FARSum: Feedback-Aware Abstractive Product Review Summarization

Ming Wang, Javid Huseynov[1], Jim Chan, and Jin Li

Amazon, New York, USA
mingww@amazon.com, husej@amazon.com, jamchan@amazon.com, jincli@amazon.com

Abstract. Retail customers read through multitude of online product reviews to make confident purchase decisions. To automate this process, we explore and evaluate several state-of-the-art (SOTA) models for summarizing product reviews along three dimensions: a summary product verdict, pros, and cons. To improve the performance of summarization from a large number of reviews per product, we propose FARSum, an efficient solution that leverages review filtering based on review recency and customer feedback including review helpful vote and review rating in the first stage. To improve context generalization across product categories, we train a BART-based model on synthetic review summaries and fine tune the model using ground-truth summary labels. We demonstrate the competitive performance of our solution vis-a-vis other SOTA models using the ROUGE metrics.

Keywords: Feedback Awareness, Context Generalization, Abstractive Text Summarization

1 Introduction

In e-commerce, organic and sponsored recommendations built from product knowledge and shopping missions help retail customers make informed and confident purchase decisions. Product reviews from online shoppers are a key source of information in this process. Customers read through a variety of reviews to assess product features, attributes, and performance across different dimensions. To get a quick sense of what's good or bad about a given product, they often organize reviews by high and low ratings, or search for reviews with "pros and cons" using keywords or other cues. This is a time-consuming process compounded with a complex user experience and broad context of the reviews. A summary of customer reviews can help facilitate the online shopping process by highlighting only the relevant product insights. This would, in turn, reduce the time to a purchase decision and the amount of context switching that distracts customers from their shopping missions. To provide informative and abstractive review summary, we will explore different solutions from extractive summarization to

[1] an Associate Professor of Practice at the Columbia University School of Prof. Studies

abstractive summarization, and propose our efficient feedback-aware abstractive summarization procedure.

Several extractive and abstractive methods have been explored for product review summarization [3–5, 9]. Extractive summarization is a subset of phrases or sentences that best represent a given document. While extractive methods are robust in a sense of preserving the sentences in the original text, they are prone to producing summaries taken out of the overall document context. Thus, in the context of product reviews, extractive summaries often surface redundant or incomplete customer feedback about a product.

In contrast, an abstractive summarization method generates a concise summary that captures the salient ideas in the original document. These ideas may not necessarily be expressed using the original phrases or sentences, but abstractive summaries generalize the overall context better than extractive ones. In recent years, especially with the advancements in the transformer-based language models, abstractive method shows it's power to conduct abstrctive summarizaiton and has become a popular choice for summarizing product reviews. As early as [9], an abstractive method was proposed for generating lists of aspects and sentiments. More advanced applications leveraged deep neural network-based learning [1–3,5]. But due to the lack of high quality ground-truth datasets, unsupervised learning or few-shot supervised learning methods didn't achieve good performance. [4] collected the largest multi-document opinion summarization dataset, known as AmaSum, to develop a model that outperforms all previous models.

The main challenge for review summarization is a large number of customer reviews surfaced per product. As an example, a single mobile phone case product on Amazon features over half a million customer reviews. At such volume, it is almost impossible or impractical to train a learning model that generalizes over all available reviews. To tackle this problem, [4] proposed SELSUM, which fits a prior network model to selecting important reviews for training a summarization model. Though this approach achieved SOTA performance, the model needs to run on each sentence to select important reviews, which increases the complexity of model inference. There have been other applications of the fine-tuned encoder-decoder models like [12,15]. By applying systematic perturbations to the model's input during the inference, these approaches generated multiple intermediate summaries per product and then summarized those. [15] proposed another two-stage multi-document summarization: a cluster of reviews is generated followed by two layers of weakly supervised model for summarization. These two efforts effectively deal with a large volume of reviews, but they are computationally costly. These approaches also neglect the fact that just based on ground-truth data, the fine tuned model cannot generalize the summarization procedure to all product domains.

In order to reduce the complexity of the model inference, we are trying to simplify the two-step review selection and generation procedure. Specifically, we propose review selection based on the helpful votes index and review ratings. To improve the model summarization generalization ability, we then train the

BART-base model on pseudo summary labels and further fine tune the trained model with manually labeled data to generate a high quality verdict summary. Additionally, we propose generating summary pros and cons to help customers discover and compare products for confident purchase decisions. Overall, our contributions can be summarized as follows:

- We introduce a new light-weight procedure for customer review filtering and selection based on customer feedback;
- We train a summarization model on pseudo summary data fine-tuned using the human-labeled data, to capture context and generate a good quality summary;
- We generate verdict, pros, and cons summaries, to help customers make final decisions.

The paper is organized as follows. We start with introducing our methodology including the review dataset, model overview, and the two main components of the model: review candidate selector and summarizer in section 2. We then conducted the model performance comparison versus relevant models in section 3. At last, we summarize our proposed procedure and draw the conclusion in section 4.

2 Methodology

2.1 Dataset

To train a supervised model for product review summarization, Brazinskas et al. [4] introduced the AmaSum dataset of 33,000 verdict, pro, and con summaries for more than 31,000 products on Amazon. These summaries were narrated by professional reviewers based on the information collected from several product review sites. We used the AmaSum dataset to train and evaluate the proposed light-weight procedure with a good baseline performance. Scaling our approach for production would require collecting more human-annotated summaries with varying patterns, styles, and lengths.

2.2 Approach Overview

In a nutshell, we selected candidate reviews based on their ratings, helpful votes, and recency for the first step. Then, we train the BART base model on pseudo summary data and further fine tune the trained model with human-labeled data for different subtasks of pros, cons, and verdict summarization. The following flowchart(Figure 1) shows the the high level processes of data preparation, model training, and inference. We will dive deep into each component in the following sub-sections.

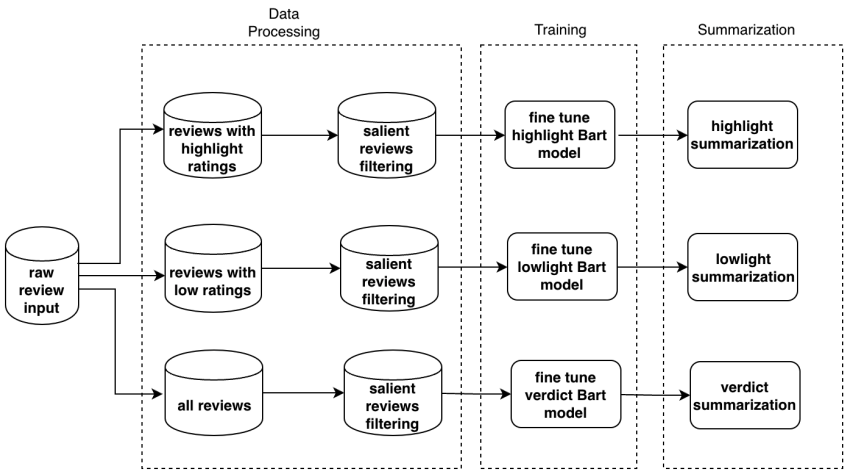

Fig. 1. Workflow of Light-Weight Review Summarization for Verdict, Pros, and Cons

2.3 Review Candidates Selector

One major challenge for customer review summarization is the large volume of reviews per product. It's challenging for a deep learning model to generalize over such a contextually diverse corpus. Fortunately, the review sections of most e-commerce websites provide useful features like ratings and helpful votes. This is organic feedback that reflects customers' sentiment and confidence in their reviews. It helps narrow down the candidates for review summary generation through the following two steps:

- Step 1: The reviews with 4 <= ratings <= 5 reflect pros, those with 1 <= ratings <= 2 reflect cons, and all reviews are considered for verdict generation. Thus, using the review ratings, we organize the reviews by sentiment to further train the model for each subtask.
- Step 2: More recent reviews with helpful votes can reflect more relevant and updated information. Thus, we re-rank the reviews using the helpful vote index first and then recency index if the reviews have same helpful vote for each product.

2.4 Summarizer

The pre-trained Transformer-based [16] models, such as Google's T5 [14] and PEGASUS [17], and Facebook's BART [10], have been successfully applied on a range of natural language generation (NLG) tasks, including the abstractive text summarization. We explored the pre-trained BART and PEGASUS models for our baseline, and chose the BART-base model. BART is a transformer-based sequence-to-sequence model, in which the encoder is similar to the BERT model

[6] and the decoder is similar to the GPT model [13]. In the encoder layer, BART adds a causal decoder to BERT's bidirectional encoder architecture and replaces BERT's fill-in-the blank task with a complex mix of pre-training tasks, where spans of text are replaced with a single mask token. [10] explains the concept as in figure 2. As shown in the figure, the original document is 'A B C D E'. The span [C, D] is masked before encoding and an extra mask is inserted before B. This results the corrupted document 'A _ B _ E' as input to the encoder. The model must learn to reconstruct the original document based on the encoder's output and previous uncorrupted tokens.

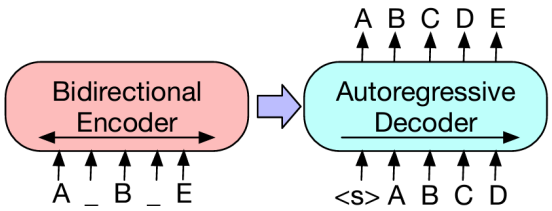

Fig. 2. Overview of BART Architecture

We could directly fine tune the BART-base model with ground-truth summaries as labels. However, due to the limited ground-truth summary data, it is hard for the model to learn and generalize over the domain knowledge for all products. This may lead to a less coherent and actionable summarization for any specific product category. In order to solve this problem, we apply a weakly supervised training step first before fine tuning on the ground-truth review labels. Specifically, we use Lexrank to extract important sentences from the corpus of reviews per product, to generate a pseudo summary label. We then select the top 10 other sentences that are contextually similar (by ROUGE score) to the the pseudo summary label. This approach helps generate as many pseudo summary, reviews pairs as we need. Consequently, the weakly supervised training will help the BART-base model learn more domain context in e-commerce setting.

Next, we fit the pre-trained model with the ground-truth data, which enables the model learning of the summary format and improved the summary quality. The process is shown in the flowchart on Figure 3, where we use one type of a summary (verdict, pros, or cons) for example.

2.5 Discussion

In this section, we briefly compare FARSum to other summary methods to highlight the novelty of our methods.

Traditional extractive summary can provide original text information, however, it cannot provide more comprehensive text information compared to abstractive summarization. The current abstractive approach started to deal with

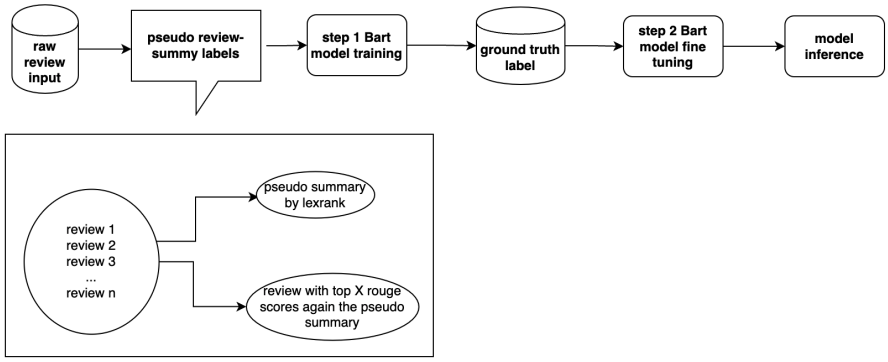

Fig. 3. Overview of Summarization Training Workflow

customer review summary problems based on few-shot supervisions. These approaches didn't achieve very good performance due to lack of high quality ground-truth data. Leveraging large opinion summary dataset AmaSum from [4], the main challenge becomes how to deal with large corpus of raw review data. Most approaches solve this problem by using a candidate selection procedure. However, the procedures used to select candidate reviews for summarization are always too complex; and they don't consider organic customer feedback information like the customer's helpfulness vote for the review.

In this paper, we simply the candidate selection step based on helpful vote index and review ratings. In the training step, we further trained the model on pseudo summary labels to allow the model to learn more domain information, which is also a weakness of previous approaches.

3 Experiments

The Recall-Oriented Understudy for Gisting Evaluation (ROUGE) is a widely used set of metrics for evaluating the quality of machine-generated summaries or translations. It assesses the contextual similarity of a model-generated text with a reference (i.e. human-generated) text. ROUGE-1 and ROUGE-2 measure the overlap in unigrams and bigrams, respectively, between the model-generated and reference texts. Similarly, ROUGE-L measures the overlap in the longest common subsequence (LCS) of n-grams between the model-generated and reference texts.

In the following subsection 3.1, we present our evaluation results based on ROUGE-1, -2, -L metrics for verdict summarization using some baseline models on AmaSum dataset. These results helped us select the base model for review summarization. More detailed precision, recall, and F1 scores on ROUGE are

discussed in subsection 3.2. Finally, a generated review summary case study is presented in subsection 3.3.

3.1 Baseline Verdict Summarization Evaluation

We evaluated the verdict summarization using the following baseline extractive and abstractive summarization models:

LEXRANK [7] is an unsupervised graph-based extractive summarization algorithm. Similar to TextRank, it uses a connectivity matrix using intra-sentence cosine similarity as the graph representation of sentences.

MEANSUM [5] and COPYCAT [3] are unsupervised abstractive summarization models that do not rely on any review-specific features. MEANSUM uses an auto-encoder where the mean of the representations of the input reviews decodes to a reasonable summary-review. COPYCAT is a SOTA abstractive summarization model that learns the latent semantic representations of reviews and generates summaries using those.

SELSUM [4] is another SOTA model that fits a prior network model to select important reviews for training a summarizer on those.

BARTFT and PEGASUSFT are two new fine-tuned models by us based on BART and PEGASUS pre-trained models, respectively. Based on the model performance and inference speed, BART was chosen as a base model for simplifying the training and inference procedure later. More information can be found in the following two sub-sections.

We didn't evaluate MEANSUM and COPYCAT, but obtained the ROUGE F1 scores from [4]. For other models listed in Table 1, we evaluated using ROUGE score on the same test set (of around on 3200 products) which is used in MEANSUM and COPYCAT. Based on the ROUGE-1, -2, -L F1 scores for verdict evaluation in Table 1[1], supervised abstractive models significantly outperform the extractive and unsupervised abstractive models. We can also learn that the fine-tuned BART model performs similar to SELSUM. The PEGASUS fine-tuned model also achieves good performance, though it is 4 times slower than the BART-based model. Based on this baseline model evaluation, we proceeded with fine tuning the BART model for our light-weight training procedure.

Table 1. ROUGE F1 Scores for Verdict Summarization

	rouge-1	rouge-2	rouge-L
lexrank	15.29	1.26	11.62
meansum	13.78	0.93	11.7
copycat	17.05	1.78	14.5
selsum	23.12	5	17.51
bartft	22.04	4.6	17.05
pegasusft	22.3	4.45	16.53

[1] Due to the table space limitation, we lowercased model names to show all model results.

3.2 Evaluation of Verdict, Pros, and Cons Summaries

Table 2. ROUGE Recall, Precision and F1 Scores for Verdict Summarization

	selsum	lexrank	bart_s1k	bart_s25k	farsum
rouge1_p	23.92	19.79	23.42	24.1	**24.12**
rouge1_r	**24.74**	15.55	20.47	22.1	22.65
rouge1_f1	**23.13**	15.29	20.62	22.04	22.31
rouge2_p	**5.24**	1.57	4.43	5.08	5.15
rouge2_r	**5.27**	1.35	3.75	4.58	4.71
rouge2_f1	**4.99**	1.26	3.81	4.6	4.7
rougeL_p	18.07	14.95	18.13	18.55	**18.51**
rougeL_r	**18.78**	11.94	16.01	17.18	17.59
rougeL_f1	**17.51**	11.62	16.04	17.05	17.23

In Table 2, we detailed the precision, recall, and F1 scores on ROUGE-1, -2, -L metrics for different models, including SELSUM, LEXRANK, BART-s* and FARSum. The BART-s* are directly fine-tuned BART models based on different training sample sizes. For example, BART-s1k is the BART model fine-tuned on 1000 random product samples. The reason for exploring different training sample size is to provide some context if ground truth data needs to be collected for training in another user case. 'farsum' means the model first trained on pseudo label, then fine-tuned on 25,000 ground truth data. With all the training samples and our light-weight review selector, the fine-tuned model (BART-s25k) is very close to the SOTA SELSUM model. Moreover, our proposed FARSum can outperform the directly fine-tuned BART model on the ground truth data.

We also learned that without too many training samples, the summarization model can still achieve very good performance like BART-s1k. See more training result with different sample size in table 3. For instance, with just 100 training samples, we can achieve ROUGE-L F1 score of 15.86 vs 17.51 for SELSUM. This provides some context if ground truth data needs to be collected for training in another user case.

Table 4 lists precision, recall, and F1 scores on the ROUGE-1, -2, -L metrics for pros summarization. The fine-tuned BART (with the same sample size as SELSUM) outperforms the SELSUM model, so does our proposed FARSum. And just like in the verdict summarization, our FARSum can further improve the ROUGE score. We can also see that with 1000 training samples our procedure can still achieve impressive ROUGE metric scores.

Table 5 shows the precision, recall and F1 scores on the ROUGE-1, -2, -L metrics for cons summarization. Comparing to the verdict and pros summaries, all models have lower ROUGE scores in cons cases. The model with smaller training samples also results in lower scores than SELSUM, but still higher than LEXRANK. This is probably due to the fact that the cons context isn't as frequent in customer reviews and, consequently, there are fewer candidate

Table 3. ROUGE Scores for Verdict Summarization on Different Training Sample Size

	bart_s100	bart_s200	bart_s500	bart_s1k	bart_s25k
rouge1_p	23.86	22.51	23.63	23.42	24.1
rouge1_r	19.06	20.04	19.86	20.47	22.1
rouge1_f1	19.96	20	20.38	20.62	22.04
rouge2_p	4.65	4.31	4.59	4.43	5.08
rouge2_r	3.59	3.72	3.71	3.75	4.58
rouge2_f1	3.8	3.74	3.86	3.81	4.6
rougeL_p	18.9	17.52	18.34	18.13	18.55
rougeL_r	15.25	15.73	15.52	16.01	17.18
rougeL_f1	15.88	15.61	15.87	16.04	17.05

Table 4. ROUGE Recall, Precision, and F1 Scores for Pros Summarization

	selsum	lexrank	bart_s1k	bart_25k	farsum
rouge1_p	**25.68**	17.53	21.45	24.59	24.52
rouge1_r	17.89	16.61	19.7	20.1	**20.69**
rouge1_f1	19.9	15.74	19.17	20.99	**21.4**
rouge2_p	**5.02**	1.44	3.62	5.01	4.94
rouge2_r	3.37	1.39	3.22	3.98	**4.07**
rouge2_f1	3.78	1.30	3.17	4.19	**4.25**
rougeL_p	**18.45**	15.17	15.14	17.36	17.12
rougeL_r	12.88	14.47	14.01	14.33	**14.59**
rougeL_f1	14.26	13.66	13.53	14.86	**14.99**

sentences than in verdict or pros cases. We can see the rating distribution in Table 6.

3.3 Summarization Examples

To demonstrate the quality of a summary generated by the proposed lightweight procedure, we present the pros summarization example for Shark Vacuum Cleaner. The top 10 candidates identified by our review selector are listed in Appendix A.1. The ground-truth summary and the summaries generated with BART-s1000 and FARSum are listed in Table 7.

In Appendix A.1, we marked the mentioned aspects in BART-25k generation in red text. From these aspects, we can see that the functionalities like "HEPA filter", "pet hair removal" are mentioned several times in the reviews and included in the BART-25k generated summary. However, if we read through all reviews, some common features of a vacuum, like "powerful suction" and "easy to use", are mentioned frequently as well. We marked these in blue text. Interestingly, these features are not mentioned in the golden summary, but appear in the summary produced by the BART-1000 model. While the abstractive summarization generalizes the context of all reviews, it still may not be able to capture all aspects within the short snippet of textual summary. To address this trade-

Table 5. ROUGE Recall, Precision, and F1 Scores for Cons Summarization

	selsum	lexrank	bart1k	bart25k	effisum
rouge1_p	13.54	13.78	14.35	**14.62**	14.41
rouge1_r	**17.46**	10.06	10.92	11.11	11.46
rouge1_f1	**13.98**	10.40	11.55	11.72	11.92
rouge2_p	2.27	1.06	2.18	**2.63**	2.45
rouge2_r	**2.89**	0.79	1.53	1.88	1.91
rouge2_f1	**2.29**	0.80	1.66	2.01	1.98
rougeL_p	10.73	11.46	12.01	**12.26**	12.12
rougeL_r	**14.24**	8.46	9.26	9.41	9.77
rougeL_f1	**11.21**	8.68	9.72	9.89	10.09

Table 6. Distribution of Reviews in Each Rating Bucket

rating	count	percentage
1	29508	12.28%
2	13781	5.74%
3	18463	7.69%
4	33175	13.81%
5	145301	60.48%

off in the future, we can explore the aspect-based summarization approaches leveraging customers' shopping missions and preferences in the future.

4 Conclusion

In this paper, we tackled the problem of abstractive summarization from a large number of customer reviews per product. We applied a two-step filtering solution that leverages the customer-generated review ratings as well as the review recency, to select reviews for training a summarization model. We further evaluated several SOTA extractive and abstractive summarization models using the ROUGE metric on the ground-truth AmaSum dataset, and chose the BART-base model for fine tuning with the selected review dataset. We proposed to train the BART model on pseudo summary labels then fine tune on a human-generated ground-truth dataset. We also explored the influence of the training dataset size, in order to estimate the volume of ground-truth data that needs to be collected. We learned that as few as 100 - 500 training samples are sufficient for fine tuning a model to produce summaries with competitive ROUGE scores.

Overall, our proposed light-weight review selection procedure outperforms the SOTA SELSUM model on pros summaries and achieves similar performance on verdict and cons summaries. FARSum can perform better than the directly fine-tuned model with ground truth data in the summarization of verdict, pros, or cons cases. As future work, since product aspects can provide more granular product information when customers compare products, this can be one direct

Table 7. Product Summarization Example

Type	Content
Golden	We especially love the simple swivel system. Great for those who don't want to use a lot of strength to handle the vacuum.
FARSum	Features a HEPA filter, dust trap, and pet hair removal tool. Can be used as a canister vacuum or as a pet vac.
BART-s1k	Lightweight and easy to maneuver. Powerful suction power. Easy to empty.

to explore. To further facilitate customer purchase decisions, we can also develop a singular contrastive review summaries for a group of similar products.

References

1. Stefanos Angelidis and Mirella Lapata. 2018. Summarizing opinions: Aspect extraction meets sentiment prediction and they are both weakly supervised. *In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3675–3686, Brussels, Belgium. Association for Computational Linguistics.
2. Arthur Bražinskas, Mirella Lapata, and Ivan Titov. 2020a. Few-shot learning for opinion summarization. *In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 4119–4135, Online. Association for Computational Linguistics.
3. Arthur Bražinskas, Mirella Lapata, and Ivan Titov. 2020b. Unsupervised opinion summarization as copycat-review generation. *In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5151–5169, Online. Association for Computational Linguistics.
4. Arthur Bražinskas, Mirella Lapata, and Ivan Titov. 2021. Learning Opinion Summarizers by Selecting Informative Reviews, *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*
5. Eric Chu and Peter Liu. 2019. Meansum: a neural model for unsupervised multi-document abstractive summarization. *In Proceedings of International Conference on Machine Learning (ICML)*, pages 1223–1232.
6. Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.
7. Gunes Erkan and Dragomir R Radev. 2004. Lexrank: Graph-based lexical centrality as salience in text summarization. *Journal of artificial intelligence research*, 22:457–479.
8. Iftah Gamzu, Hila Gonen, Gilad Kutiel, Ran Levy, Eugene Agichtein. 2021. Identifying helpful sentences in product reviews. *NAACL*
9. Minqing Hu and Bing Liu. 2004. Mining and summarizing customer reviews. *In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177. ACM.

10. Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. BART: Denoising sequence-to-sequence pretraining for natural language generation, translation, and comprehension. *In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7871–7880, Online. Association for Computational Linguistics.
11. Thi Nhat Anh Nguyen, Mingwei Shen, Karen Hovsepian. 2021. Unsupervised class-specific abstractive summarization of customer reviews, , *ACL-IJCNLP 2021 Workshop on e-Commerce and NLP (ECNLP)*
12. Nadav Oved, Ran Levy, 2021. PASS: Perturb-and-select summarizer for product reviews. *ACL-IJCNLP 2021*
13. Alec Radford, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. 2018. "Improving language understanding by generative pre-training."
14. Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *J. Mach. Learn. Res.*, 21:140:1–140:67.
15. Ori Shapira and Ran Levy. 2020. Massive Multi-Document Summarization of Product Reviews with Weak Supervision. *AMLC*.
16. Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *In Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017*, December 4- 9, 2017, Long Beach, CA, USA, pages 5998–6008.
17. Yao Zhao, Mohammad Saleh, and Peter J. Liu. 2020a. PEGASUS: pre-training with extracted gap-sentences for abstractive summarization. *In Proceedings of the 37th International Conference on Machine Learning*, ICML 2020, 13-18 July 2020, Virtual Event, volume 119 of Proceedings of Machine Learning Research, pages 11328–11339. PMLR.

A Appendix

A.1 Raw Reviews

– I am a Shark convert. This vacuum is amazing. I used it on just the first floor of our house last, only 3 days after vacuuming with our old vacuum, and the *dust bin* was almost entirely filled! I don't know whether to be grossed out or thrilled? FYI - I do have 3 kids but we *don't have any pets*, don't wear shoes in the house, and I vacuum on the regular. This shocked me! This vacuum SUCKS - in the good way! I love that you can take it on the stairs *without the weight* of the entire vacuum also - huge advantage :) Would absolutely recommend!!.",
– "We're a busy outdoor family - the Shih Tzu is in and out as many times as you can count, bringing in leaves, *dirt*, stones, and today mud (they were working on a water main down the street and he HAD to supervise - very dirty feet) This machine gets it all...love it!.",
– 'I have had this vacuum for over a year. *I have two dogs and a cat*. One of my dogs is a german "shedder" and my cat is long hair. *This vacuum picks*

up the hair and cat litter without any problem. It doesnt́ blow dust or dump litter when you stop it. Iv́e tried several other, more expensive brands and always ended up running into problems. This was does an amazing job, the *lift away feature* makes it perfect for stairs and couches as well. Best vacuum Iv́e ever owned, it isnt́ even a comparison.',

- "Great suction! By far the best vacuum I've owned. I have *2 cats, a dog* and 3 kids. My only complaint is that my golden retrievers hair seems to wrap around the brush and I need to clean it out every now and then. Perhaps hold out for a price drop. I put it in my cart for $200 and watched it for awhile and finally purchased when I saw it drop to $150.",
- 'ok the vacuum works good but you cant take the plate off the bottom of the brush without a special driver—cant find one in any location. guess shark sells one as they made it to prevent average guy from being able to *clean out the hair* and carpet runners from the beater brush? so once you get this clogged up good luck pulling individual strands from the brush instead of being able to pull free and cut away spun threads from brush.',
- "My wife and I have bought 3 or 4 Dyson's over the years the last two broke pretty quickly. One of them was the Dyson Animal vacuum. Big, bulky, and expensive. I was never really sold on them. *We have a husky and she drops fur all year long*. So, when I went looking for a replacement I didn't want to break the bank but still wanted one that would hold up for more than a year. This Shark is the best vacuum I have ever had! *Better suction, light weight, strong spinning brush and easy to clean*. I will never buy a Dyson again.",
- 'I have three dogs, two cats and a boyfriend that live with me so I am constantly cleaning my home and this vacuum is a game changer. *The pet hair attachment works wonders on my black couch*. I thought I would need to replace my carpets but after shampooing them and using the brush setting on this vacuum they look like different carpets. *It's light and the dust trap is easy to empty*.',
- "This is a ver good vacuum, especially at this price point. We had some lead dust issues in our house and *needed a HEPA vacuum ASAP*. This was the best deal on the market. *Not only does it feature a HEPA filter*, it has a ton of incredible features—a versatile rotating head, *a lift-away feature so you can use it as a canister vacuum, several options for hoses*, hardwood and carpet suction settings. It works great and we're very happy with it.",
- 'Did not think I would be really impressed, but I was. I used my old vacuum and then the Shark on the same carpet and wood floor. The dirt the Shark picked up made me think I had really been not cleaning my floors! I love it. *It is lighter, easier to maneuver and a dream to empty*. It takes a bit of redoing to get the lift away feature to work but so what? I think this is a wonderful product. Good by Hoover, Eureka and Sears. I will keep my tank vacuum for special tasks, but the Shark is the go to choice.',
- "Absolutely impressed with this vacuum. The suction is so very impressive compares to my store-bought vacuum. *It's light and it glides and reaches every crevice. It was easy to put together*. And comes with a bag to put the

other parts. *Lots of attachments for dog hair* and cool features to detach and clean hard to reach high places. I also love the feature where I can turn the switch to bare floor and roll right onto my wood paneled floor and get up the dirt. There is also an attachment with two washable pads for dusting bare floors! What more could you ask for? You couldn't. This vacuum has it all covered!! I love my shark!