Reading Like HER: Human Reading Inspired Extractive Summarization

Anonymous EMNLP-IJCNLP submission

Abstract

In this work, we re-examine the problem of extractive summarization for long text. Inspired by the reading cognition of human being, the process of extracting summary can be divided into two stages: 1) the *rough reading* looking for sketched information; 2) the subsequent *careful reading* repeatedly selecting essential sentences as summary. To simulate such two-stage procedures, we propose a novel approach for extractive summarization. The overall framework is formulated as a contextual bandit trained with policy gradient reinforcement learning. We adopt a variant of convolutional neural network to encode gist of paragraphs to mimic rough reading, and a bandit policy with an adapted termination mechanism is devised to analogy careful reading. Experiments on the CNN and DailyMail datasets demonstrate that our proposed model can provide high-quality summaries with varied number of sentences at diversed positions and significantly outperform the state-of-the-art extractive methods.

1 Introduction

Automatic text summarization has wide popularity in NLP applications such as producing digests, headlines and reports. Among the supervised methods, two main types are usually explored, namely abstractive and extractive summarizations (Nenkova et al., 2011). Compared with abstractive approaches, extractive methods can be applied in more realistic situations as they are faster, simpler and more reliable on grammar as well as semantic information (Yao et al., 2018).

Recent works (Cheng and Lapata, 2016; Nallapati et al., 2017; Yasunaga et al., 2017; Feng et al., 2018) consider extractive summarization as a sequence labeling task, where each sentence is individually processed and determined whether it should be extracted or not. Various neural networks are widely used to label each sentence and trained using cross-entropy loss to maximize the likelihood of the ground-truth labeled sequences, which may derive the mismatch between the cross-entropy objective function and the evaluation criterion. To solve such issue, some reinforcement learning based methods (Wu and Hu, 2018; Narayan et al., 2018; Yao et al., 2018) directly optimize the evaluation metric by combining cross-entropy loss with rewards from summary evaluation measures and train model parameters with policy gradient reinforcement learning. However, they still sequentially process text and usually focus on earlier sentences over later ones due to the sequential nature of selection (Dong et al., 2018).

Although great efforts have been devoted to this filed, most of the existing approaches neglect human's nature of reading text and forming summaries. Human beings are very good at refining the main idea of a given text and implement it based on their reading cognitive process, which in general includes *pre-reading*, *reading* and *post-reading* (Avery and Graves, 1997; Saricoban, 2002; Toprak and Almacıoğlu, 2009; Pressley and Afflerbach, 2012). In the *pre-reading* stage, they roughly preview the whole text to form an initial cognition and extract general but coarse-grained information at the meantime. Based on such prior knowledge, the subsequent *reading* stage is a conscious process that focuses on target-specific purposes to search fine-grained details through repeated skimming and scanning. For *post-reading*, re-reading is performed to check the missed details. The three-stage reading process makes it effective in capturing essential sentences of text as the extracted summarization.

Inspired by the above human being's reading cognitive process, in this paper, we re-examine the problem of extractive summarization and propose

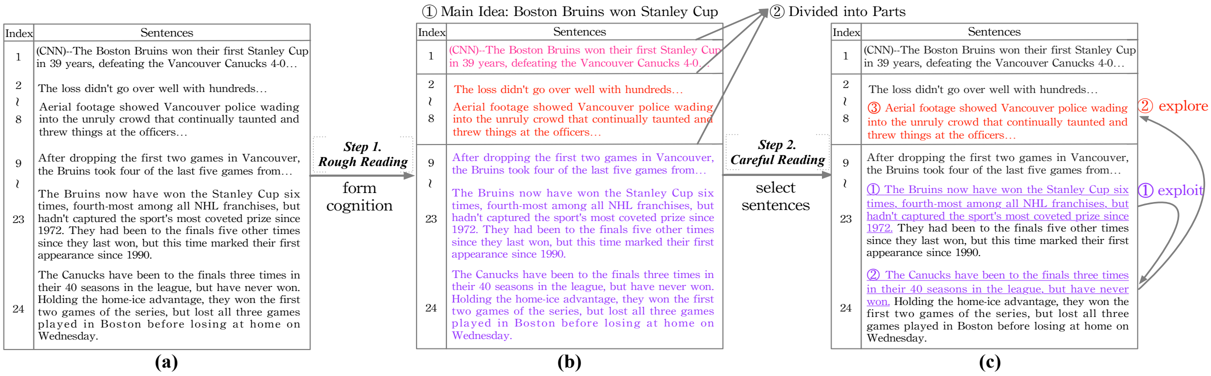

Figure 1: A case of how human being extracts summary. The article is from CNN/DailyMail dataset.

a new approach **HER** (**H**uman-b**E**ing-**R**eading inspired extractive summarization). It consists of two subsequent stages called *rough reading* and *careful reading*. In *rough reading*, coarse-grained information of the original context is identified to form a general cognition. A specific case is shown in Figure 1. In Figure 1 (a) and (b), after browsing on the whole article, the main idea is outlined and the text is roughly divided into three parts based on the gist of paragraphs at the meantime. Each part describes related but not the same contents. To implement the *rough reading* process, we device a hierarchical neural network to encode sentence vectors and derive a document representation as global feature for the main idea. Meanwhile, a variant convolutional neural network (CNN) is utilized to focus on different paragraphs and encode local features to present various parts of the text.

During *careful reading*, the model searches for specific but important details through re-readings to cover the content and extract essential fine-grained information as the final summary. For instance, as shown in Figure 1 (c), after rough reading, two sentences close to the main idea "Boston Bruins won Stanley Cup" may be selected firstly. Then an earlier and more detailed sentence about "fans rioting" is appended to the summary by performing re-reading. It is a combination of people's *reading* and *post-reading* process. To accomplish this, we propose a neural network to decode each sentence to a real-number score. A bandit policy with an adapted termination mechanism is then devised to form summary based on sentence scores.

In our HER model, the whole process is formulated as a contextual bandit which we train an agent to solve using policy gradient reinforcement learning. The agent takes an action which is a to-be-selected sentence set, and then receives a reward based on the correlation between extractive summary and gold-standard reference summary. The contributions of our work are as follows:

- We propose a brand new extractive summarization method that simulates human being's reading cognitive process. The whole framework is formulated as a contextual bandit problem in which two stages named rough reading and careful reading are devised, respectively.
- In rough reading, a hierarchical neural network is used to encode the whole document and a CNN-based network is adopted to capture paragraphical features. In careful reading, a bandit policy with an adapted termination mechanism is devised to flexibly select various but proper numbers of sentences as summaries.
- Experiments conducted on the CNN and DailyMail datasets show our proposed model outperforms the state-of-the-art extractive methods and provides high-quality summaries with varied number of sentences at diverse positions.

2 The HER Model

In this section, we introduce the overall framework of our model HER. We formulate extractive summarization as a contextual bandit trained using policy gradient reinforcement learning where an agent chooses a sentence set as an action based on sampled context and then receives a reward. As illustrated in Figure 2, the framework can be divided into two stages: rough reading and careful reading. During rough reading, a document is encoded into sentence vectors $\{S_1, S_2, \ldots, S_N\}$ and a feature set denoted as F, which includes one global feature D_0 describing the whole documentary information and K local features $\{L_1, L_2, \ldots, L_K\}$

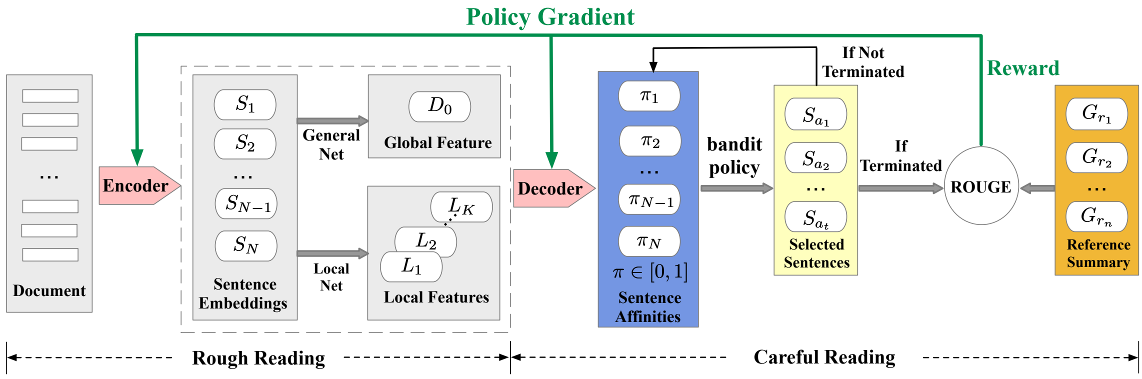

Figure 2: The overall framework of HER is formulated as a contextual bandit and can be divided into a two-stage process containing rough reading and careful reading.

depicting paragraphical contents[1]. In careful reading, sentence vectors are decoded into real-number scores called sentence affinities $\{\pi_1, \pi_2, \dots, \pi_N\}$, which can be considered as an estimation of sentence correlation to cover the context. Then a bandit policy is used to repeatedly choose unique sentence until the termination mechanism is triggered. Next, we will detail the preliminaries in Sec. 2.1, the rough reading stage in Sec. 2.2, and the careful reading stage in Sec. 2.3. The training process is illustrated in Sec. 2.4.

2.1 Preliminaries

In contextual bandit, a context is sampled and shown to the agent, the agent selects an action and then receives a reward which partly depends on the sampled context. The agent's goal is to quickly learn to choose actions which yield more favorable distribution over rewards.

In our model, extractive summarization can be formulated as a contextual bandit. Specifically, context including a document as well as its extra information F extracted from rough reading is sampled and shown to the agent. The goal of the task is to extract a subset of $M \in [1, N]$ sentences out of a N-length document and M is various for different documents. So a sequential sentence indices $a = a_1, \dots, a_M$ can be considered as an action where $a_t \in [1, N], t \in [1, M]$ and there are $2^N - 1$ optional actions in total. After an action a is taken, which means $\{S_{a_1}, \dots, S_{a_M}\}$ is formed as the extractive summary implemented through careful reading, the agent will receive a reward $R(a; G)$. G is the manually-created and gold-standard summary of document D. $R(a; G)$ represents the match degree between G and the extractive summary induced by a. Under such perspective, the problem for us is to train an agent which is based on document features from rough reading and takes an action to form extractive summarization in careful reading.

2.2 Rough Reading

The rough reading aims to produce the feature set F including global and local features to form a general cognition on a given document. F is considered as the contextual information in our framework. First, a hierarchical biLSTM is used to encode a document and obtain hidden sentence embeddings $\{S_1, S_2, \dots, S_n\}$, where $S_i \in \mathbb{R}^{d_s}$ and d_s is the embedding size. Second, global feature $D_0 \in \mathbb{R}^{d_s}$ is computed by an average pooling of all the sentence vectors, which can also be represented by the last hidden unit of the sentence-level biLSTM. Third, we devise a variant of the CNN architecture to refine gist of different paragraphs and generate multiple local features on the sentence level, which is different from previous methods (Kim, 2014; Narayan et al., 2017; Yao et al., 2018) processing on the word level. In detail, a combination of N sentence vectors is,

$$S_{1:N} = S_1 \oplus S_2 \oplus \dots S_N, \quad (1)$$

where $\oplus$ is the concatenation operator and $S_{1:N} \in \mathbb{R}^{N \times d_s}$. A convolution operation involves a filter $w \in \mathbb{R}^{H \times d_s}$ to produce a feature $c \in \mathbb{R}^{N-H+1}$ sliding a window of H sentences. Then a max-over-time pooling operation is applied to produce $\hat{c} = \mathbf{max}\{c\}$ considered as the salient local feature within filter w. We altogether generate K local features by using multiple filters varying K window sizes to present different parts of context.

[1]Although pre-trained models like BERT (Devlin et al., 2018) or Skip-Thought (Kiros et al., 2015) can also be used as encoders, they might be inappropriate in our model as the update taken by policy gradient may make little difference on these large pre-trained models.

2.3 Careful Reading

The careful reading adopts an ϵ-greedy policy to select one sentence at each step based on the contextual information provided by the rough reading and sentence affinities computed by the sentence decoder. Such process will be repeated until the termination mechanism is triggered and all the selected sentences can be formed as a summary.

2.3.1 Sentence Decoder

In order to extract high-quality summaries, we compute the sentence affinities, which is observed effective in Dong et al. (2018), by a sentence decoder. The sentence affinities are calculate by the following principles: (1) **Salience** (The sentences whose meanings are close to the central idea should be emphasized); (2) **Coverage** (The sentences that match different paragraphical information should be encouraged); (3) **Redundancy** (The unselected sentences which are similar to already extracted ones should be inhibited).

Therefore, we utilize a decoder $\mathbf{Dec_1}$ to highlight the sentences presenting Salience or Coverage by applying the global and local features from rough reading. Additionally, a second encoder $\mathbf{Dec_2}$ is devised to screen the sentences that might have Redundancy. Specifically,

$$S'_t = S_t \oplus D_0 \oplus L_{1:K}, t = 1, \ldots, N, \tag{2}$$

$$q_1 = \mathbf{Dec_1}(S'_{1:N}), \tag{3}$$

$$q_2 = \mathbf{Dec_2}(S'_{1:N} \times (1 - q_1)), \tag{4}$$

where $\mathbf{Dec_1}$ and $\mathbf{Dec_2}$ are both constructed with a multilayer perceptron. We use $q = \alpha q_1 + (1 - \alpha)q_2$ as the final sentence affinities where α is a pre-specified constant.

2.3.2 Bandit Policy

Recall that an action contains a set of selected sentences in our framework. Hence the agent chooses one sentence every time until it believes all the selected sentences are good enough to cover the document contents. Selecting sentences with high affinities as summary could be an intuitive choice. However, some low-value sentences might also be important but are easily ignored. They should be explored to form the summary as well. Hence we adopt ϵ-greedy to extract one sentence every time until terminated. Specifically, the agent samples a sentence index a_t from the multinomial distribution with sentence affinities $\{\pi_1, \pi_2, \ldots, \pi_N\}$ as probabilities, where $a_t \in [1, N]$ and N is the document length. The new sentence S_{a_t} will be appended in the summary with a probability of $1 - \epsilon$. Otherwise the agent randomly picks one sentence with a probability of ϵ as an exploration. ϵ-greedy policy could raise the probability of low-value sentences to be extracted, and each extractive step in our method is a repeated sampling-without-replacement.

2.3.3 Termination Mechanism

In HER, we propose a termination mechanism that is independent on future rewards to make our model flexible in extracting summary with various numbers of sentences. This mechanism relies on the distribution of sentence affinities $\pi_{1:N}$.

$$r \sim Bernoulli(1, \max(\hat{p}, \max(\pi_{1:N}))), \tag{5}$$

$$\hat{p} = \frac{\max(\pi_{1:N}) - \min(\pi_{1:N})}{\max(\pi_{1:N})}, \tag{6}$$

where $r \in \{0, 1\}$ is sampled from the binomial distribution and $r = 0$ terminates the sentence extraction. With this mechanism, the agent will stop extraction with high probability as long as the differences among affinities are small enough or all the sentence affinities are very low.

2.4 Training

After the agent takes an action a, we can derive an summary induced by a out of a document D. Then the agent would receive a reward $R(a; G)$ where G is the gold-standard summary of D. $R(a; G)$ is computed by the average of three variants of ROUGE (Lin, 2004). To balance precision and recall, we use F-score here,

$$\begin{aligned} R(a; G) = \frac{1}{3}(&\text{ROUGE-1}_f(a; G) \\ +&\text{ROUGE-2}_f(a; G) + \text{ROUGE-L}_f(a; G)). \end{aligned} \tag{7}$$

We represent the whole extractive neural network as $p_\theta(\cdot|D)$ containing the encoder in rough reading and the decoder in careful reading. The goal of our model is to find parameters θ of p_θ to produce high-quality summary and maximize the rewards (c.f. Eq. (8)). But we cannot obtain gradient to maximize Eq. (8) with gradient ascent as it is discretely sampled. So we use the likelihood ratio gradient estimator from reinforcement learning and stochastic optimization (Williams, 1992; Sutton et al., 2000) to acquire the gradient by Eq. (9).

We use $Q(D)$ in Eq. (10) to construct $p_\theta(a|D)$ following Dong et al. (2018), where $z(D) =$

$\sum_t \pi(D)_t$ and ϵ is the exploration probability of the ϵ-greedy denoted in Sec. 2.3.2. M is the number of extracted sentences this is determined jointly by the termination mechanism and the document context. $Q(D)^{\frac{1}{M}}$ is adopted to present $p_\theta(a|D)$ to avoid extracting fewer or more sentences when maximizing the objective function. Hence, $\bigtriangledown_\theta \log p_\theta(a|D)$ in Eq. (9) can be easily computed.

$$J(\theta) = E[R(a;G)] \tag{8}$$

$$\bigtriangledown_\theta J(\theta) = E[\bigtriangledown_\theta \log p_\theta(a|D) R(a;G)] \tag{9}$$

$$Q(D) = \prod_{j=1}^{M} \left(\frac{\epsilon}{N-j+1} + \frac{(1-\epsilon)\pi(D)_{a_j}}{z(D) - \sum_{k=1}^{j-1} \pi(D)_{a_k}}\right) \tag{10}$$

However, the exact document distribution is unknown and we cannot evaluate the expected value in Eq. (9). So we use sampling to estimate it instead. Given a document-summary pair (D, G), our HER samples B summaries induced by $a^1, \ldots, a^B$ from $p_\theta(\cdot|D)$ and obtain all the gradients, then the average value is considered as the estimation. As sample-based gradient estimate may have high variance, we use a baseline for variance reduction. The gradient of the objective function is finally represented as,

$$\bigtriangledown_\theta J(\theta) \approx \frac{1}{B} \sum_{b=1}^{B} \bigtriangledown_\theta \log p_\theta(a^b|D)(R(a^b, G) - \bar{r}) \tag{11}$$

where we choose self-critical reinforcement learning to acquire the baseline $\bar{r}$ following Ranzato et al. (2015); Rennie et al. (2017); Paulus et al. (2017); Dong et al. (2018) computed by greedy encoding $\bar{r} = R(a_{greedy}; G)$. More concretely, $a_{greedy} = \text{argmax} p_\theta(a|D)$ and this baseline satisfies that the probability of a sampled sequence would be increased when the summary it induces is better than what is obtained by greedy decoding. The procedure of HER is shown in Algorithm 1.

Algorithm 1 Procedure of HER

Input A randomly sampled (D, G) pair
$S_{1:N}, D_0, L_{1:K} = \text{Encoder}(D)$ ▷ Rough Reading
$\pi_{1:N} = \text{Decoder}(S_{1:N}, D_0, L_{1:K})$ ▷ Careful Reading
Initialize $loss_{1:B} = 0, sumloss = 0$
for $c = 1, \ldots, B$ **do**
  $t \leftarrow 0$
  while $t \leq |D|$ and NotTerminated **do**
    if $random() < \epsilon$ **then**
      randomly select a_t
    else
      $a_t \sim P_r(N, \pi_{1:N})$
    $loss_b \mathrel{*}= \left(\frac{\epsilon}{N-t+1} + \frac{(1-\epsilon)\pi_{a_t}}{\sum_j \pi_j - \sum_{k=1}^{t-1} \pi_{a_k}}\right)$
    $t \leftarrow t + 1$
  $loss_b \leftarrow \log(loss_b)/t$
  Compute reward $R_b = R(a; G)$
$\bar{r} = R(a_{greedy}; G)$ ▷ Training
for $b = 1, \ldots, B$ **do**
  $sumloss \mathrel{+}= (loss_b * \frac{\bar{r} - R_b}{R_b})$
$sumloss \leftarrow sumloss/B$
$\theta \leftarrow \theta + \lambda \bigtriangledown_{sumloss}$

3 Experiment Settings

In this section we present our experimental setup for evaluating the performance of the proposed HER, including the datasets, evaluation protocol, baselines and implementation details.

Datasets We evaluated our models on three datasets: the CNN, the DailyMail and the combined CNN/DailyMail (Hermann et al., 2015; Nallapati et al., 2016) and use the standard splits of Hermann et al. (2015) for training, validation, and testing (90,266/1,220/1,093 documents for CNN and 196,961/12,148/10,397 for DailyMail) with the same setting in See et al. (2017).

Evaluation We evaluate summarization quality using F_1 ROUGE (Lin, 2004) including unigram and bigram overlap (ROUGE-1 and ROUGE-2) to assess informativeness and the longest common subsequence (ROUGE-L) to assess fluency with the reference summaries. We obtain ROUGE scores using a faster python implementation[2] for training and evaluation, and the standard pyrouge package[3] for testing following Dong et al. (2018).

Baselines We compare our proposed HER against four kinds of extractive methods: (1) Lead-3 model simply selects the first three sentences. (2) NN-SE (Cheng and Lapata, 2016) and SummaRuNNer (Nallapati et al., 2017) are sequence labeling task and trained with cross-entropy loss. (3) Refresh (Narayan et al., 2018), DQN (Yao et al., 2018) and RNES (Wu and Hu, 2018) extract summary via reinforcement learning. (4) BANDITSUM (Dong et al., 2018) considers the task as a contextual bandit but fails to simulate human reading recognition process.

Implementation Details We initialize word embeddings with 100-dimension Glove embeddings (Pennington et al., 2014). In rough reading,

[2] https://github.com/pltrdy/rouge

[3] Pyrouge is a Python package. We compute all ROUGE scores with parameters "-a -c 95 -m -n 4 -w 1.2." Refer to https://pypi.python.org/pypi/pyrouge/0.1.3

the encoder is hierarchical and each layer is a two-stacked BiLSTM with a hidden size of 200. Therefore, sentence vectors and the document representation D_0 have a dimension of 400. For the variant CNN, we adopt filter windows H in $\{1, 2, 3\}$ with 100 feature maps each and generate $K = 3$ local representations for each document. In careful reading, we set $\alpha = 0.5$ for sentence decoder and $\epsilon = 0.1$ for bandit policy. We also bound the minimum and maximum number of selected sentence to be 1 and 10 for terminal mechanism. During training, we use the optimizer Adam (Kingma and Ba, 2014) with a learning rate of 10^{-5}, beta parameters as $(0, 0.999)$ and a weight decay of 10^{-6} to maximize the objective function following Dong et al. (2018). We employ gradient clipping of 1 for regularization and sample $B = 20$ times for each document. We train our model within two epochs. During the test, we choose the whole document as the extractive summary if its length is less than 3 sentences since the local features cannot be obtained through the CNN-based network.

4 Experimental Results

4.1 Quantitative Analysis

We first report the ROUGE metrics on the combined CNN/DailyMail test sets in Table 1 and the separate results in Table 2. We can get several observations from the tables.

Firstly, our model generally performs the best and even surpasses 42 on ROUGE-1 score on the combined CNN/DailyMail dataset. It also shows better results on the separate datasets. We argue that global and local features from rough reading can help extract summaries by capturing deep contextual relations, and the designed structure in careful reading makes it more flexible in selecting sentence sets. Hence a two-stage framework based on the human's reading cognition is more appropriate for extractive summarization.

Secondly, directly optimizing the evaluation metric by combining cross-entropy loss with rewards may improve the extractive results. RL-based methods, Refresh (Narayan et al., 2018) and RNES (Wu and Hu, 2018), perform better than the sequence labeling methods like SummaRuNNer (Nallapati et al., 2017). BANDITSUM (Dong et al., 2018) generally performs better than the other baselines, and it reports that framing the extractive summarization based on contextual bandit is more suitable than sequential labeling setting and has more search space than other RL-based methods (Narayan et al., 2018; Yao et al., 2018; Wu and Hu, 2018).

Model	**ROUGE**		
	R1	R2	RL
Lead-3	40.0	17.5	36.2
SummaRuNNer	39.6	16.2	35.3
DQN	39.4	16.1	35.6
Refresh	40.0	18.2	36.6
RNES	41.3	**18.9**	37.6
BANDITSUM	41.5	18.7	37.6
HER	**42.3**	**18.9**	**37.9**

Table 1: Results on the combined CNN/DailyMail test sets. We report F1 scores of ROUGE-1 (R1), ROUGE-2 (R2), and ROUGE-L (RL). The result of Lead-3 is taken from Dong et al. (2018).

Model	CNN			DailyMail		
	R1	R2	RL	R1	R2	RL
Lead-3	28.8	11.0	25.5	41.2	18.2	37.3
NN-SE	28.4	10.0	25.0	36.2	15.2	32.9
Refresh	30.4	11.7	26.9	41.0	18.8	37.7
BANDITSUM	**30.7**	**11.6**	27.4	42.1	18.9	38.3
HER	**30.7**	11.5	**27.5**	**42.7**	**19.0**	**38.5**

Table 2: Results of the test sets on the CNN and DailyMail datasets separately.

4.2 Ablation Test

Next, we conduct ablation test by removing the modules of the proposed HER step by step. Firstly, we replace the automatic termination mechanism with a fixed extracting strategy that always selects three sentences for every document and present the model as HER-3. Based on HER-3, we also remove bandit policy, local net, general net gradually, and denote them as HER-3 w/o policy, HER-3 w/o policy & local net and HER-3 w/o policy & rough reading. The results are reported in Table 3 and it proves the effectiveness of each proposed module. Firstly, HER constructed with an automatic termination mechanism is more flexible and reliable in extracting various numbers of sentences varying different documents. Secondly, HER use ϵ-greedy to select sentences in order to raise the exploration chances on discovering important but easily ignored information. Thirdly, general cognition from rough reading process is useful in extractive summarizarion.

4.3 A Closer Look

To verify whether our proposed model HER can simulate human beings' reading cognitive process, and whether such simulation are inherently helpful on extractive summarization, we conduct extensive evaluations that try to answer the following three questions.

Model	ROUGE		
	R1	R2	RL
HER	**42.3**	**18.9**	**37.9**
HER-3	42.0	18.5	37.6
HER-3 w/o policy	41.7	18.3	37.1
HER-3 w/o policy&L	41.2	18.4	37.0
HER-3 w/o policy&F	40.6	18.2	36.9

Table 3: The results of ablation test on the test split. L and F are short for local net and rough reading.

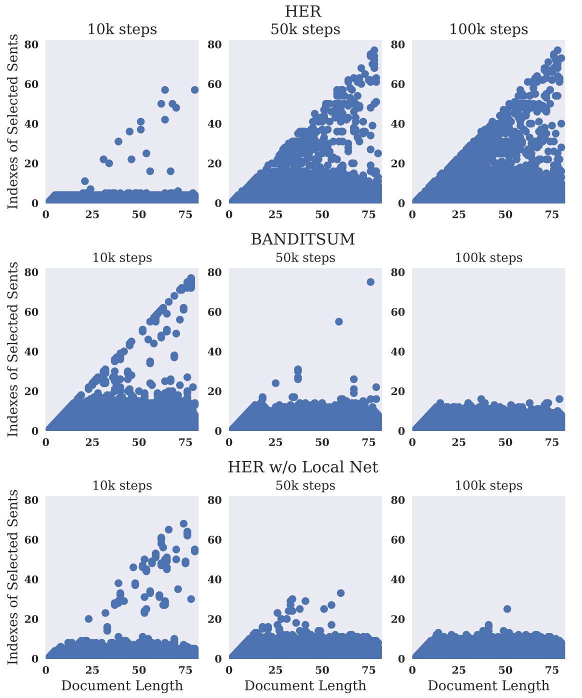

Figure 3: The statistics of model HER, BANDITSUMT (Dong et al., 2018), HER w/o Local Net on the selected sentences' indexes varying different document lengths. This is reported on the documents the length of which is less than 80 of the test split.

(1) Can CNN-based network extract local features of different paragraphs?

In Figure 3, we report the distribution of selected sentences' positions on our proposed model HER, BANDITSUM and HER w/o Local Net. Each model is shown on testing after training 10k, 50k, 100k steps. We observe that all the three models can focus on different parts of the context to form summary at first and BANDITSUM performs the best after training 10k steps. However, with training steps growing, BANDITSUM and HER w/o Local begin to prefer earlier sentences. HER, on the other hand, can focus on various paragraphs and extract information from different parts of the texts with constant training. The contextual bandit (CB) based frameworks seems to be able to attend on various parts of the contexts to some

Index	Sentence	Affinity	HER	HER w/o policy
2	Two spotted leopards, two Macaque monkeys and a brown bear will be returned to Marian Thompson…	0.873	yes	yes
3	He set off a wide scare in October when he released 50 potentially dangerous animals from his farm before shooting himself.	0.872	yes	yes
4	Of the 50 animals Thompson released, 48 were killed by law enforcement, while two primates were killed by the other animals, zoo officials said.	0.767	**no**	**yes**
13	State officials have no legal power to inspect the cages before the animals are returned…	0.297	**yes**	**no**

Figure 4: A case on sentence selection of HER and HER w/o policy. The article is from CNN dataset. The highlighted indices indicate the corresponding sentences should be extracted as summary.

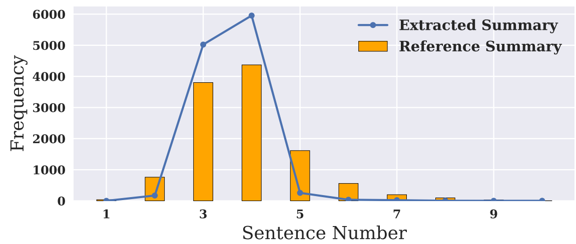

Figure 5: The statistics on extracted sentence number of our model. Frequency is the number of documents.

degree in the beginning. However, with constant training, both BANDITSUM and HER w/o Local start focusing on earlier sentences since the nature that sentences similar to the main idea usually lie on the head of the text. As our proposed HER is equipped with a variant CNN that extracts local features, our model can focus on gist of paragraphs rather than only the first several sentences, which also encourages the exploration on extracting information from various positions more flexibly.

(2) Can the proposed bandit policy discover low-score but easily ignored information?

To answer this question, we demonstrate a detailed case on sentence selection in Figure 4. We observe that although the 4th sentence has a high affinity, it should not be included in the summary since its meaning is close to the 3rd sentence which has already been extracted. Instead, the 13th sentence is supposed to be selected though it has low affinity. Since our HER adopts the ϵ-greedy policy, it can explore such sentence and extract it out correctly.

(3) Can HER extract varied but proper numbers of sentences?

We answer this question by drawing the frequency distribution of extracted sentence numbers by our model on the test set of combined CNN/DailyMail, and Figure 5 exhibits the results. We observe that the frequency distribution of extracted sentence number is basically similar to

Model	Overall	Coverage	Non-Redundancy
HER w/o Dec_2	2.88	2.81	2.81
HER w/o L	3.02	2.96	2.75
BANDITSUM	2.06	2.07	**1.91**
HER	**1.81**	**1.75**	1.97

Table 4: Average rank of human evaluation in terms of overall performance, coverage, and non-redundancy. L is short for local Net. Lower score is better.

that of the gold-standard summary. Compared with BANDITSUM which extracts fixed number of sentences, our model shows more flexibility and extensibility on extractive summarization.

4.4 Human Evaluation

Lastly, we conduct a qualitative evaluation. Following Wu and Hu (2018), we randomly sample 50 documents from the test set on the combined CNN/DailyMail dataset and ask three volunteers to evaluate the summaries extracted by HER w/o Dec_2, HER w/o Local Net, BANDITSUM and HER, respectively. HER w/o Dec_2 only uses Eq. (3) to compute sentence affinities without inhibiting redundant sentences. For each document-summary pair, they are asked to rank the output of each system on three aspects, namely overall quality, coverage and non-redundancy. Notice that the best one will be marked rank 1 and so on, and two system would be ranked the same if their extracted summaries are identical. We report the average results in Table 4 and it shows that our HER is leading than BANDITSUM on overall quality and coverage. Additionally, HER w/o Dec_2 performs the worst on non-redundancy as it does not specialize these unselected sentences which are similar to already extracted ones. Furthermore, HER w/o Local Net takes on bad performance on coverage because the local features can focus on paragraphical messages and help to refine thorough information.

5 Related Work

Extractive Text Summarization Researchers have developed many statistical methods for automatic extractive summarization. Traditional methods learn to score each sentence dependently (Erkan and Radev, 2004; Mihalcea and Tarau, 2004; Wong et al., 2008). Recently neural network based extractive methods (Cheng and Lapata, 2016; Nallapati et al., 2017; Feng et al., 2018; Shi et al., 2018) usually consider extractive summarization as sequence labeling tasks and aim to minimize the cross-entropy objective function. Narayan et al. (2017) utilizes side information to help sentence classifier while Yasunaga et al. (2017) computes the salience of each sentence for selection with graph convolutional networks. In addition, reinforcement learning based methods (Wu and Hu, 2018; Narayan et al., 2018; Yao et al., 2018) have been proposed to directly optimize the evaluation metric ROUGE by combining cross-entropy loss with rewards from policy gradient reinforcement learning. Dong et al. (2018) considered extractive summarization as a contextual bandit and it performs well especially when good summary sentences appear late in the source document. Recently, Nallapati et al. (2017); Chen and Bansal (2018); Hsu et al. (2018) propose unified models and combine the advantages of both extractive and abstractive methods.

Human Reading-inspired Strategy in NLP Recently, several pioneer researches began to study how to adapt human reading cognition process, usually including pre-reading, reading and post-reading (Avery and Graves, 1997; Saricoban, 2002; Toprak and Almacıoğlu, 2009; Pressley and Afflerbach, 2012), into various NLP-related applications. For example, Li et al. (2018) solved document-based question answering and by simulating human being's reading strategy. Yang et al. (2019) applied it for abstractive summarization, Zheng et al. (2019) simulated human behavior for reading comprehension, and Lei et al. (2019) utilized human-like semantic cognition for aspect-level sentiment classification. In this paper, we attempt to perform extractive summarization under the inspiration of human reading recognition.

6 Conclusion

Inspired by the reading cognition of human beings, we propose HER, a two-stage method, to mimic how people extract summaries. The whole learning process is formulated as a contextual bandit trained with policy gradient reinforcement learning. In rough reading, two neural networks are taken to encode coarse-grained information. In careful reading, repeatedly reading are conducted to select fine-grained sentences as summary. Experiments on two real-world datasets demonstrate that our proposed model can significantly outperform the state-of-the-art extractive methods on summary quality, coverage and non-redundancy.

References

Patricia G Avery and Michael F Graves. 1997. Scaffolding young learners' reading of social studies texts. *Social Studies and the Young Learner*, 9(4):10–14.

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.

Jianpeng Cheng and Mirella Lapata. 2016. Neural summarization by extracting sentences and words. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, pages 484–494.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Yue Dong, Yikang Shen, Eric Crawford, Herke van Hoof, and Jackie Chi Kit Cheung. 2018. Banditsum: Extractive summarization as a contextual bandit. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3739–3748.

Günes Erkan and Dragomir R Radev. 2004. Lexrank: Graph-based lexical centrality as salience in text summarization. *Journal of artificial intelligence research*, 22:457–479.

Chong Feng, Fei Cai, Honghui Chen, and Maarten de Rijke. 2018. Attentive encoder-based extractive text summarization. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 1499–1502. ACM.

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.

Wan-Ting Hsu, Chieh-Kai Lin, Ming-Ying Lee, Kerui Min, Jing Tang, and Min Sun. 2018. A unified model for extractive and abstractive summarization using inconsistency loss. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 132–141.

Yoon Kim. 2014. Convolutional neural networks for sentence classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1746–1751.

Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization.

Ryan Kiros, Yukun Zhu, Ruslan R Salakhutdinov, Richard Zemel, Raquel Urtasun, Antonio Torralba, and Sanja Fidler. 2015. Skip-thought vectors. In *Advances in neural information processing systems*, pages 3294–3302.

Zeyang Lei, Yujiu Yang, Min Yang, Wei Zhao, Jun Guo, and Yi Liu. 2019. A human-like semantic cognition network for aspect-level sentiment classification. In *AAAI*.

Weikang Li, Wei Li, and Yunfang Wu. 2018. A unified model for document-based question answering based on human-like reading strategy. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

Chin-Yew Lin. 2004. Rouge: A package for automatic evaluation of summaries. *Text Summarization Branches Out*.

Rada Mihalcea and Paul Tarau. 2004. Textrank: Bringing order into text. In *Proceedings of the 2004 conference on empirical methods in natural language processing*.

Ramesh Nallapati, Feifei Zhai, and Bowen Zhou. 2017. Summarunner: A recurrent neural network based sequence model for extractive summarization of documents. In *AAAI*, pages 3075–3081.

Ramesh Nallapati, Bowen Zhou, Cicero dos Santos, Caglar Gulcehre, and Bing Xiang. 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.

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.

Shashi Narayan, Nikos Papasarantopoulos, Shay B Cohen, and Mirella Lapata. 2017. Neural extractive summarization with side information. *arXiv preprint arXiv:1704.04530*.

Ani Nenkova, Kathleen McKeown, et al. 2011. Automatic summarization. *Foundations and Trends® in Information Retrieval*, 5(2–3):103–233.

Romain Paulus, Caiming Xiong, and Richard Socher. 2017. A deep reinforced model for abstractive summarization. *arXiv preprint arXiv:1705.04304*.

Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543.

Michael Pressley and Peter Afflerbach. 2012. *Verbal protocols of reading: The nature of constructively responsive reading*. Routledge.

Marc'Aurelio Ranzato, Sumit Chopra, Michael Auli, and Wojciech Zaremba. 2015. Sequence level training with recurrent neural networks. *arXiv preprint arXiv:1511.06732*.

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.

Arif Saricoban. 2002. Reading strategies of successful readers through the three phase approach. *The Reading Matrix*, 2(3).

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.

Jiaxin Shi, Chen Liang, Lei Hou, Juanzi Li, Zhiyuan Liu, and Hanwang Zhang. 2018. Deepchannel: Salience estimation by contrastive learning for extractive document summarization. *arXiv preprint arXiv:1811.02394*.

Richard S Sutton, David A McAllester, Satinder P Singh, and Yishay Mansour. 2000. Policy gradient methods for reinforcement learning with function approximation. In *Advances in neural information processing systems*, pages 1057–1063.

Elif Leyla Toprak and Gamze Almacıoğlu. 2009. Three reading phases and their applications in the teaching of english as a foreign language in reading classes with young learners. *Journal of language and Linguistic Studies*, 5(1):20–36.

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

Kam-Fai Wong, Mingli Wu, and Wenjie Li. 2008. Extractive summarization using supervised and semi-supervised learning. In *Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1*, pages 985–992. Association for Computational Linguistics.

Yuxiang Wu and Baotian Hu. 2018. Learning to extract coherent summary via deep reinforcement learning. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

Min Yang, Qiang Qu, Wenting Tu, Ying Shen, Zhou Zhao, and Xiaojun Chen. 2019. Exploring human-like reading strategy for abstractive text summarization. In *AAAI*.

Kaichun Yao, Libo Zhang, Tiejian Luo, and Yanjun Wu. 2018. Deep reinforcement learning for extractive document summarization. *Neurocomputing*, 284:52–62.

Michihiro Yasunaga, Rui Zhang, Kshitijh Meelu, Ayush Pareek, Krishnan Srinivasan, and Dragomir Radev. 2017. Graph-based neural multi-document summarization. In *Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017)*, pages 452–462.

Yukun Zheng, Jiaxin Mao, Yiqun Liu, Zixin Ye, Min Zhang, and Shaoping Ma. 2019. Human behavior inspired machine reading comprehension. In *SIGIR*.