Summarization experiment results on the test sets with the ROUGE and BERTSCORE rewards. The model with no reward is a summarizer model trained only on the cross-entropy objective.

DATASET		CNN-DAILY MAIL			NEW YORK TIMES		
REWARD	CRITERIA	Negative	Self-Critical	Tie	Negative	Self-Critical	Tie
ROUGE	Relevance to Topic	25	21	34	21	14	45
	Fluency	17	21	42	14	12	54
BERTSCORE	Relevance to Topic	19	15	46	22	18	40
	Fluency	16	11	51	18	15	47

Table 3: Human evaluation ratings of two annotators on 40 sampled summaries from each dataset, comparing reinforced summarization models trained with the negative baseline (ours) vs. the self-critical baseline.

the topic, and outputs the most relevant sentence as the generated summary. The third baseline model is **SumFocus** (Vanderwende et al., 2007), an unsupervised probabilistic model for topic-focused summarization. The fourth baseline is **Oracle-1**, which greedily searches for the sentence with the highest ROUGE score with the reference summary. This baseline model is meant as an upper-bound for extractive summarization models.

We experiment with three abstractive summarization baseline models that use BART Large. The first baseline model is trained on cross-entropy only. The second and third baseline models are trained with the **ROUGE** and **BERTSCORE** (DSR) rewards respectively, with the self-critical method.

Results and Discussion. We show the results of our experiments in Tables 2. Our proposed approach outperforms the cross-entropy-only BART baseline, but also the two self-critical approaches across both datasets. This shows that negative examples are a good reward baseline in topic-focused summarization. We notice a significant jump in performance in ROUGE-L F1 especially (about 2.5 points), and an increase in ROUGE-1 and ROUGE-2

as well, compared to the cross-entropy-only baseline. Our BERTSCORE-rewarded model achieves the highest F_{BERT} scores, with a slight increase from its self-critical counterpart.

We hire two annotators to judge the fluency and topic relevance of the 40 sampled summary pairs, and ask them to compare our models with their self-critical counterparts. The annotators are not informed about which model generated which summary. Results in Table 3 show that our model's summaries are generally more relevant to the topic, and that our BERTSCORE models are more fluent.

6 Conclusions

We propose a deep reinforcement learning approach to topic-focused abstractive summarization. We introduce a new baseline for rewards, based on negative examples collected from independent summary sentences. We show through experiments that our proposed approach outperforms the baseline of self-critical reinforcement learning in the optimized reward metric, and human annotators find our model generates summaries that are more relevant to the topic.

References

David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022.

Florian Böhm, Yang Gao, Christian M Meyer, Ori Shapira, Ido Dagan, and Iryna Gurevych. 2019. Better rewards yield better summaries: Learning to summarise without references. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3101–3111.

Adrien Bougouin, Florian Boudin, and Béatrice Daille. 2013. Topicrank: Graph-based topic ranking for keyphrase extraction.

Asli Celikyilmaz, Antoine Bosselut, Xiaodong He, and Yejin Choi. 2018. Deep communicating agents for abstractive summarization. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1662–1675.

Yen-Chun Chen and Mohit Bansal. 2018. Fast abstractive summarization with reinforce-selected sentence rewriting. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 675–686.

Hoa Trang Dang. 2005. Overview of duc 2005. In *Proceedings of the document understanding conference*, pages 1–12.

Hoa Trang Dang. 2006. Overview of duc 2006. In *Proceedings of the document understanding conference*.

Daniel Deutsch and Dan Roth. 2019. Summary cloze: A new task for content selection in topic-focused summarization. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3711–3720.

Surabhi Gupta, Ani Nenkova, and Dan Jurafsky. 2007. Measuring importance and query relevance in topic-focused multi-document summarization. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions*, pages 193–196.

Karl Moritz Hermann, Tomas Kocisky, Edward Grefenstette, Lasse Espeholt, Will Kay, Mustafa Suleyman, and Phil Blunsom. 2015. Teaching machines to read and comprehend. In *Advances in neural information processing systems*, pages 1693–1701.

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

Piji Li, Lidong Bing, and Wai Lam. 2018. Actor-critic based training framework for abstractive summarization. *arXiv preprint arXiv:1803.11070*.

Siyao Li, Deren Lei, Pengda Qin, and William Yang Wang. 2019. Deep reinforcement learning with distributional semantic rewards for abstractive summarization. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 6038–6044, Hong Kong, China. Association for Computational Linguistics.

Chin-Yew Lin. 2004. Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, pages 74–81.

Ramesh Nallapati, Feifei Zhai, and Bowen Zhou. 2017. Summarunner: a recurrent neural network based sequence model for extractive summarization of documents. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pages 3075–3081.

Shashi Narayan, Shay B Cohen, and Mirella Lapata. 2018. Ranking sentences for extractive summarization with reinforcement learning. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1747–1759.

Preksha Nema, Mitesh M Khapra, Anirban Laha, and Balaraman Ravindran. 2017. Diversity driven attention model for query-based abstractive summarization. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1063–1072.

Ramakanth Pasunuru and Mohit Bansal. 2018. Multi-reward reinforced summarization with saliency and entailment. 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 646–653.

Romain Paulus, Caiming Xiong, and Richard Socher. 2018. A deep reinforced model for abstractive summarization. In *International Conference on Learning Representations*.

Steven J Rennie, Etienne Marcheret, Youssef Mroueh, Jerret Ross, and Vaibhava Goel. 2017. Self-critical sequence training for image captioning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7008–7024.

Stephen E Robertson, Steve Walker, Susan Jones, Micheline M Hancock-Beaulieu, Mike Gatford, et al. 1995. Okapi at trec-3. *Nist Special Publication Sp*, 109:109.

Evan Sandhaus. 2008. The new york times annotated corpus. *Linguistic Data Consortium, Philadelphia*, 6(12):e26752.

Thomas Scialom, Sylvain Lamprier, Benjamin Piwowarski, and Jacopo Staiano. 2019. Answers unite! unsupervised metrics for reinforced summarization models. *arXiv preprint arXiv:1909.01610*.

Eva Sharma, Luyang Huang, Zhe Hu, and Lu Wang. 2019. An entity-driven framework for abstractive summarization. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3271–3282.

Lucy Vanderwende, Hisami Suzuki, Chris Brockett, and Ani Nenkova. 2007. Beyond sumbasic: Task-focused summarization with sentence simplification and lexical expansion. *Information Processing & Management*, 43(6):1606–1618.

Xiaojun Wan. 2008. Using only cross-document relationships for both generic and topic-focused multi-document summarizations. *Information Retrieval*, 11(1):25–49.

Xiaojun Wan and Jianguo Xiao. 2009. Graph-based multi-modality learning for topic-focused multi-document summarization. In *Twenty-First International Joint Conference on Artificial Intelligence*. Citeseer.

Xiaojun Wan, Jianwu Yang, and Jianguo Xiao. 2007. Manifold-ranking based topic-focused multi-document summarization. In *IJCAI*, volume 7, pages 2903–2908.

Li Wang, Junlin Yao, Yunzhe Tao, Li Zhong, Wei Liu, and Qiang Du. 2018. A reinforced topic-aware convolutional sequence-to-sequence model for abstractive text summarization. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, pages 4453–4460.

Ronald J Williams. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 8(3-4):229–256.

Min Yang, Qiang Qu, Ying Shen, Qiao Liu, Wei Zhao, and Jia Zhu. 2018. Aspect and sentiment aware abstractive review summarization. In *Proceedings of the 27th international conference on computational linguistics*, pages 1110–1120.

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*.