Generating Attractive and Authentic Copywriting from Customer Reviews

Yu-Xiang Lin
Academia Sinica
yuxianglin@iis.sinica.edu.tw

Wei-Yun Ma*
Academia Sinica
ma@iis.sinica.edu.tw

Abstract

The goal of product copywriting is to capture the interest of potential buyers by emphasizing the features of products through text descriptions. As e-commerce platforms offer a wide range of services, it's becoming essential to dynamically adjust the styles of these auto-generated descriptions. Typical approaches to copywriting generation often rely solely on specified product attributes, which may result in dull and repetitive content. To tackle this issue, we propose to generate copywriting based on customer reviews, as they provide firsthand practical experiences with products, offering a richer source of information than just product attributes. We have developed a sequence-to-sequence framework, enhanced with reinforcement learning, to produce copywriting that is attractive, authentic, and rich in information. Our framework outperforms all existing baseline and zero-shot large language models, including LLaMA-2-chat-7B and GPT-3.5[1], in terms of both attractiveness and faithfulness. Furthermore, this work features the use of LLMs for aspect-based summaries collection and argument allure assessment. Experiments demonstrate the effectiveness of using LLMs for marketing domain corpus construction. The code and the dataset is publicly available at: https://github.com/YuXiangLin1234/Copywriting-Generation.

1 Introduction

Copywriting is essential in E-commerce, as it assists online stores in showcasing and marketing their products to prospective buyers considerably. Research has shown that systems capable of automatically generating persuasive copywriting can significantly improve performance metrics on e-commerce platforms (Zhang et al., 2022; Guo et al., 2022a,b), reflecting the value of the automated copywriting creation solutions.

Typical methods of copywriting composition tend to focus only on the predefined attributes of a product (Wang et al., 2017; Zhang et al., 2022; Guo et al., 2022a,b), leading to content that might be monotonous and unengaging. Instead, testimonials have been proven to induce greater interest among readers (Zulkifly and Firdaus, 2014), indicating the effectiveness of copywriting incorporating diverse and authentic customer use cases. Inspired by these viewpoints, our approach emphasizes creating engaging and reliable copywriting from customer reviews. These reviews provide valuable, firsthand insights about the product, thus offering a more detailed and informative perspective than mere product attributes.

To boost the persuasiveness of produced narration, our work endeavors to integrate various dimensions such as factual accuracy, allure in sentence composition, and the richness of information in copywriting. Achieving these objectives simultaneously in a supervised manner presents significant challenges. In contrast, reinforcement learning (RL) algorithms have demonstrated proficiency in handling complex tasks, particularly in the areas of summarization (Stiennon et al., 2020) and text style transfer (Gong et al., 2019). Our work represents a synthesis of these two tasks, driving us to leverage RL techniques to attain the goal.

We apply RL algorithms with three dedicated but complementary reward models, responsible for attractiveness, faithfulness, and information density, respectively. For attractiveness, we propose to learn a sentence allure evaluator from pairwise comparisons adjudicated by GPT-3.5[2]. Our experiments indicate that using the win rate as a metric from these pairwise comparisons to train a simple regression model aligns more closely with ground truth compared to the binary classification model, which is the method typically used in studies focus-

*Corresponding Author
[1]In this work, we use gpt-3.5-turbo-0613.
[2]https://chat.openai.com/

Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), pages 4629–4642
June 16-21, 2024

ing on learning human preferences (Gleize et al., 2019; Ziegler et al., 2019; Stiennon et al., 2020).

Besides, we find that focusing solely on attractiveness in sequence generation might compromise the fidelity, which has been observed in attractive headline generation studies (Song et al., 2020; Chen et al., 2023). To address this, we incorporate a textual entailment model to improve the faithfulness of produced copywriting. Also, to enrich the information of created copywriting, we employ a natural language inference model fine-tuned on the extensive QNLI dataset (Wang et al., 2018), helping to promote the meaningful content and avoid the appearance of gorgeous but meaningless arguments. Our experiments demonstrate that our framework surpasses the baseline in performance, effectively balancing attractiveness, faithfulness, and information richness in the produced content.

In sum, our contributions are mainly threefold:

- Our framework introduces an innovative approach to creating compelling, genuine, and information-dense copywriting. This method derives content from off-the-shelf customer reviews, instead of relying on seller-provided product materials.
- We discover that optimizing our reward model based on win rates achieves greater congruence with human judgment compared to binary classification optimization.
- We develop a new dataset focused on restaurant review summarization equipped with allure assessment scores with the help of GPT-3.5, reflecting the effectiveness of using LLMs for marketing domain dataset construction.

2 Related Work

2.1 Copywriting Generation

Previous work generates product copywriting based on product attributes or descriptions (Wang et al., 2017; Zhang et al., 2022; Guo et al., 2022b). Xiao and Munro (2019) treat the product title generation task as a named entity recognition problem. Chen et al. (2019) build an encoder-decoder framework to generate personalized product descriptions by integrating information about product aspects, user categories, and external knowledge base.

Besides, some work aims to mine the opinions among customer reviews. Wu et al. (2016) introduces a CNN-based classifier to select reviews based on pre-defined aspects and group them by sentiment. Akhtar et al. (2017) extracts sentences in reviews based on similarity with given keywords by Latent Dirichlet Allocation modeling (Blei et al., 2003) but does not focus on the attractiveness. Compared to existing work, we first build an automatic copywriter to create attractive and authentic copywriting, based on customer reviews.

2.2 Attractive Arguments Generation

Many studies have been conducted with the aim of enhancing the appeal of generated sequences. Xu et al. (2019) train a CNN-based sensationalism scorer to make Pointer Generator (See et al., 2017) create sensational headlines. Song et al. (2020) train a popularity predictor from the click of view and take the popularity score and ROUGE-L score (Lin, 2004) to improve the model. A follow-up study claims that the number of clicks may be affected by trending topics, making click rate not a suitable popularity indicator (Chen et al., 2023). For example, news about politicians may receive much interest during the election. They alleviate this issue by introducing a writing technique called forward-reference (FR) (Blom and Hansen, 2015) into the headline generator. FR aims to motivate the reader's curiosity so that they want to learn more, which is also implicitly employed in earlier research on generating appealing headlines in the form of questions (Zhang et al., 2018).

Despite their remarkable achievement, the concepts of FR cannot be directly applied to our use case. The writing technique aims to capture the reader's attention, while what we focus on is persuasive copywriting to attract potential buyers. In fact, the most common popularity indicator coming with customer reviews is the star rating. However, a single review may encompass opinions of multiple products or aspects, with the star ratings generally reflecting the overall satisfaction of customers with the shop, making it an inappropriate attractiveness evaluator for individual aspect.

2.3 Aspect-based Summarization

Text summarization has been a focal point of NLP research for a long time. Numerous researchers dig into this field and develop multifarious methods to sum up long documents automatically (Tas and Kiyani, 2007; Allahyari et al., 2017). Aspect-based summarization is usually considered when extracting information tailored to different aspects from customer feedback (Hu and Liu, 2004). Summarizing reviews in an abstract fashion may be more appropriate due to the colloquial format and potential spelling errors in customer feedback. Frermann and Klementiev (2019) modify the attention mechanism to ensure the model focuses on the information about target keywords. CTRLsum (He et al., 2022) builds a framework for automatic keyword extraction during summarization . They find that the pretending aspect of the source document generates summaries related to this aspect.

Considering that a single review may touch upon numerous aspects, there's a need for aspect-focused copywriting, which is particularly significant in certain scenarios (such as discussing "dishes" in the context of restaurant reviews from Yelp). This requirement for aspect-specific copywriting can be viewed as an unique use case of aspect-based summarization technology. In this work, we adopt the method mentioned in CTRLsum for aspect-based control.

3 Dataset

Regarding the abstractive summarization benchmark, previous work has predominantly focused on the CNN/Daily Mail dataset (Nallapati et al., 2016) and the XSum dataset (Narayan et al., 2018). These datasets comprise articles on a broad spectrum of topics and lack relevance to business content, making them unsuitable for our purposes. Additionally, there are lots of existing datasets comprised of real-world customer reviews. Yelp platform provides a large-scale dataset[3] encompassing customer feedback toward a wide range of business entities. Amazon platform offers multi-language review collections sourced from their e-commerce platform (Keung et al., 2020). Nonetheless, these datasets are primarily composed of reviews without corresponding summaries. The absence of a corpus poses a challenge in producing marketing content.

Recently, numerous studies show that summaries generated by LLMs such as GPT-3.5 or LLaMA-2 (Touvron et al., 2023), are on par with human-produced summaries (Goyal et al., 2023; Zhang et al., 2023; Bang et al., 2023). Also, Yang et al. (2023) shows the remarkable capabilities of GPT-3.5 to conduct text summarization conditioned on given keywords. Built upon these works, we leverage GPT-3.5 to address the lack of corpus.

First, we sample 3622 restaurant reviews from Yelp dataset and then leverage the off-the-shelf

[3]https://www.yelp.com/dataset

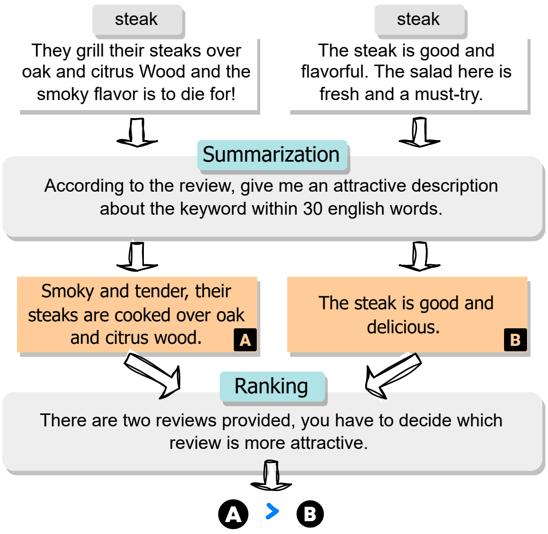

Figure 1: The illustration for dataset construction. For simplicity, the requirements of the response format are skipped. We initially replaced "description" with "summary", resulting in a collection of monotonous text.

keyphrase extraction tool[4] to draw out aspects. The tool is built on the methods utilizing contextualized embeddings (Sahrawat et al., 2020; Kulkarni et al., 2021). After that, we ask GPT-3.5 to write summaries based on specified aspects and individual customer reviews. These samples are split into training/validation/testing sets in the ratio of 7:1:2.

After generating summaries of reviews, we prompt GPT-3.5 to rank summaries regarding the same aspect. We separate arguments belonging to different splits, ensuring that there was no comparison between two different groups. We find some comparison results violate transitive relation, so a cycle breaking technique (Sun et al., 2017) is introduced to clean the dataset. The final version of the ranking dataset consists of 22602/408/1292 samples in training/dev/testing set. The pipeline for dataset construction is illustrated in Figure 1, with the prompts we used for GPT-3.5.

4 Methodology

As previous work indicates the proficiency of RL algorithms in handling complex tasks (Stiennon et al., 2020; Gong et al., 2019), especially the success of Reinforcement Learning from Human Feedback (RLHF) (Ziegler et al., 2019; Stiennon et al., 2020; Bakker et al., 2022; Ouyang et al., 2022), we leverage RL to enhance our model, aiming to align the created copywriting to human preference.

[4]https://huggingface.co/ml6team/keyphrase-extraction-distilbert-inspec

A typical pipeline of RLHF comprises 3 steps: (1) supervised fine-tuning, (2) reward modeling, and (3) reinforcement learning. We follow these steps as previous work suggests.

4.1 Supervised Fine-tuned Copywriter

We begin with a pre-trained BART-Large model (Lewis et al., 2019). To steer the model toward a proficient copywriter, we fine-tune model θ by maximizing the likelihood of reference summaries $\hat{y}$ conditioned on source review x and designated aspect k. That is, minimizing the cross-entropy loss at each decoding step as

$$L(\theta) = -\frac{1}{n}\sum_{i=1}^{n} \log p(\hat{y}_i|\hat{y}_{1:i-1}, x, k, \theta). \quad (1)$$

4.2 Reward Modeling

To further enhance the quality of generated copywriting through reinforcement learning, we leverage the following reward models to conduct argument assessment from multiple perspectives:

Allure Reward First, we build an allure evaluator to gauge the attractiveness of generated copywriting y conditioned on the given aspect k. Conceptually, allure assessment determines the attractiveness of arguments based on the writing style, word choice, and persuasiveness.

We learn the allure reward model from binary annotation from GPT-3.5 mentioned in Section 3. Related work takes learning pairwise comparison as a binary classification task (Gleize et al., 2019; Stiennon et al., 2020). Inspired by them, we initially built a Siamese network as our allure reward model and optimized it using cross-entropy classification loss, but found that this baseline did not yield ideal results on the testing set. Hence, we turn to compute the win-rate w_i from all M pairwise annotation $\hat{y}_i$ engage in, as the alluring score of $\hat{y}_i$. The formula is shown as

$$w_i = \frac{\Sigma_{j=1}^{M} v_{ij}}{M}, \quad (2)$$

where $v_{ij} = 1$ if $\hat{y}_i$ is more attractive than $\hat{y}_j$ and otherwise 0 .

We then learn a regression model to predict the win rate and rank arguments according to their win rate, from highest to lowest. The method significantly improves binary classification accuracy and reduces root mean square error (RMSE). The experimental results are shown in Section 5.1. We adopt this regression model as our allure reward model R_a, and take the predicted win rate as the allure reward

$$r_a = R_a(k, y) \quad (3)$$

Veracity Reward Previous work shows there is a high correlation between entailment and faithfulness (Falke et al., 2019; Chen et al., 2023). Inspired by them, we leverage a DeBERTa-v3 model (Laurer et al., 2023) trained on 33 datasets such as WANLI (Liu et al., 2022) and MultiNLI (Williams et al., 2018), and thus owning rich knowledge and remarkable zero-shot capabilities, to serve as our veracity reward model. In this work, we adopt the model provided by the authors[5]. Since the model outputs two values representing "entailment" and "not entailment," we choose the output logits of the "entailment" head as our veracity reward

$$r_v = R_v(x, y). \quad (4)$$

The statement refers to estimating whether review x logically entails output y, with an objective to compose copywriting that accurately reflects the source review without any misrepresentation.

Information Reward Intuitively, the copywriter optimized with the above two rewards might generate gorgeous but pointless template-based arguments such as "The food is delicious and perfect." or "The food is a mouthwatering delight that will leave you craving for more." To tackle this issue, we encourage the copywriter to provide more information, like commendable characteristics or the difference between similar products on the market. We leverage a model trained on a large-scale dataset named QNLI (Wang et al., 2018) to attain the purpose. QNLI dataset consists of question-paragraph pairs, and labels indicate whether the answer to the question can be found in the paragraph. With some common facets of restaurant reviews, we query the QNLI model to check whether information about these facets is in the generated copywriting. Additionally, to ensure the fidelity to the source review, we also check what the fact review contains. The information reward is formulated as

$$r_i = \frac{|\Sigma_{f \in F} R_i(f, k, x) \wedge R_i(f, k, y)|}{|\Sigma_{f \in F} R_i(f, k, y)|}, \quad (5)$$

where $|\cdot|$ means the counting function, F refers to the set of queries about different facets, such as

[5] https://huggingface.co/MoritzLaurer/deberta-v3-large-zeroshot-v1.1-all-33

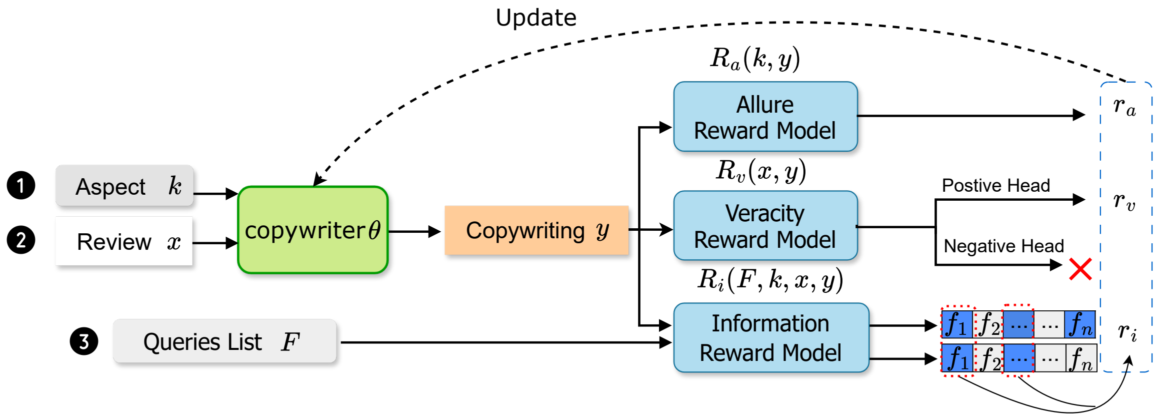

Figure 2: The illustration of the proposed framework. Given an aspect k, a review x, and a list of queries F, we enhance the supervised finetuned model θ by the Proximal Policy Optimization (PPO) algorithm with three dedicated reward models.

"How does k look?" or "How does k compare to similar ones on the market?". Our queries are provided in Appendix B. Conceptually, if the information is included in source documents, it should be in generated summaries. In this case, there might not be a difference between an attractive sentence and an unappealing one. But with the help of the other two RMs, we are able to create a proficient copywriter producing persuasive copywriting.

4.3 Reinforcement Learning

With the supervised fine-tuned model θ, the allure RM R_a, the veracity RM R_v, and the information RM R_i, we further enhance θ with the Proximal Policy Optimization (PPO) algorithm (Schulman et al., 2017), as illustrated in Figure 2. PPO shows the remarkable capability of guiding the policy toward a desired direction with high sampling efficiency. We treat θ as policy and optimize θ as

$$\mathbb{R}(x, y, k, F) = \alpha r_a + \beta r_v + \gamma r_i - \log \frac{\theta(y|x, k)}{\theta_{old}(y|x, k)}, \tag{6}$$

where the last term refers to a penalty, θ_{old} is the original version of θ before RL optimization. The KL-divergence-based penalty is meant to prevent θ moving too far from θ_{old} and thus stabilize the training. α, β and γ are weight of different rewards. For simplicity, we set α, β and γ to be 1 in all following experiments.

5 Experiments

All experiments are conducted on an Nvidia Tesla-V100 GPU. We select the model that performs the best on a validation set. The implementation details are listed in Appendix A.

5.1 Allure Reward Model

We initially built a Siamese network as our allure reward model and optimized it using cross entropy classification loss as suggested in related work (Gleize et al., 2019; Stiennon et al., 2020). However, we find that this baseline does not yield ideal results on the testing set. We turn to learn a regression model with the win rate for each argument and get a better result. The experiments are conducted on a BART-base model (Lewis et al., 2019) and DeBERTa-v3-base model (He et al., 2021), with searching learning rate in $[1e-5, 3e-5, 5e-5, 7e-5, 9e-5]$. we report the best results about root mean square error (RMSE) and binary classification accuracy on testing set in Table 1.

It's reasonable that the regression model exhibits a lower RMSE metric than the Siamese network since RMSE is computed based on win rate. However, the gap between these two types of models is too large despite the existence of a sigmoid layer. Surprisingly, the accuracy of the DeBERTa-v3 regression model is much higher than that of the Siamese network. According to the experiments, we select the DeBERTa-v3 regression model, which exhibits the best performance, as our allure reward model in the following experiments.

5.2 Baselines for Copywriting Generation

We compare our proposed framework to the following baselines. We choose works built on a BART-large model to alleviate the effect of different backbones. Summaries in our dataset are both attractive and faithful to source review inherently since we prompt GPT-3.5 to generate such samples. Hence, baseline fine-tuned on them should highlight their

	Model	Acc. ↑	RMSE ↓
BART	Siamese Network	66.25 %	0.52
	Regression Network	66.25 %	**0.37**
DeBERTa	Siamese Network	65.09 %	0.54
	Regression Network	**69.50** %	**0.37**

Table 1: Experiment about allure reward model

strong abilities to generate attractive and authentic copywriting. We first compare our model to the supervised fine-tuning (SFT) model for checking the effect of reinforcement learning. Besides, an aspect-based summarization model called CTRL-sum (He et al., 2022) shows good controllability to paraphrase the sentence based on given aspects. Some work utilizes ROUGE score to enhance the quality of the abstract summarization and achieve notable success (Pasunuru and Bansal, 2018; Paulus et al., 2018), making us put the model optimized by ROUGE reward into comparison. We compute ROUGE-L score as a reward for training this baseline, as suggested in their paper.

We also incorporate LLMs in comparison since LLMs demonstrate remarkable zero-shot capabilities on multiple tasks (Zhang et al., 2023; Goyal et al., 2023). Given a keyword k and source review x, we prompt LLaMA-2-chat-7B (Touvron et al., 2023) to generate attractive and faithful summaries, serving as one of the baselines. The prompt employed here is the same as the summarization part in Figure 1 and adjusted to the suggested format of LLaMA-2-Chat. The last baseline is summaries generated by GPT-3.5. Note that they are ground truth for supervised fine-tuning.

5.3 Human evaluation

Since we use the large language model for dataset construction, it's more reliable for conducting human evaluation. In this experiment, we randomly sample 100 customer reviews and the corresponding copywriting from the testing set for comparison. Given a customer review x, an aspect k, and two copywriting y_1, y_2, The instructions for evaluators are (1) Which one is more attractive and makes you want to purchase/enjoy k? (2) Which one is more faithful to the original review x? (3) Which one is more fluent? The evaluators are asked to select *first copywriting*, *second copywriting*, or *tie* for each pair and each question. We give \$0.05 for each pair of data because of the simplicity of annotation. We use evaluators in the United States and the United Kingdom, with a greater than 95% approval rate.

The evaluation protocol follows previous work (Zhao et al., 2020; Chen et al., 2023) to evaluate the generated copywriting's attractiveness, faithfulness, and fluency by pairwise comparison. Results are reported in Table 2. For example, the copywriting generated by our framework is 55.33%/34.33%/10.33% better than/worse than/equal to the output of the SFT model on faithfulness. The value of faithfulness in the table thus becomes 34.33% - 55.33% = -21%. Results show that our framework outperforms all baselines, including zero-shot LLMs and fine-tuned BART model, in terms of attractiveness and faithfulness, with comparable fluency. This implies the positive impact of RL finetuning. The results also reflects the controllability of our framework because the judgment criteria for attractiveness are based on the given aspect.

	Model	ATRC ↑	FAITH ↑	FLCY ↑
LLM	GPT-3.5	-10.0%	-3.3%	3.0%
	LLaMA-2-7B-Chat	-25.0%	-19.3%	3.3%
BART	SFT	-15.0%	-21.0%	-13.3%
	CTRLsum	-17.0%	-5.3%	-2.7%
	ROUGE	-13.7%	-4.7%	3.0%
	Ours	-	-	-

Table 2: Human evaluation results of pairwise comparison. We compare copywriting generated by our framework and baseline models in terms of attractiveness (ATRC), faithfulness (FAITH), and fluency (FLCY).

5.4 Automatical Metrics

ROUGE Score ROUGE score (Lin, 2004) is a conventional metric to measure the quality of text summarization. It counts word-level overlap to estimate the similarity between the generated summary and the reference one. ROUGE score has been shown to be unconvincing in abstractive summarization task (Schluter, 2017) . Our framework aims to build a copywriter capable of generating attractive copywriting. Using the ROUGE score to measure the performance is inappropriate, but we still provide ROUGE-1, ROUGE-2, ROUGE-L in Table 3 for reference. Most baseline models exhibit comparable ROUGE scores, except for LLaMA-2-7B-Chat, which may be attributed to the zero-shot generation. SFT achieves the highest ROUGE score, which is reasonable since it closely imitates the writing style of reference summaries, although it yields the worst results comprehensively. Interestingly, the model optimized by the ROUGE reward exhibits a similar ROUGE score with other baselines despite the increasing ROUGE reward during the training process. We think the reason is the broad knowledge GPT-3.5 owns, and the dataset is constructed in a zero-shot manner, resulting in a different distribution between the training and testing sets.

	Model	R_1	R_2	R_L
LLM	GPT-3.5 (gold)	-	-	-
	LLaMA-2-7B-Chat	18.50	3.55	15.00
BART	SFT	**28.25**	**8.13**	**23.94**
	CTRLsum	27.22	7.52	23.30
	ROUGE	27.26	7.63	23.40
	Ours	27.28	7.34	23.18

Table 3: ROUGE score computed with reference summary generated by GPT-3.5. R_1, R_2, R_L are the ROUGE-1, ROUGE-2, ROUGE-L score respectively.

Perplexity Perplexity (PPL) is commonly used to evaluate text fluency (Wang and Wan, 2018; Dathathri et al., 2020). The metric refers to the likelihood of pre-trained language models (PLMs) to produce the sequence. Lower PPL means better fluency of the sentences. We compute PPL in Table 4 with two publicly available PLMs, GPT-2 (Radford et al., 2019) and LLaMA-2-7B (Touvron et al., 2023). Results show that our framework achieves the lowest PPL, the highest fluency.

	Model	GPT-2	LLaMA-2-7B
LLM	GPT-3.5	67.10	32.59
	LLaMA-2-7B-Chat	71.88	46.36
BART	SFT	49.75	26.86
	CTRLsum	48.93	26.89
	ROUGE	52.05	26.10
	Ours	**48.63**	**25.23**

Table 4: Perplexity (PPL) computed by GPT-2 and LLaMA-2-7B. Lower PPL means better fluency.

5.5 Ablation Study

We conduct the following ablation analyses to probe the impact of components in our framework.

Human Evaluation We repeat the experiments in Section 5.3. Since the enhancement of RL over the SFT model has been shown in previous experiments, we learn three models removing allure RM R_a, veracity RM R_v, and information RM R_i in this section. Results are demonstrated in Table 5. Intuitively, models optimized without R_a and R_v yield the degradation of attractiveness and faithfulness. The information reward model seems to be helpful for faithfulness. We claim that the purpose of R_i is to increase the information density of generated copywriting that the source document mentioned. Hence, with the help of R_v and R_i, copywriters tend to offer customers more accurate information about different facets. In this case, copywriting with many "facts" will be more faithful to that with only empty words.

	ATRC ↑	FAITH ↑	FLCY ↑
Ours	-	-	-
W/o R_a	-14.0%	-0.0%	-5.7%
W/o R_v	-8.3%	-14.0%	-6.7%
W/o R_i	2.0%	-8.7%	-2.3%

Table 5: Ablation study with human evaluation in terms of attractiveness (ATRC), faithfulness (FAITH), fluency (FLCY).

Information Score In Table 6, we compute the information score of generated summaries in the testing set as Equation 5 by QNLI model. The experiment is to further confirm the impact of the information reward model R_i. Additionally, whether the information score is proportional to the sentence length is questionable. We write down the mean and standard deviation of unigram counts for each baseline. All baseline models generated copywriting within 30 words, as required in Section 3 for generated SFT training data by GPT-3.5. Results show that our framework achieves the highest information reward with around 3 words longer. Instead, the model optimized without information reward produces longer sentences on average but with the lowest information score, verifying the significance of R_i.

5.6 Qualitative Results

As an example, we randomly select a sample with two aspects in the testing set. Table 9 shows the source review and corresponding copywriting. Due to the page limit, we only list the output of the two strongest baselines in the above experiments, GPT-3.5 and CTRLsum. Copywriting generated

	Model	Avg.	Std.	INFO ↑
LLM	GPT-3.5	19.09	6.34	56.45
	LLaMA-2-7B-Chat	17.03	12.04	55.42
BART	SFT	19.14	5.81	57.57
	CTRLsum	15.71	4.81	56.79
	ROUGE	20.55	7.38	56.48
	Ours	22.57	8.00	**61.06**
	W/o R_i	23.28	8.55	56.48

Table 6: Length of generated sentences and information score. Avg. and Std. refer to the average counts of uni-grams and the standard deviation of sentence length, respectively. INFO refers to the information score computed by QNLI model.

by other baselines is listed in Table 10. For both aspects, "Steak" and "Tampa," we can see our copywriter provides more information mentioned in the source review, such as "Sides were veggies and sweet potatoes" and "delicious veggies and sweet potatoes." And for the aspect "Tampa", GPT-3.5 and CTRLsum say that there is a diverse option for meals, but our framework lists what the dining options include. For attractiveness, we find that GPT-3.5 tends to produce gorgeous but advanced words, making it hard to attract turker at first glance. This might be the reason why our framework performs better than GPT-3.5 regarding attractiveness and faithfulness.

6 Comparisons with Human-written Copywriting

We have taken the following steps to bolster the experiments with some human-written copywriting to address concerns regarding using LLMs for training data collection and the word length constraints applied to LLM baselines.

We employ Amazon Mechanical Turk workers to create copywriting samples. Specifically, we gather 100 human-written copywriting pieces from testing samples used in Section 5.3. To ensure the quality of these human-generated samples, we meticulously filter out 20 significantly inappropriate or off-topic responses. The instructions given to the turkers are as follows: *Please write an attractive but faithful summary based on the given keyword and customer review. The summary will be used as product copywriting. You should paraphrase the sentence to make it more attractive and attract potential buyers to buy/enjoy this keyword. But it should be faithful to the original review.*

6.1 Human Evaluation

After incorporating the human-written summaries, we conduct a human evaluation with these samples regarding three key aspects, the same as conditions outlined in Section 5.3. The comparative results are shown in Table 7, showcasing that copywriting generated by our framework is superior quality compared to human writers.

	ATRC ↑	FAITH ↑	FLCY ↑
Human-written	-2.09%	-0.81%	-6.25%
Ours	-	-	-

Table 7: Human evaluation results of pairwise comparison. We compare copywriting generated by our framework and human-written one in terms of attractiveness (ATRC), faithfulness (FAITH), and fluency (FLCY).

6.2 Will Language Models Generate Sentences Humans Will Produce?

To further alleviate the concern about the ability of LLMs to produce copywriting as persuasive and attractive as human authors, we employ Fast-detectgpt (Bao et al., 2023), a state-of-the-art toolkit for machine-generated text detection, to verify the human-like quality of the copywriting generated by our framework. Fast-detectgpt is proficient in detecting machine-generated text in a zero-shot manner, making it an ideal tool for this assessment.

Adopting Fast-detectgpt, we analyze the copywriting produced by our framework to determine the likelihood of it being identified as machine-generated. The results presented in Table 8 indicate low probabilities of the text being flagged as machine-generated or 'Fake.' This implies that the text generated by our model closely resembles human-written content, both in style and substance.

	probability ↓
Human-written	**33.22**%
Ours	37.44%

Table 8: probabilities of the text are machine-generated

7 Conclusion

In this paper, we present an innovative framework with an elegant but effective reward mechanism, designed to automatically generate attractive and

<table>
<tr><th>Aspect</th><th>Steak</th><th>Tampa</th></tr>
<tr><td>Source Review</td><td colspan="2">We got diverted to Tampa and decided to make the drive to Orlando, but first we needed somewhere to eat. Wanting to avoid chain restaurants, yelp steered us toward this little local gem! They grill their steaks over oak, and citrus Wood and the smoky flavor is to die for! We are big fans of our smoker back home and cook our steaks this way as well. It's rare to find steak cooked over woodsmoke in most places, and this was worth the side trip! The sides were veggies and sweet potatoes. The wines on offer were so reasonably priced - most under $20 a bottle, but a decent selection. They just don't mark them up 5x like most places! Bookmarking this place for a return trip next time I am in Tampa for sure!</td></tr>
<tr><td>GPT-3.5</td><td>Juicy, tender, and perfectly grilled, this oak and citrus wood-smoked steak delivers a mouthwatering smoky flavor. Cooked to perfection, it's a rare gem worth savoring, a true delight for steak enthusiasts.</td><td>An enchanting city in Florida known for its diverse dining options, Tampa captivates visitors with its local gems and the tantalizing aroma of woodsmoked steaks</td></tr>
<tr><td>CTRLsum</td><td>Mouthwatering, tender, and smoky steaks grilled over oak and citrus Wood, bursting with a delectable flavor that will leave you craving for more.</td><td>Vibrant city with diverse culinary options, including a local gem serving smoky steaks and delicious sides.</td></tr>
<tr><td>Ours</td><td>Juicy, smoky, and grilled over oak and citrus wood, our steaks at this local gem in Tampa are truly exceptional. A rare delight cooked over woodsmoke, it's worth the side trip! The sides were veggies and sweet potatoes.</td><td>Vibrant Florida city with smoky steaks, affordable wines, delicious veggies and sweet potatoes. A must-visit!</td></tr>
</table>

Table 9: The generated examples of different models. Due to the page limit, we provide the copywriting generated by each baseline model in Appendix C.

faithful copywriting. Extensive experiments are conducted to verify the superior capabilities of the proposed method, compared to existing baselines including zero-shot Large Language Models (LLMs), fine-tuned BART models, and even human writers. Moreover, our findings reveal that optimizing the reward model using win-rate leads to a higher alignment with human judgment than the traditional binary classification optimization method. This approach enhances the model's ability to reflect human preferences more accurately.

In addition to these developments, we have created an unique dataset comprising summaries and corresponding ranking data, which we believe will be valuable for future research in this field. Our dataset is built with the help of GPT-3.5 and includes curated content specifically tailored to advancing the capabilities of automatic copywriting systems, demonstrating the effectiveness of using LLMs for marketing domain dataset construction. Through this work, we aspire to expand the potential and effectiveness of E-commerce platforms, offering new avenues for engaging and authentic content generation.

Scientific Artifacts

All experiments are conducted upon PyTorch (Paszke et al., 2019), and all pretrained models other than GPT-3.5 are obtained from HuggingFace (Wolf et al., 2019). We also use NLTK (Bird et al., 2009) package for word tokenization and sentence tokenization. We adopt these artifacts for their intended use.

Limitation and Potential Risks

Our framework is not applicable when customer reviews are lacking, especially when the product is unavailable on the market. The proposed approach underscores our key point: generating copywriting from customer reviews is a novel and valuable direction. Just as pre-release copywriting sourced seller-defined attributes has merits, post-release revisions based on actual customer experiences can provide a wealth of information beyond mere product attributes. Both pre-release generation and post-release refinement of copywriting are essential, complementing each other in effectively marketing a product.

Also, despite our framework exhibiting the best attractiveness and faithfulness, there is still room for improvement. For attractiveness, the binary classification accuracy and the agreement with the ground truth of our allure model are around 70 percent. For faithfulness, our framework produces some fake information once in a while. We believe these issues can be alleviated by developing a more accurate reward evaluator. Building upon the rapid development of LLMs, Kwon et al. (2023) prompts LLMs to compute reward, which might be the direction of our future work.

Acknowledgements

We are grateful for the insightful and valuable comments from anonymous reviews. Thanks also to Yin-Hsiang Liao for the preliminary discussions. This work is supported by the National Science and Technology Council of Taiwan under grant numbers NSTC112-2221-E-001-025. Thanks to the National Center for High-performance Computing (NCHC) for providing computational and storage resources.

References

Nadeem Akhtar, Nashez Zubair, Abhishek Kumar, and Tameem Ahmad. 2017. Aspect based sentiment oriented summarization of hotel reviews. *Procedia Computer Science*, 115:563–571. 7th International Conference on Advances in Computing Communications, ICACC-2017, 22-24 August 2017, Cochin, India.

Mehdi Allahyari, Seyedamin Pouriyeh, Mehdi Assefi, Saeid Safaei, Elizabeth D. Trippe, Juan B. Gutierrez, and Krys Kochut. 2017. Text summarization techniques: A brief survey. *International Journal of Advanced Computer Science and Applications*, 8(10).

Michiel Bakker, Martin Chadwick, Hannah Sheahan, Michael Tessler, Lucy Campbell-Gillingham, Jan Balaguer, Nat McAleese, Amelia Glaese, John Aslanides, Matt Botvinick, and Christopher Summerfield. 2022. Fine-tuning language models to find agreement among humans with diverse preferences. In *Advances in Neural Information Processing Systems*, volume 35, pages 38176–38189. Curran Associates, Inc.

Yejin Bang, Samuel Cahyawijaya, Nayeon Lee, Wenliang Dai, Dan Su, Bryan Wilie, Holy Lovenia, Ziwei Ji, Tiezheng Yu, Willy Chung, Quyet V. Do, Yan Xu, and Pascale Fung. 2023. A multitask, multilingual, multimodal evaluation of chatgpt on reasoning, hallucination, and interactivity.

Guangsheng Bao, Yanbin Zhao, Zhiyang Teng, Linyi Yang, and Yue Zhang. 2023. Fast-detectgpt: Efficient zero-shot detection of machine-generated text via conditional probability curvature. *arXiv preprint arXiv:2310.05130*.

Steven Bird, Ewan Klein, and Edward Loper. 2009. *Natural language processing with Python: analyzing text with the natural language toolkit*. " O'Reilly Media, Inc.".

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

Jonas Nygaard Blom and Kenneth Reinecke Hansen. 2015. Click bait: Forward-reference as lure in online news headlines. *Journal of Pragmatics*, 76:87–100.

Chih Yao Chen, Dennis Wu, and Lun-Wei Ku. 2023. HonestBait: Forward references for attractive but faithful headline generation. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 4810–4824, Toronto, Canada. Association for Computational Linguistics.

Qibin Chen, Junyang Lin, Yichang Zhang, Hongxia Yang, Jingren Zhou, and Jie Tang. 2019. Towards knowledge-based personalized product description generation in e-commerce. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '19, page 3040–3050, New York, NY, USA. Association for Computing Machinery.

Sumanth Dathathri, Andrea Madotto, Janice Lan, Jane Hung, Eric Frank, Piero Molino, Jason Yosinski, and Rosanne Liu. 2020. Plug and play language models: A simple approach to controlled text generation. In *International Conference on Learning Representations*.

Tobias Falke, Leonardo F. R. Ribeiro, Prasetya Ajie Utama, Ido Dagan, and Iryna Gurevych. 2019. Ranking generated summaries by correctness: An interesting but challenging application for natural language inference. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 2214–2220, Florence, Italy. Association for Computational Linguistics.

Lea Frermann and Alexandre Klementiev. 2019. Inducing document structure for aspect-based summarization. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 6263–6273, Florence, Italy. Association for Computational Linguistics.

Martin Gleize, Eyal Shnarch, Leshem Choshen, Lena Dankin, Guy Moshkowich, Ranit Aharonov, and Noam Slonim. 2019. Are you convinced? choosing the more convincing evidence with a Siamese network. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 967–976, Florence, Italy. Association for Computational Linguistics.

Hongyu Gong, Suma Bhat, Lingfei Wu, JinJun Xiong, and Wen-mei Hwu. 2019. Reinforcement learning based text style transfer without parallel training corpus. 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 3168–3180, Minneapolis, Minnesota. Association for Computational Linguistics.

Tanya Goyal, Junyi Jessy Li, and Greg Durrett. 2023. News summarization and evaluation in the era of gpt-3.

Xiaojie Guo, Shugen Wang, Hanqing Zhao, Shiliang Diao, Jiajia Chen, Zhuoye Ding, Zhen He, Jianchao Lu, Yun Xiao, Bo Long, et al. 2022a. Intelligent online selling point extraction for e-commerce recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 12360–12368.

Xiaojie Guo, Qingkai Zeng, Meng Jiang, Yun Xiao, Bo Long, and Lingfei Wu. 2022b. Automatic controllable product copywriting for e-commerce. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 2946–2956.

Junxian He, Wojciech Kryscinski, Bryan McCann, Nazneen Rajani, and Caiming Xiong. 2022. CTRLsum: Towards generic controllable text summarization. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 5879–5915, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Pengcheng He, Jianfeng Gao, and Weizhu Chen. 2021. Debertav3: Improving deberta using electra-style pretraining with gradient-disentangled embedding sharing.

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*, KDD '04, page 168–177, New York, NY, USA. Association for Computing Machinery.

Phillip Keung, Yichao Lu, György Szarvas, and Noah A. Smith. 2020. The multilingual Amazon reviews corpus. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 4563–4568, Online. Association for Computational Linguistics.

Mayank Kulkarni, Debanjan Mahata, Ravneet Arora, and Rajarshi Bhowmik. 2021. Learning rich representation of keyphrases from text. *arXiv preprint arXiv:2112.08547*.

Minae Kwon, Sang Michael Xie, Kalesha Bullard, and Dorsa Sadigh. 2023. Reward design with language models. *arXiv preprint arXiv:2303.00001*.

Moritz Laurer, Wouter Van Atteveldt, Andreu Casas, and Kasper Welbers. 2023. Less Annotating, More Classifying: Addressing the Data Scarcity Issue of Supervised Machine Learning with Deep Transfer Learning and BERT-NLI. *Political Analysis*, pages 1–33.

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

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

Alisa Liu, Swabha Swayamdipta, Noah A. Smith, and Yejin Choi. 2022. WANLI: Worker and AI collaboration for natural language inference dataset creation. In *Findings of the Association for Computational Linguistics: EMNLP 2022*, pages 6826–6847, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Ramesh Nallapati, Bowen Zhou, Cicero dos Santos, Çağlar Gulçehre, 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, Berlin, Germany. Association for Computational Linguistics.

Shashi Narayan, Shay B. Cohen, and Mirella Lapata. 2018. Don't give me the details, just the summary! topic-aware convolutional neural networks for extreme summarization. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 1797–1807, Brussels, Belgium. Association for Computational Linguistics.

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul F Christiano, Jan Leike, and Ryan Lowe. 2022. Training language models to follow instructions with human feedback. In *Advances in Neural Information Processing Systems*, volume 35, pages 27730–27744. Curran Associates, Inc.

Ramakanth Pasunuru and Mohit Bansal. 2018. Multi-reward reinforced summarization with saliency and entailment. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, pages 646–653, New Orleans, Louisiana. Association for Computational Linguistics.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. 2019. Pytorch:

An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc.

Romain Paulus, Caiming Xiong, and Richard Socher. 2018. A deep reinforced model for abstractive summarization. In *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*. OpenReview.net.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.

Dhruva Sahrawat, Debanjan Mahata, Haimin Zhang, Mayank Kulkarni, Agniv Sharma, Rakesh Gosangi, Amanda Stent, Yaman Kumar, Rajiv Ratn Shah, and Roger Zimmermann. 2020. Keyphrase extraction as sequence labeling using contextualized embeddings. In *Advances in Information Retrieval: 42nd European Conference on IR Research, ECIR 2020, Lisbon, Portugal, April 14–17, 2020, Proceedings, Part II 42*, pages 328–335. Springer.

Natalie Schluter. 2017. The limits of automatic summarisation according to ROUGE. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pages 41–45, Valencia, Spain. Association for Computational Linguistics.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

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, Vancouver, Canada. Association for Computational Linguistics.

Yun-Zhu Song, Hong-Han Shuai, Sung-Lin Yeh, Yi-Lun Wu, Lun-Wei Ku, and Wen-Chih Peng. 2020. Attractive or faithful? popularity-reinforced learning for inspired headline generation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(05):8910–8917.

Nisan Stiennon, Long Ouyang, Jeffrey Wu, Daniel Ziegler, Ryan Lowe, Chelsea Voss, Alec Radford, Dario Amodei, and Paul F Christiano. 2020. Learning to summarize with human feedback. In *Advances in Neural Information Processing Systems*, volume 33, pages 3008–3021. Curran Associates, Inc.

Jiankai Sun, Deepak Ajwani, Patrick K. Nicholson, Alessandra Sala, and Srinivasan Parthasarathy. 2017. Breaking cycles in noisy hierarchies. In *Proceedings of the 2017 ACM on Web Science Conference*, WebSci '17, pages 151–160, New York, NY, USA. ACM.

Oguzhan Tas and Farzad Kiyani. 2007. A survey automatic text summarization. *PressAcademia Procedia*, 5(1):205–213.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, et al. 2023. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*.

Leandro von Werra, Younes Belkada, Lewis Tunstall, Edward Beeching, Tristan Thrush, and Nathan Lambert. 2020. Trl: Transformer reinforcement learning. `https://github.com/lvwerra/trl`.

Alex Wang, Amanpreet Singh, Julian Michael, Felix Hill, Omer Levy, and Samuel R Bowman. 2018. Glue: A multi-task benchmark and analysis platform for natural language understanding. *arXiv preprint arXiv:1804.07461*.

Jinpeng Wang, Yutai Hou, Jing Liu, Yunbo Cao, and Chin-Yew Lin. 2017. A statistical framework for product description generation. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 187–192.

Ke Wang and Xiaojun Wan. 2018. Sentigan: Generating sentimental texts via mixture adversarial networks. In *IJCAI*, pages 4446–4452.

Adina Williams, Nikita Nangia, and Samuel Bowman. 2018. A broad-coverage challenge corpus for sentence understanding through inference. 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 1112–1122, New Orleans, Louisiana. Association for Computational Linguistics.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, and Jamie Brew. 2019. Huggingface's transformers: State-of-the-art natural language processing. *CoRR*, abs/1910.03771.

Haibing Wu, Yiwei Gu, Shangdi Sun, and Xiaodong Gu. 2016. Aspect-based opinion summarization with convolutional neural networks. In *2016 International Joint Conference on Neural Networks (IJCNN)*, pages 3157–3163.

Joan Xiao and Robert Munro. 2019. Text summarization of product titles. In *eCOM@ SIGIR*.

Peng Xu, Chien-Sheng Wu, Andrea Madotto, and Pascale Fung. 2019. Clickbait? sensational headline generation with auto-tuned reinforcement learning. In *Proceedings of the 2019 Conference on Empirical*

Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pages 3065–3075, Hong Kong, China. Association for Computational Linguistics.

Xianjun Yang, Yan Li, Xinlu Zhang, Haifeng Chen, and Wei Cheng. 2023. Exploring the limits of chatgpt for query or aspect-based text summarization.

Ruqing Zhang, Jiafeng Guo, Yixing Fan, Yanyan Lan, Jun Xu, Huanhuan Cao, and Xueqi Cheng. 2018. Question headline generation for news articles. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, CIKM '18, page 617–626, New York, NY, USA. Association for Computing Machinery.

Tianyi Zhang, Faisal Ladhak, Esin Durmus, Percy Liang, Kathleen McKeown, and Tatsunori B. Hashimoto. 2023. Benchmarking large language models for news summarization.

Xueying Zhang, Yanyan Zou, Hainan Zhang, Jing Zhou, Shiliang Diao, Jiajia Chen, Zhuoye Ding, Zhen He, Xueqi He, Yun Xiao, et al. 2022. Automatic product copywriting for e-commerce. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 12423–12431.

Chao Zhao, Marilyn Walker, and Snigdha Chaturvedi. 2020. Bridging the structural gap between encoding and decoding for data-to-text generation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 2481–2491, Online. Association for Computational Linguistics.

Daniel M Ziegler, Nisan Stiennon, Jeffrey Wu, Tom B Brown, Alec Radford, Dario Amodei, Paul Christiano, and Geoffrey Irving. 2019. Fine-tuning language models from human preferences. *arXiv preprint arXiv:1909.08593*.

Hani Zulaikha Zulkifly and Norsham Firdaus. 2014. Persuasion and the online consumers: Investigating copywriting strategies in native advertisements. *International Journal of Social Science and Humanity*, 4(6):430.

A Implementation Details

Allure Reward Model For the training of the allure RM, we adopt a debate-v3 model as the backbone and set the batch size to 32, with a learning rate of 2e-5. Since the regression model and Siamese network converge earlier, we train them for 5 epochs.

Information Reward Model Since the QNLI dataset is included in a well-known benchmark named GLUE (Wang et al., 2018), there are numerous off-the-shelf models provided in the huggingface platform (Wolf et al., 2019); we select one of them for all experiments[6]. The reported accuracy of this model on the QNLI dev set is 93.21%.

Reinforcement Learning We adopt the trl implementation provided by the huggingface platform (von Werra et al., 2020) for this part. The learning rate is set to be 1e-6, and the batch size is set to be 32. We fine-tune the model for 20 epochs. Although it seems that there is still room for enhancement. We observe instability in the training process after 20 epochs and might be unable to reproduce sometimes. The training procedure takes around 12 hours on a single Nvidia Tesla-V100 GPU.

B Queries for Information Reward

Given an aspect k, we define some common facets of restaurant review as following:

- What is the price of k?
- What is the cooking method of k?
- How does the k taste?
- How does the k smell?
- How does the k look?
- How is the quality of k?
- How is the seasoning of k?
- How is the portion size of k ?
- How is the cleanliness of k
- What are the components of the k?
- What is the opinion of the people about the k?
- How does k compare to similar ones on the market?

Despite the template-based queries not being suitable for all scenarios, it's reasonable to let the model handle itself because there is no "answer" to unreasonable queries in the world. According to our preliminary experiments, the QNLI model indeed owns this ability.

C Generated Examples

The examples generated by each baseline model are shown in Table 10.

[6] https://huggingface.co/cross-encoder/qnli-electra-base

Aspect	Steak	Tampa
Source Review	We got diverted to Tampa and decided to make the drive to Orlando, but first we needed somewhere to eat. Wanting to avoid chain restaurants, yelp steered us toward this little local gem! They grill their steaks over oak, and citrus Wood and the smoky flavor is to die for! We are big fans of our smoker back home and cook our steaks this way as well. It's rare to find steak cooked over woodsmoke in most places, and this was worth the side trip! The sides were veggies and sweet potatoes. The wines on offer were so reasonably priced - most under $20 a bottle, but a decent selection. They just don't mark them up 5x like most places! Bookmarking this place for a return trip next time I am in Tampa for sure!	
GPT-3.5	Juicy, tender, and perfectly grilled, this oak and citrus wood-smoked steak delivers a mouthwatering smoky flavor. Cooked to perfection, it's a rare gem worth savoring, a true delight for steak enthusiasts.	An enchanting city in Florida known for its diverse dining options, Tampa captivates visitors with its local gems and the tantalizing aroma of woodsmoked steaks
LLaMA-2-7B-Chat	Savory, smoky steak grilled over oak and citrus wood, a true delight for steak lovers.	Tampa's hidden gem for steak lovers: wood-grilled masterpieces in a cozy local setting.
SFT	Smoky, charred steak cooked over oak and citrus wood, a rare delicacy in Tampa. Worth the side trip!	Tampa: A local gem with smoky steaks over oak and citrus wood , delicious sides, affordable wines, and exceptional service. A must-visit!
CTRLsum	Mouthwatering, tender, and smoky steaks grilled over oak and citrus Wood, bursting with a delectable flavor that will leave you craving for more.	Vibrant city with diverse culinary options, including a local gem serving smoky steaks and delicious sides.
ROUGE	Juicy, smoky, and grilled over oak and citrus wood, our steaks at this local gem in Tampa are truly exceptional. A rare and mouthwatering delight cooked over smoky flavor, worth the side trip and the friendly service.	Vibrant Florida city with smoky steaks, affordable wines, and delicious sides. A must-visit destination.
Ours	Juicy, smoky, and grilled over oak and citrus wood, our steaks at this local gem in Tampa are truly exceptional. A rare delight cooked over woodsmoke, it's worth the side trip! The sides were veggies and sweet potatoes.	Vibrant Florida city with smoky steaks, affordable wines, delicious veggies and sweet potatoes. A must-visit!

Table 10: The generated examples of different models.