A Hierarchical End-to-End Model for Jointly Improving Text Summarization and Sentiment Classification

Shuming Ma, Xu Sun, Junyang Lin, Xuancheng Ren

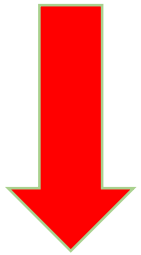
Peking University, Beijing, China

Introduction

- **Text summarization** aims at generating a summary with the major points of the original text.
- **Sentiment classification** can be regarded as a special type of summarization which ''summarizes'' the text into a even more abstract fashion, i.e., a sentiment class.
- Both text summarization and sentiment classification aim at mining **the main ideas** of the text.

Introduction

- **Text summarization** aims at generating a summary with the major points of the original text.
- **Sentiment classification** can be regarded as a special type of summarization which "summarizes" the text into a even more abstract fashion, i.e., a sentiment class.
- Both text summarization and sentiment classification aim at mining **the main ideas** of the text.



Joint learning of text summarization and sentiment classification

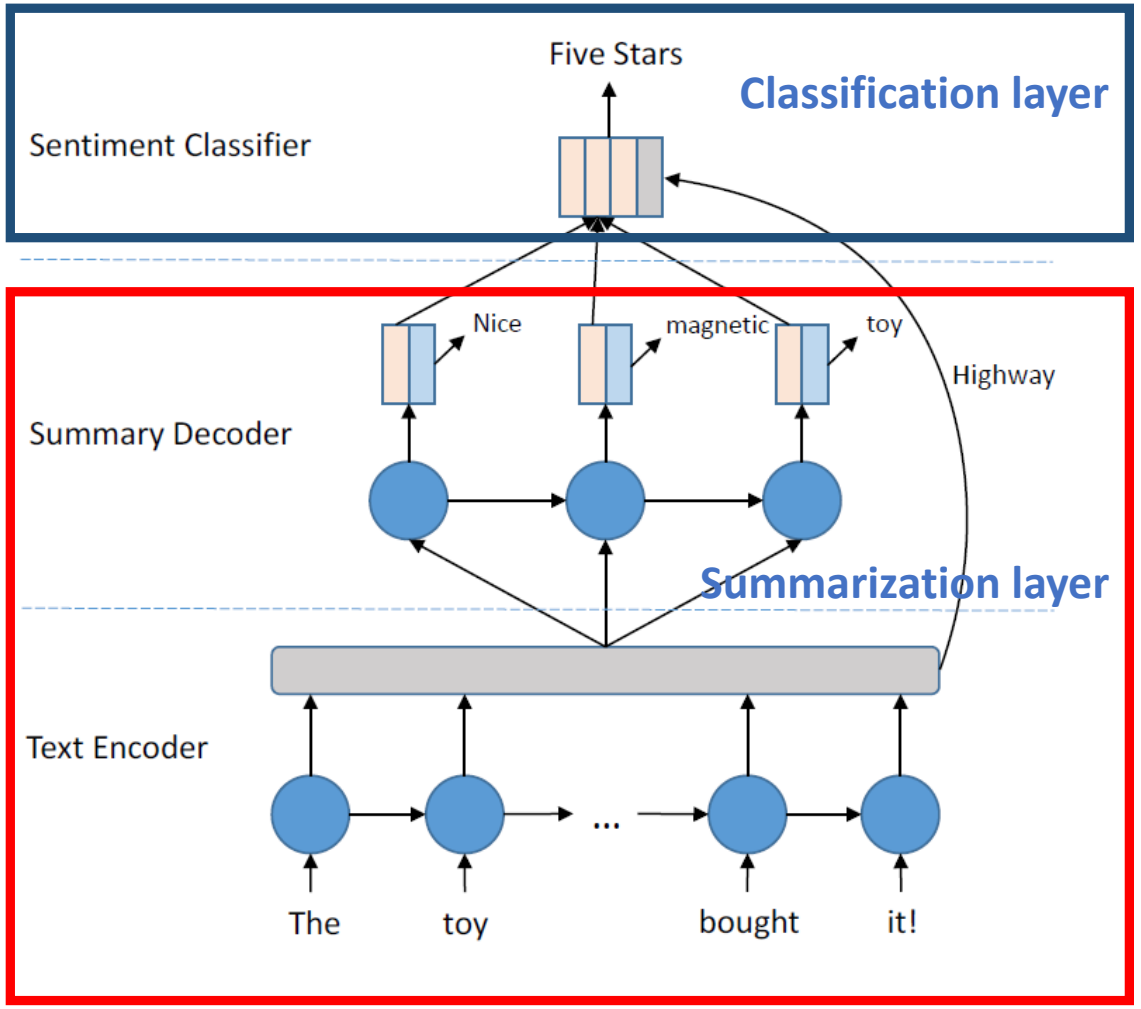
Model Overview

Summarization Layer:

- Compresses the original text into short sentences
- Point out the important and informative words
- Remove the redundant and misleading information

Sentiment Classification Layer:

- Further ``summarizes'' the texts into a sentiment class
- Provide a more significant supervision signal for text summarization
- Guides the summarization component to capture the sentiment tendency



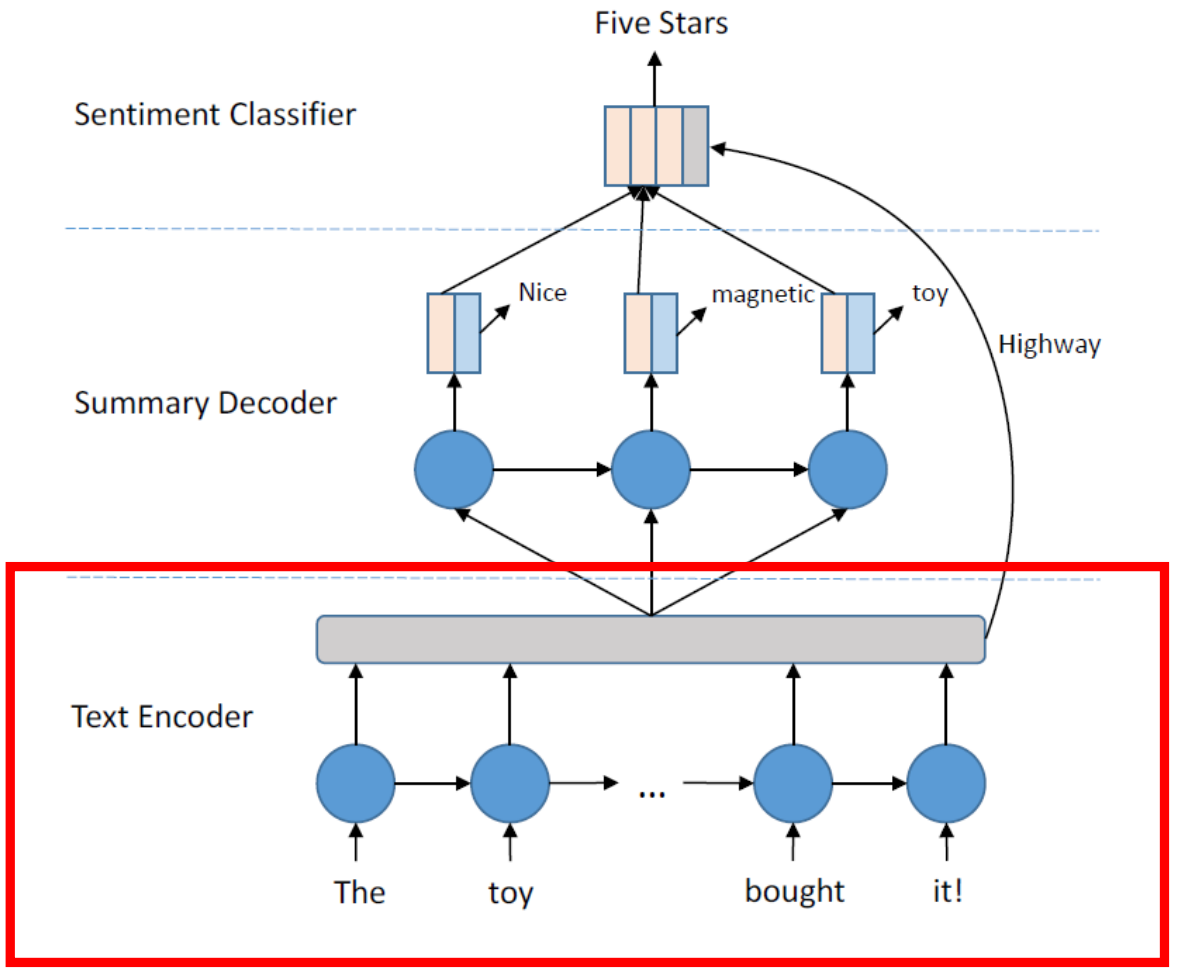
Text Encoder

BiLSTM Encoder:

$$\vec{h}_t = \vec{f}(x_t, \vec{h}_{t-1})$$

$$\overleftarrow{h}_t = \overleftarrow{f}(x_t, \overleftarrow{h}_{t+1})$$

$$h_t = \vec{h}_t + \overleftarrow{h}_t$$



Summary Decoder with Multi-View Attention

Uni-directional LSTM:

$$s_t = f(y_{t-1}, s_{t-1})$$

Multi-view Attention:

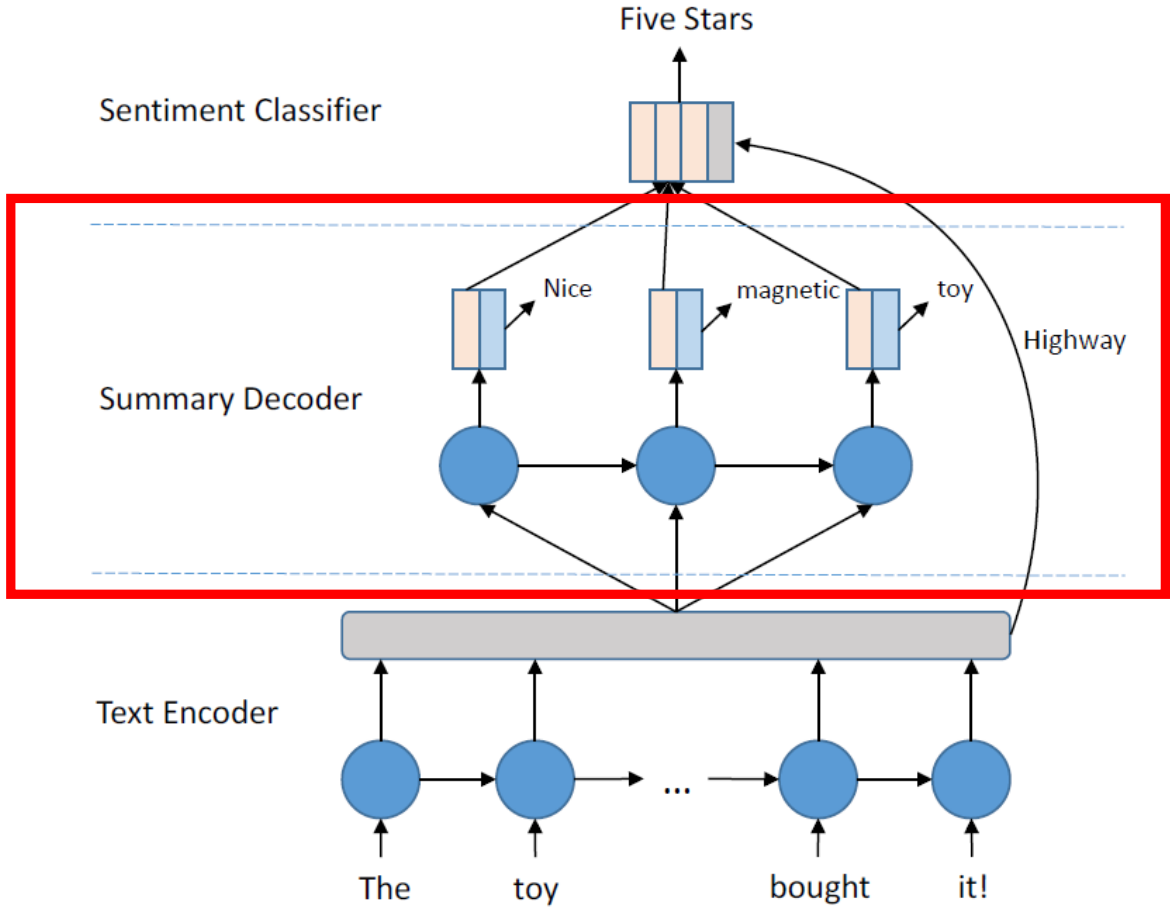
$$v_t^{(c)} = \sum_{i=1}^{N} \alpha_{ti} h_i$$

$$\alpha_{ti} = \frac{e^{g(s_t, h_i)}}{\sum_{j=1}^{N} e^{g(s_t, h_j)}}$$

$$g(s_t, h_i) = \tanh\left(s_t^T W_t h_i\right)$$

Word Generator:

$$p(y_t|x) = softmax(W_g v_t^{(c)} + b_g)$$



Summary-Aware Sentiment Classifier

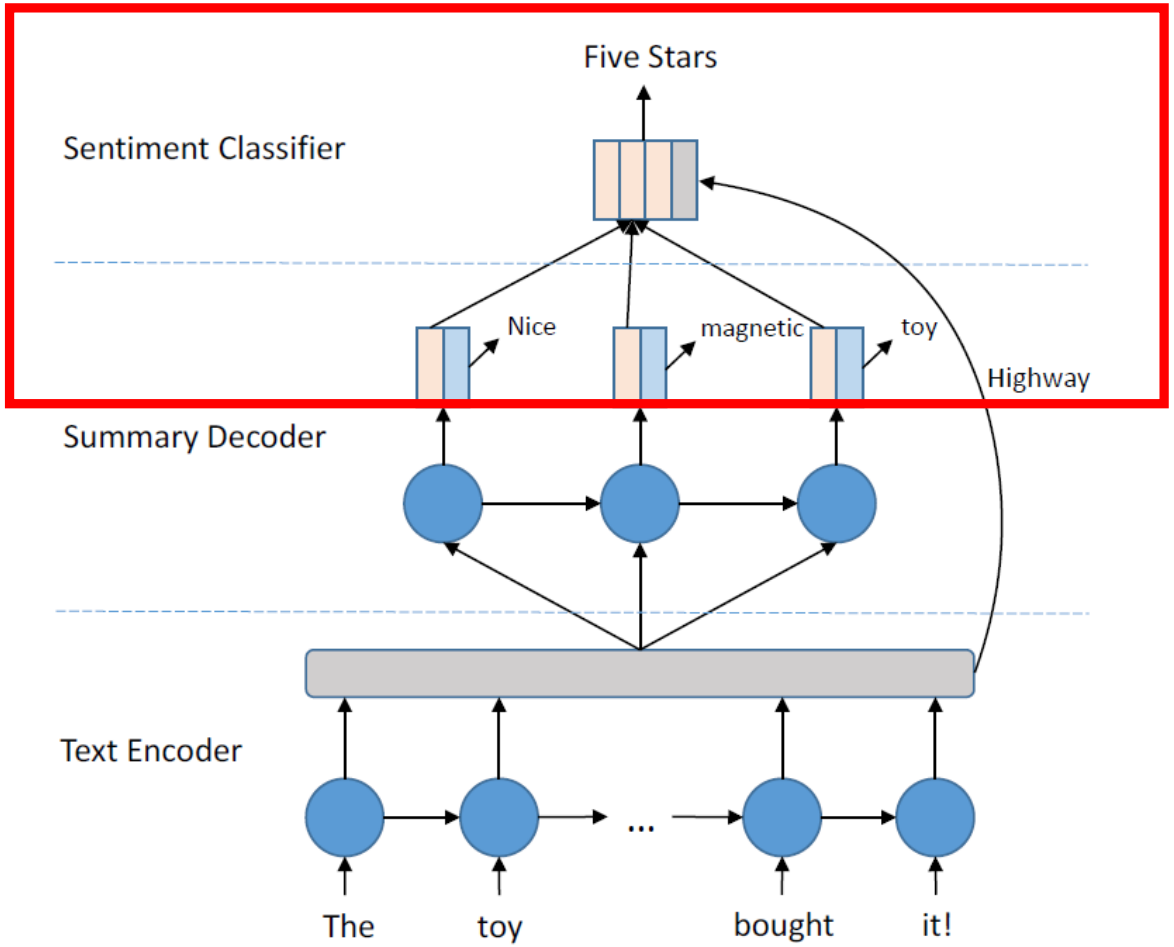
Highway and Max-pooling:

$$\boldsymbol{v^{(t)}} = [v_1^{(t)}, v_2^{(t)}, ..., v_M^{(t)}]$$

$$\boldsymbol{r} = \max(\boldsymbol{v^{(t)}} \oplus \boldsymbol{h})$$
$$= \max([v_1^{(t)}, v_2^{(t)}, ..., v_M^{(t)}, h_1, h_2, ..., h_L])$$

Classfier:

$$p(l|x) = softmax(W_c\boldsymbol{r} + b_c)$$



Experiments

Dataset

Amazon SNAP Review Dataset (SNAP):

- This dataset is part of Stanford Network Analysis Project (SNAP), and is provided by He and McAuley [2016].
- The dataset consists of reviews from amazon, and contains product reviews and metadata from Amazon, including 142.8 million reviews spanning May 1996 - July 2014.
- We select three domains of product reviews: **Toys & Games, Sports & Outdoors, and Movie & TV.**

Evaluation Metrics

- For **abstractive summarization**, our evaluation metric is ROUGE score, which is popular for summarization evaluation. We use **ROUGE-1** (unigram), **ROUGE-2** (bi-gram) and **ROUGE-L** (LCS) as the evaluation metrics in the reported experimental results.

- For **sentiment classification**, the evaluation metric is per-label **accuracy**. We evaluate the accuracy of both five-class sentiment, of which the sentiment is classified into 5 class, and two-class sentiment, of which the sentiment is either positive or negative.

Result of Text Summarization

Toys & Games	RG-1	RG-2	RG-L
S2S [Hu *et al.*, 2015]	14.05	2.47	15.75
S2S-att [Hu *et al.*, 2015]	16.23	4.27	16.01
S2S-att + BiLSTM	16.32	4.43	16.27
HSSC (this work)	**18.44**	**5.00**	**17.69**

Sports & Outdoors	RG-1	RG-2	RG-L
S2S [Hu *et al.*, 2015]	13.38	2.59	13.18
S2S-att [Hu *et al.*, 2015]	15.70	3.61	15.53
S2S-att + BiLSTM	15.75	3.64	15.68
HSSC (this work)	**17.85**	**4.77**	**17.59**

Movie & TV	RG-1	RG-2	RG-L
S2S [Hu *et al.*, 2015]	10.98	2.34	10.77
S2S-att [Hu *et al.*, 2015]	12.17	3.08	11.77
S2S-att + BiLSTM	12.33	3.22	11.92
HSSC (this work)	**14.52**	**4.84**	**13.42**

Table 1: Comparison between our model and the sequence-to-sequence baseline for abstractive summarization on the Amazon SNAP test sets. The test sets include three domains: Toys & Gamse, Sports & Outdoors, and Movie & TV. RG-1, RG-2, and RG-L denote ROUGE-1, ROUGE-2, and ROUGE-L, respectively.

Result of Sentiment Classification

Toys & Games	5-class	2-class
CNN	70.5	90.2
BiLSTM	70.7	90.9
BiLSTM + S2S-att	70.9	90.9
HSSC (this work)	**71.9**	**91.8**

Sports & Outdoors	5-class	2-class
CNN	72.0	91.5
BiLSTM	71.9	91.6
BiLSTM + S2S-att	72.1	91.9
HSSC (this work)	**73.2**	**92.1**

Movie & TV	5-class	2-class
CNN	66.9	86.0
BiLSTM	67.8	86.2
BiLSTM + S2S-att	68.0	86.6
HSSC (this work)	**68.9**	**88.4**

Table 2: Comparison between our model and the sequence-to-sequence baselines for sentiment classification on the Amazon SNAP test sets. The test sets include three domains: Toys & Gamse, Sports & Outdoors, and Movie & TV. 5-class and 2-class denote the accuracy of five-class sentiment and two-class sentiment, respectively.

Ablation Study

Toys & Games	5-class	RG-L
w/o Multi-View	70.9	16.47
w/o Highway	70.1	16.06
HSSC (Full Model)	71.9	17.69

Sports & Outdoors	5-class	RG-L
w/o Multi-View	72.0	16.36
w/o Highway	71.5	15.73
HSSC (Full Model)	73.2	17.59

Movie & TV	5-class	RG-L
w/o Multi-View	68.1	12.34
w/o Highway	67.7	12.01
HSSC (Full Model)	68.9	13.42

Table 3: Ablation study. 5-class denotes the accuracy of five-grained sentiment, and RG-L denotes ROUGE-L for summarization.

Visualization of Multi-View Attention

(1)	i saw this movie 11 times in the theater and i think that this is one of the best movies ever made and the best movie made about christ and his passion . god bless all those responsible for the creation of this powerful film .
(2)	my daughter , who is now 8 years old , received this as a christmas gift when she was 2 . it has been ready many times , and since been passed along to my son who is now 4 . my children enjoy the tactile quality of the monkeys faces . it is helpful learning counting when there is something they can feel . i have always enjoyed reading the sing song story . it does not take long to read , and after all these years i pretty much have it memorized . a great book , very fun .
(3)	this mattress is too narrow to be comfortable . you fit on it fine but because of the air , i found that it was a balancing act to switch positions . i tried more and less air to no effect . i think if you sleep on your back and stay in that position it would be fine but unfortunately that is not how i sleep . the strong vinyl smell does go away after airing out though .

(a) Sentiment view of the original text.

Visualization of Multi-View Attention

(1)	i saw this movie 11 times in the theater and i think that this is one of the best movies ever made and the best movie made about christ and his passion . god bless all those responsible for the creation of this powerful film .
(2)	my daughter , who is now 8 years old , received this as a christmas gift when she was 2 . it has been ready many times , and since been passed along to my son who is now 4 . my children enjoy the tactile quality of the monkeys faces . it is helpful learning counting when there is something they can feel . i have always enjoyed reading the sing song story . it does not take long to read , and after all these years i pretty much have it memorized . a great book , very fun .
(3)	this mattress is too narrow to be comfortable . you fit on it fine but because of the air . i found that it was a balancing act to switch positions . i tried more and less air to no effect . i think if you sleep on your back and stay in that position it would be fine but unfortunately that is not how i sleep . the strong vinyl smell does go away after airing out though .

(b) Summary view of the original text.

Conclusion

- We propose a model to generate both the sentiment labels and the human-like summaries, hoping to summarize the opinions from the coarse-grained sentiment labels to the fine-grained word sequences.
- Experimental results show that our model achieves better performance than the baseline systems on both abstractive summarization and sentiment classification.

Thank you!

If you have any question, please feel free to contact me:
shumingma@pku.edu.cn