

Personalized Abstract Review Summarization Using Personalized Key Information-Guided Network



Nidhin S. Dharan and R. Gowtham

Abstract We are proposing a personalized summarization model, which generates an abstractive summary of a random review based on the preference of a specific user. The summary will account the user's preference on different aspects present in the review. We put forward a Personalized Key Information Guided Network (PKIGN) that pools both extractive and abstractive methods for summary generation. Specifically, keywords present in the review are extracted which are specific to that user, and these keywords are used as key information representation to guide the process of generating summaries. Additionally, Pointer-Guide mechanism is employed for obtaining long-term value for decoding. We evaluate our model on a new Trip-Advisor hotel review dataset, comprising of 140,874 reviews from 41,600 users. Combining the results from both human evaluation and quantitative analysis, it is seen that our model achieves better performance than existing models on personalized review summarization in case of hotel reviews.

Keywords Extractive and abstractive summarization · Encoded · PKIGN · Pointer-guide

1 Introduction

Review summarization has gained a great success owing to the introduction of models sequence-to-sequence [1], transformers [2] and their variants [3]. The main objective of review summarization is to create a condensed summarization of the single or multiple reviews. With the exponential growth of e-commerce websites, it has been widely researched (Fig. 1).

This paper discusses introducing the concept of personalization to review summarization, which has not been discussed extensively in previous research. A model has

N. S. Dharan (✉) · R. Gowtham

Department of Computer Science Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

R. Gowtham

e-mail: r_gowtham@cb.amrita.edu

Review : The location is good, the room is cool, the price is reasonable, but the breakfast is pricey. the check in and check out is pretty quick. The staff is friendly. Most of the stuff in the mini-bar is free. The room I stayed in has no view, but that's fine for me, as I don't spend a lot of time in the room	
User	Summary
Sachin	affordable price, room is nice
Rohit	Great location
Virat	friendly staff and nice room

Fig. 1 The main motivation behind personalized key information-guided network is that for the same review different users are likely to create different summaries according to their own preferences

been built on creating personalized summarization based on important words [4] which does not provide good results. In case of a review, different users may care about different aspects according to their personal preferences. In case of our dataset which consists of hotel reviews, different aspects have been identified as location, room, value, facility, service and food. User A may care about the aspects service and room more than price, while user B may care about price. Therefore, we are proposing a model which takes into account user-specific aspects while creating summaries.

Personalized review summarization has a wide range of application across all online consumer products, such as TripAdvisor and Zomato. Users write their reviews across all of these platforms, and one function is to provide summarizes of other reviews according to their preferences or aspects. Using classical models for summarization, every user will view the same summary for every review. Using our proposed model, for the same review, different users will be able to see different summaries, thus providing a personalized service to each and every customer.

To perform personalized review summarization, we propose personalized key information-guided network which is based on sequence-to-sequence and key information guide network. Our model has major updates done in two parts.

Firstly, we create a corpus of all the reviews and summaries from a user and find out the most common words used by each user. Each user will talk more about the aspect they care about. By this method, we find out which aspects each user cares about.

Secondly, for each review, we are extracting keywords using TextRank [5], and we are only keeping keywords containing using specific keywords. By these methods, these user-specific aspects will be given more importance while generating summary.

To validate our approach, we have taken a dataset from paper. With quantitative and human evaluation approaches, we present that our model has achieved better

results for personalized review summarization in the case of hotel reviews. Our contributions for this project are as follows:

- To the best of our knowledge, we are the first ones to propose a personalized key information-guided network by using user-specific aspect keywords from reviews for personalized review summarization.
- For evaluating our model, we have created a novel dataset Hotel Reviews.

2 Related Works

Abstractive summarization has been studied across numerous papers throughout many years. Abstractive text summarizations have been extensively used for review summarization. RNNs are mainly leveraged for most of the natural language processing tasks as a result of very promising results. After the introduction of Encoder Decoder models in neural machine translation [6], a neural attention encoder–decoder model with feed-forward networks was introduced for abstract summarization [7]. In this model, an attention mechanism is added to the framework by taking into account the context vectors in hidden state of the encoder that helps in decoding the target sequences and achieved state-of-the-art results for DUC-2004 and Gigaword, two sentence-level summarization datasets. The basic architecture of our model is inspired from the sequence-to-sequence model. The approach which is focused on the attention mechanism has been augmented with recurrent decoders [3], abstract meaning representations [8], hierarchical networks [9] and variational autoencoders [10] improved performance on the respective datasets. A segment-to-segment neural transduction model [11] for sequence-to-sequence framework. The model introduces a latent segmentation which determines correspondences between tokens of the input text and the output text. Experimentation performed on the proposed transduction model shows good results on the Gigaword dataset.

While the sequence-to-sequence model with attention was getting promising results, some problems still existed. In these models, for each time step, the decoder uses a fixed target vocabulary to the given probability distribution. These types of situations can lead to OOV word errors. One method of solving this is where the size of target vocabulary is increased, but this will result in increase of computational complexity needed to find out the Softmax function across all possible words in the target vocabulary. To solve this, a model which uses soft copy mechanism was introduced in PGN [12] model which uses pointer generation network, which is a hybrid network, which lets us both copy the words from keywords and also generate the words from target vocabulary. PGN was able to achieve state-of-the-art results on CNN/Daily Mail dataset. The soft copy mechanism, mentioned in the above architecture, is added to our model with personalized keyword.

Another drawback of the traditional method is that there is no way to filter out the secondary information. In the existing methods, all the information given at the encoder is passed on to the decoder state for generation without checking if they are useful or not. This can often lead to the model focusing on unimportant information

while generating summaries. SEASS [13] network uses a selective mechanism to control the information flow which highlights important information and release the burden of the decoder which performed significantly better on ROGUE score for English Gigaword, DUC 2004 and MSRATC test sets. Therefore, focusing on the keywords while summarizing will improve summarization, and this principle is used in our model. Other models use dual-encoding [14] propose an LSTM-CNN approach which create new sentences by searching more fine-grained fragments than sentences and semantic phrases. The dual-encoding approach is consisting of two main stages. The first stage extracts phrases from source sentences which is followed by a stage that generates text summaries using deep learning. Existing models which create abstractive summary based on the opinions and arguments [15] also fail to filter the personalized information. There have been other methods which summarize texts [16–18] and analyze product reviews [19] but fail in personalization. Our model is inspired from Guiding Generation for Abstractive Text Summarization Based on Key Information Guide Network [20] which used key information-guided mechanism and soft copy mechanism. This model was validated on CNN/Daily Mail dataset. The main difference in our model is that we have devised a way to extract personalized information from the users' previous reviews and use this information for key information-guided mechanism and soft copy mechanism.

Another downside is that the encoder-decoder methods do not provide better results for longer texts. Transformer model [2] was introduced using stacked self-attention and point-wise fully connected layers for both the encoder and decoder. This model achieved state-of-the-art results in WMT 2014 English–German dataset. The main advantage of this model is parallelizable and requiring significantly less time to train. We tried implementing a guided mechanism in the transformer, but it did not provide better results.

User-aware sequence network [21] is used for personalized review summarization. This model is based on S2S network with user-aware encoder and user-aware decoder, where selective mechanism is used in user-aware encoder to highlight user-specific information in the review. USN achieved state-of-the-art rogue for personalized summarization. The user-aware decoder identify the user writing style and uses a soft copy mechanism to obtain summaries. But our model differs from this user-aware sequence network as we use personalized keywords to guide our summarization model. We have devised a method to extract user preferences which is unique to our model.

3 Problem Formulation

Suppose that we have a corpus with N user-review-summary triplets, where user u writes a review x and a summary y . Review x is a sequential input where $x = \{x_1, x_2, \dots, x_i, \dots, x_n\}$ is a sequence of n number of words, and i is the index of the input words. Summary y is the shorter output where $y = \{y_1, y_2, \dots, y_i, \dots, y_m\}$ of

Table 1 Dataset description for Tripdata

Reviews	140,874
Summaries	140,874
Users	41,745
Reviews/user	3.37

m number of words where $n > m$. The aim of our model is to generate a summary y from review x by attending to u 's aspects preference on summarizing reviews.

4 Dataset

We use hotel review dataset from Trip Advisor [22]. We create a new dataset Tripdata from this data. Tripdata is collected from identifying manipulated offerings on review portals [22], which was collected from TripAdvisor which is a travel review website. The data contains user-generated reviews along with author names and titles. The title of a review provides summarized information of the review. Moreover, there are many noisy samples found in the data. Different filtering techniques are used to obtain a clean data such as:

- i. Review length filter is used to remove reviews less than twenty five words and more than five hundred words. This is done to remove reviews that are too short and too long
- ii. Title length filter is used to remove titles less than five words. This is done to remove titles which are too short
- iii. Aspect-based filter is used to remove titles which do not have any aspects relating to hotel reviews. For hotel review data, we have mentioned six aspects along with their seed words in Table 1. The seeds words mentioned are expanded with boot strapping method by aspect segmentation algorithm [23]. Finally, the samples where the seed words are not present in the title are removed.

Statistics for Tripdata are provided in Table 1. We randomly split the dataset into 5000 user-review-summary triplets for test, 1500 user-review-summary triplets for validation and the rest for training purpose.

5 Our Model

The classical encoder–decoder network works with the review text as the input and the summary text as the output, where there is limited control over generation and key information may be missing in the summary. In addition to this, we also want the summary to be guided by personalized key information. That is where personalized key information-guided network is introduced shown in Fig. 2.

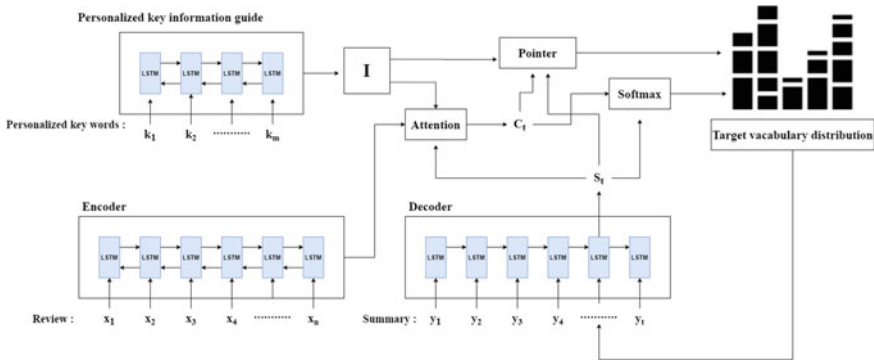


Fig. 2 Personalized key information-guided network architecture

5.1 Preprocessing

In detail, the first step is to identify which aspect each user cares about. Since we are working exclusively on hotel reviews, we can specify which aspects we should look for [21]. In addition to the seed words mentioned [21], we have also included some words with respect to this use case (Table 2). Firstly, we remove the stopwords from all the reviews (to avoid noise) and create a corpus with all the reviews and summaries for each user. Secondly, we find the most commonly used 30 words from each corpus (users will talk more about the aspects they care about). Then, we identify which aspects these words relate to if any. By this method, we identify aspect preference for each user.

The second step is to keywords using TextRank algorithm for each review. Then, we filter out the keywords which are not related to the aspects mentioned for the specific user. By this method, we will be able to get personalized key information for each review.

Table 2 Aspect words and their keywords

Aspect	Keywords
Location	'location,' 'traffic,' 'minute,' 'walk,' 'property,' 'noise,' 'loud'
Service	'server,' 'service,' 'welcome,' 'staff,' 'management,' 'friendly,' 'front desk,' 'helpful,' 'help,' 'courteous'
Room	'room,' 'bed,' 'clean,' 'cleanliness,' 'dirty,' 'bathroom,' 'rooms,' 'suite,' 'décor,' 'spacious,' 'suite'
Value	'value,' 'price,' 'quality,' 'worth,' 'rate,' 'cost,' 'luxury,' 'cheap,' 'bargain,' 'affordable,' 'money,' 'overpriced,' 'expensive,' 'budget'
Facility	'pool,' 'parking,' 'internet,' 'wifi,' 'vibe,' 'facilities,' 'amenities,' 'hotel,' 'fitness,' 'gym'
Food	'delicious,' 'breakfast,' 'coffee,' 'restaurant,' 'eatery,' 'food,' 'restaurants,' 'bar'

5.2 Personalized Key Information Guided Network

The architecture we use here is similar to the one used in guiding generation for abstractive text summarization based on key information guide network [20]. The difference between these models is that while the KIGN used the whole keywords as a guided mechanism, our model used only the personalized key information for guiding the network.

The traditional encoder–decoder model works with source text as input and summary is the output. In this method, the summarization method is hard to control because of the lack of a guided mechanism. So we propose adding two enchantment to traditional sequence-to-sequence model: personalized attention mechanism and pointer mechanism.

Firstly, with the help of TextRank algorithm, we extract personalized keywords for the reviews. As displayed in Fig. 1, the personalized keywords are passed individually to a bi-LSTM present in personalized key information-guided network, and then, we join the last forward hidden state and backward hidden state as the personalized key information representation I :

$$I = \frac{h_1^{\leftarrow}}{h_n^{\rightarrow}} \quad (1)$$

Personalized Attention Mechanism: Traditional attention mechanism uses decoder state to attain the attention distribution of encoder hidden states which makes it hard to have a guided mechanism. The personalized key information representation I as the input onto to tradition attention models (Eq. 2) and the personalized attention mechanism is shown in Eq. 3

$$e_{ti} = v^t \tanh(W_h h_i + W_s s_t) \quad (2)$$

$$e_{ti} = v^t \tanh(W_h h_i + W_s s_t + W_I I) \quad (3)$$

where W_I is learnable parameter. We will be using e_{ti} to obtain the latest attention distribution and context vector c .

$$\alpha_t^e = \text{softmax}(e_t) \quad (4)$$

$$c_t = \sum_{i=1}^N \alpha_{ti}^e h_i \quad (5)$$

An advantage of our personalized key information network is that it makes sure that more focus is given to the personalized keywords. So, more focus will be given to the personalized aspects and prior knowledge is given to the model.

Pointer mechanism: Some of the keywords might be missing from the target vocabulary, which might result in the summaries losing key information. So, we introduce pointer generation network which is a hybrid network which lets us both copy the words from keywords and also generate the words from target vocabulary. We use the personalized key information I , the context vector c_t and decoder state s_t and uses them to calculate a soft switch p_{key} , which makes a decision on whether to generate words from target vocabulary or reproducing words from the input text:

$$P_{\text{key}} = \sigma(w_I^T I + w_c^T c_t + w_{s_t}^T s_t + b_{\text{key}}) \quad (6)$$

where w_I^T , w_c^T , $w_{s_t}^T$ and b_{key} are learnable parameter, σ the sigmoid function.

Our pointer mechanism, where the personalized key information is included, has the capacity to recognize personalized keywords. The attention distribution is used as the probability of the input word a_i and the probability distribution to predict the successive word was obtained:

$$P(y_t = a) = P_{\text{key}} P_v(t = a) + (1 - P_{\text{key}}) \sum_{i:a_i=a} \alpha_{ti}^e \quad (7)$$

Note that if a is an OOV word, then P_v is zero. The main advantage of pointer generation is the ability to produce OVV words with respect to the personalized keyword.

We reduce the maximum likelihood loss at every individual decoding time step during training, which is mainly used for creating sequences. We define y_b^* as the target word for each decoding time step b and the loss is given as

$$L = \frac{-1}{T} \sum_{t=0}^T \log P(y_b^* | y_1^*, \dots, y_{t-1}, x) \quad (8)$$

5.3 Experiments

All experiments has been conducted on Tripdata which has 134,374 training triplets, 5000 test triplets and 1500 validation triplets. Two bidirectional LSTMs of dimension 256 are used in encoder, and an LSTM of embedding 256 is used for decoder. We use GloVe pretrained model for word embedding with dimension 300. A vocabulary size of 63,172 words is used for both source and target texts. W truncate the review to 450 token, personalized keywords to 30 tokens and the summary to 50 tokens for training. The dropout used [24] with probability $p = 0.2$. During training, we use loss on the validation set to implement early stopping and also apply gradient clipping [25] with range $[-4, 4]$. At test time, we use beam search by setting a beam size of 7

for producing summaries. We use Adam as our optimizing algorithm and by setting the batch size to 128. Our model was trained on 300 training iterations.

5.4 *Evaluation Methods*

We exploit ROUGE [26] metric for evaluating our model. ROUGE scores reported in this paper are computed by Pyrouge package.

5.5 *Comparison Methods*

As far as we know, all previous review summarization studies focused on the multi-review summarization scenario, which is essentially different from our task. Here, we compare with several methods which are popular in abstractive text summarization approaches.

- **S2S** is sequence to sequence model with attention. For this model, an attention mechanism is added to the framework by taking into account context vectors in hidden state of the encoder that helps in decoding the target sequences and achieved state-of-the-art results for DUC-2004 and Gigaword datasets.
- **SEASS** [13] adopts a selective network to select important information from review into S2S + Att. This model uses a selective mechanism to control the information flow which highlights important information and release the burden of the decoder which performed significantly better on ROGUE score for English Gigaword, DUC 2004 and MSRATC test sets.
- **PGN** [12] adopts a copy mechanism to copy words from review when generating summarization into S2. This model which uses pointer generation network, which is a hybrid network, which lets us both copy the words from keywords and also generate the words from target vocabulary. PGN was able to achieve state-of-the-art results on CNN/Daily Mail dataset.
- **User-Aware Sequence Network** [21] is used for personalized review summarization. This model is based on S2S network with user-aware encoder and user-aware decoder where selective mechanism is used in user-aware encoder to highlight user-specific information in the review. USN achieved state-of-the-art rouge for personalized summarization.

Table 3 ROUGE $F1$ scores on the test set for various models

Models	Rogue-1	Rogue-2	Rogue-L
S2S	28.15	14.85	28.88
SEASS	29.17	15.22	29.21
PGN	32.53	19.72	33.63
USN	32.89	20.11	33.48
Our model	34.36	21.51	34.28

6 Results

6.1 Review Summarization

Our results are shown in Table 3. Our model is evaluated with standard ROUGE score, taking the $F1$ scores for ROUGE-1, ROUGE-2 and ROUGE-L. We observe that the S2S model has the lowest score since it employs a basic encoder decoder model for summarization. S2S model. For SEASS, we add the selective mechanism, and the scores improve slightly. However, it is important to filter the review text using a selective mechanism for summarization. Therefore, we use guided mechanism with personalized keywords, and it is seen to be efficient. Our model has a gain of 6.21 ROUGE1, 6.66 ROUGE-2 and 5.4 ROUGE-L on. The PGN model, soft copy mechanism is added to the encoder decoder model, performs better than the previous models. The copy mechanism is incorporated on our model as we observed that copying word from the input text is shown to improve the summarization. Our model outperforms PGN by 1.83 ROUGE1, 1.79 ROUGE-2 and 0.6 ROUGE-L.

We also used USN model where modelling is done using user-related characteristics. This performed better than all the other models. So, we know that incorporating user specific information in our model, and using that as a guidance mechanism will improve summarization. Therefore, we have used the personalized key information guide network in our model and we have achieved 34.36 ROUGE1, 21.51 ROUGE-2 and 34.28 ROUGE-L. Our model has exceeded the baseline models with the implementation of personalized attention mechanism and pointer mechanism over a sequence to sequence model.

6.2 Human Evaluation of Personalization

Personalized key information-guided network is a personalized model which also capture aspect preference of individual reviewer. Important aspects for each user are found out in the preprocessing steps. Therefore, we want to identify if these aspects are present in the summaries generated by our model.

We make use of six aspects already mentioned in preprocessing, and for our use case of hotel review, we add a label to describe the overall attitude towards the hotel.

Table 4 Aspect-level precision, recall and *F1* for various models

Models	Precision	Recall	<i>F1</i> -score
S2S	0.502	0.501	0.501
SEASS	0.512	0.512	0.512
PGN	0.533	0.542	0.5376
USN	0.565	0.572	0.568
Our model	0.615	0.625	0.619

Then, the generated summaries are labeled with the above-mentioned labels. The generated summaries are labeled as follows:

Example 1: excellent customer service (**service**).

Example 2: Great rooms and perfect location (**rooms, location**).

To perform this human evaluation, 1400 user reviews are randomly sampled from our test set. We produce summaries for the reviews using our personalized key information-guided network model also along with those predictions are done using other models such as S2S, SEASS and PGN. Then, we manually label generated summaries and the user preferences. This is done to see how many of the user’s preferences are present in the generated summary. While labeling, we check whether the user labels are present in the review and if not, those labels are removed from user aspects. After all the predicted summaries are labeled, we compute aspect level precision, recall and *F1* score for different models and shown in Table 4. We observe that our model performs better than the other existing models which shows that our model captures personalized features better. This is because of the presence of personalized attention mechanism and pointer mechanism which helps us capture user aspect preferences.

6.3 Case Study

Figure 3 shows an example of review between our model and USN. The output from both the models are compared with the gold summary. The review talks about aspects such as location, room, service and value. But the gold summary is a general text with no aspects. Here, it is observed that our model was able to introduce aspects into the summary and make it more meaningful. Both USN and our model were able to include aspects in our summary. We identified three user-specific aspects namely *service*, *room* and *value* from previous reviews. But it seen that USN covers two aspects, *location* and *service*, of which one, *location*, is not mentioned in user aspects. For our model, it is seen that the summary included information from the three user-specific aspects such as *service*, *room* and *value*. Therefore, our model was able to summarize the review, taking it account the user’s preferences.

Review:	I read some of the past reviews and was hesitant to stay here. I usually stay close to the Seattle Center on 5th and Roy. This location is very comfortable for me to travel. The rooms, which are pretty dated, had a modern feel to them. There was a nice sofa bed in the room. The Heavenly bed was indeed nice. The internet was easy to use, but there is a charge for the service. The charges were too high for me. Housekeeping service was great; the attendant was very courteous and professional. Front desk personnel was friendly and cordial. And I did receive a phone call a few hours after I checked in to see if all was in order. I did use the business center and the gentleman running the office was attentive. Room service was surprisingly fast and efficient. Since it was pretty cold to go out at 10pm on Saturday night, I was very impressed with the efficiency and quality of service with the room service. Food was not bad.
Gold summary :	scared of the reviews enjoyed the stay
User aspect :	'service', 'room', 'value'
USN :	good location and great room service
Our Model :	nice room , bad value , great friendly service

Fig. 3 Comparison of the output between the two personalized models on a hotel review are given. Actual summary given by the user is shown as **Gold** summary and the user's preferences are given in **User aspect**

7 Conclusion

In this project, we address personalized review summarization and propose a Personalized Key Information Guided Network to account for user aspect preference into personalized review summarization. At first, we use extractive methods to get personalized keywords as additional input for reviews which accounts for the user preference. Secondly, the important feature of this model is the personalized key information-guided network which helps include the personalized features in the summarization used along with the pointer generation network. To validate our model, we have created a Tripdata dataset containing of hotel reviews from the TripAdvisor website. From experimentation, our model was seen to be performing better than existing and traditional models for the case of personalized summarization.

As future enhancement to the existing model, we can introduce a transformer model with a guided network which would help in summarization of long texts and also introduce parallelize the operations. In addition to that, instead of setting predefined words to identify each aspect, we could introduce a network which identify aspect words from their semantic meaning. This would help in generalizing the model across many domains not just the hotel review dataset.

References

1. M.-T. Luong, H. Pham, C.D. Manning, Effective approaches to attention-based neural machine translation. arXiv preprint [arXiv:1508.04025](https://arxiv.org/abs/1508.04025) (2015)
2. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need, in *Advances in Neural Information Processing Systems (NIPS)*, Dec

- 2017, pp. 5999–6009
3. S. Chopra, M. Auli, A.M. Rush, Abstractive sentence summarization with attentive recurrent neural networks, in *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (Association for Computational Linguistics, 2016), pp. 93–98
 4. R. M6ro, M. Bielikov6, Personalized text summarization based on important terms identification, in *23rd International Workshop on Database and Expert Systems Applications* (IEEE, 2012), pp. 1529–1588
 5. R. Mihalcea, P. Tarau, Textrank: bringing order into texts, in *Proceedings of EMNLP 2004* (Association for Computational Linguistics, Barcelona, 2004), pp. 404–411
 6. D. Bahdanau, K. Cho, Y. Bengio, Neural machine translation by jointly learning to align and translate. arXiv preprint [arXiv:1409.0473](https://arxiv.org/abs/1409.0473) (2014)
 7. A.M. Rush, S. Chopra, J. Weston, A neural attention model for abstractive sentence summarization. arXiv preprint [arXiv:1509.00685](https://arxiv.org/abs/1509.00685) (2015)
 8. S. Takase, J. Suzuki, N. Okazaki, T. Hirao, M. Nagata, Neural headline generation on abstract meaning representation, in *Empirical Methods in Natural Language Processing* (2016)
 9. R. Nallapati, B. Zhou, C. dos Santos, C. Gulc6ehre, B. Xiang, Abstractive text summarization using sequence-to-sequence RNNs and beyond, in *Computational Natural Language Learning* (2016)
 10. Y. Miao, P. Blunsom, Language as a latent variable: discrete generative models for sentence compression, in *Empirical Methods in Natural Language Processing* (2016)
 11. L. Yu, J. Buys, P. Blunsom, Online segment to segment neural transduction, in *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing* (Association for Computational Linguistics, Austin, TX, 2016), pp. 1307–1316
 12. A. See, P.J. Liu, C.D. Manning, Get to the point: summarization with pointer-generator networks. arXiv preprint [arXiv:1704.04368](https://arxiv.org/abs/1704.04368) (2017)
 13. Q. Zhou, N. Yang, F. Wei, M. Zhou, Selective encoding for abstractive sentence summarization, in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers) (2017), pp. 1095–1104
 14. K. Yao, L. Zhang, D. Du, T. Luo, L. Tao, Y. Wu, Dual encoding for abstractive text summarization. *IEEE Trans. Cybern.* (2018), pp. 1–12
 15. W. Lu, L. Wang, Neural network-based abstract generation for opinions and arguments, in *NAACL* (2016), pp. 47–57
 16. D. Krishnan, P. Bharathy, Anagha, M. Venugopalan, A supervised approach for extractive text summarization using minimal robust features, in *2019 International Conference on Intelligent Computing and Control Systems (ICCS)* (2019), pp. 521–527. <https://doi.org/10.1109/ICCS45141.2019.9065651>
 17. K. Shalini, H.B. Ganesh, M.A. Kumar, K. Soman, Sentiment analysis for codemixed Indian social media text with distributed representation, in *2018 International Conference on Advances in Computing, Communications and Informatics, ICACCI* (IEEE, 2018), pp. 1126–1131
 18. N. Lalithamani, R. Sukumaran, K. Alagammnai, K.K. Sowmya, V. Divyalakshmi, S. Shanmugapriya, A mixed-initiative approach for summarizing discussions coupled with sentimental analysis, in *Proceedings of the 2014 International Conference on Interdisciplinary Advances in Applied Computing* (ACM, 2014), p. 5
 19. M.V.K. Kiran, R.E. Vinodhini, R. Archanaa, K. Vimalkumar, User specific product recommendation and rating system by performing sentiment analysis on product reviews, in *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)* (2017), pp. 1–5. <https://doi.org/10.1109/ICACCS.2017.8014640>
 20. C. Li, W. Xu, S. Li, S. Gao, Guiding generation for abstractive text summarization based on key information guide network, in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, vol. 2 (2018), pp. 55–60
 21. J. Li, H. Li, C. Zong, Towards personalized review summarization via user-aware sequence network, in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33 (2019), pp. 6690–6697

22. J. Li, M. Ott, C. Cardie, Identifying manipulated offerings on review portals, in *EMNLP* (2013)
23. H. Wang, Y. Lu, C. Zhai, Latent aspect rating analysis on review text data: a rating regression approach, in *SIGKDD* (2010), pp. 783–792
24. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **15**(1), 1929–1958 (2014)
25. R. Pascanu, T. Mikolov, Y. Bengio, On the difficulty of training recurrent neural networks, in *ICML* (2013), pp. 1310–1318
26. C.-Y. Lin, Rouge: a package for automatic evaluation of summaries, in *Text Summarization Branches Out* (2004)