Preprints.org

Article

Leveraging AIGC and Human-Computer Interaction Design to Enhance Efficiency and Quality in E-commerce Content Generation

[Xue Song](#) *

Article

Leveraging AIGC and Human-Computer Interaction Design to Enhance Efficiency and Quality in E-Commerce Content Generation

Xue Song

Department of Technology - User Experience, Fanatics Inc., San Mateo, USA; xsong@fanatics.com

Abstract: In light of the accelerated growth of e-commerce, the generation of high-quality and efficient content has emerged as a pivotal factor in enhancing user experience and business value. However, conventional methods of content creation are prone to inefficiencies and creative constraints. This paper puts forth a comprehensive model based on the combination of existing generative pre-trained models and human-computer interaction design, with the objective of enhancing the efficiency and quality of e-commerce content generation. In particular, the initial stage of the process involves the utilisation of multi-modal data fusion technology, which facilitates the integration of diverse input sources, including images, videos, and textual data. This approach enhances the model's capacity to comprehend the multifaceted nature of products and services, thereby optimising the generation process. Secondly, a cross-domain sentiment analysis and recommendation engine was designed, which combined the self-attention mechanism in deep learning with the objective of analysing the popularity of the generated content according to consumer behaviour in real time and adjusting the generation strategy. Furthermore, the real-time feedback loop mechanism has been innovatively introduced, which enables the dynamic optimisation of content generation in accordance with user evaluation and click-through rate. This facilitates an improvement in the personalisation and accuracy of content. The experimental results demonstrate that the model markedly enhances the efficiency of content generation and user satisfaction on the e-commerce platform.

Keywords: generative pre-trained models; multimodal data fusion; human-computer interaction design; sentiment analysis

1. Introduction

In the context of the accelerated growth of e-commerce, particularly within the context of small and medium-sized businesses (SMBs), the capacity to differentiate oneself in a highly competitive market has emerged as a pivotal challenge for enterprises. The content presented on e-commerce platforms, including product descriptions, marketing copy, and visual materials, plays a pivotal role in attracting consumers, which subsequently impacts the sales performance and brand image of the products in question. Due to their limited financial and human resources, small and micro enterprises frequently lack the capacity to invest the time, cost, and labour required to create content in-house [1]. Consequently, the generation of content in a rapid and high-quality manner has become a central requirement for these enterprises, with the objective of enhancing their competitive position in the market.

Generative AI, a technology that rapidly generates high-quality content, has made significant progress in a number of fields in recent years. In particular, the potential of generative AI has been demonstrated in the generation of text, images, videos and other forms of content. In the context of e-commerce content generation, generative AI has the potential to rapidly generate creative and engaging product descriptions, ad copy, product introduction videos, and more, based on the input product information [2]. Nevertheless, despite the considerable advantages of generative AI in

content generation, it still encounters certain challenges, particularly in the tuning of model parameters and the uncertainty of the generated results. Small and micro enterprises frequently lack the requisite technical support and professionals to handle complex AI models. This may result in generated content that does not align with the desired brand style, user expectations, or market demand, which in turn may affect user experience and business value [3].

Conversely, as human-computer interaction (HCI) technology advances, an increasing number of studies are concentrating on reducing the barrier to entry for sophisticated technologies through the optimisation of user interface and interaction design. Conventional AI models frequently necessitate the input of experts to modify parameters, resulting in outputs that are less amenable to control [4]. The combination of generative AI with optimised human-computer interaction design, particularly the simplification of the complex adjustment process of AI models into a visual operation interface, facilitates the more flexible and expedient utilisation of generative AI by SMEs. This enables the swift modification of the style, emotion, and format of generated content, thereby enabling the generation of high-quality content that aligns with the needs of enterprises. In this way, the technical difficulty of using generative AI is reduced, and enterprises are enabled to exercise more precise control over the generated content, thereby improving the overall effect [5].

Furthermore, a significant challenge in the field of e-commerce content generation is the enhancement of personalisation and accuracy in content. The diversity of consumer needs and preferences is increasing, and businesses must adapt their marketing strategies and content generation in accordance with user interests, behaviours, and feedback. The conventional approach to e-commerce content generation relies on the use of fixed templates and static rules, which often proves inadequate for matching generated content with the immediate needs of users [6]. In recent years, the use of real-time feedback mechanisms based on user behaviour has emerged as a significant method for optimising content generation. By analysing data such as click-through rates, purchase behaviour, reviews, and more, generative AI can adjust generated content in real time to improve its relevance and appeal. The combination of these two mechanisms allows for the enhancement of content personalisation, as well as the continuous optimisation of content strategies by companies operating in highly competitive markets, with the additional benefit of increased user engagement and satisfaction [7].

Nevertheless, for a considerable number of SMEs, while generative AI and data analytics technologies can provide solutions, technical barriers and implementation costs remain significant challenges. While some large enterprises and technology companies are able to invest resources in developing and maintaining complex AI systems, for most small and medium-sized enterprises, the question of how to effectively leverage these advanced technologies has become a significant constraint on their development. The challenge, therefore, is to make these advanced technologies affordable and efficient for more small and medium-sized enterprises by simplifying technology operations and optimising user experience.

2. Related Works

Deng et al. [8] put forth a personalised answer generation method based on multi-perspective preference modelling. The research primarily concentrates on enhancing the quality of answer generation on e-commerce platforms by modelling the heterogeneous preferences of users. By analysing the historical behaviour and preference information of users, a framework for multi-perspective preference modelling is proposed, which effectively improves the quality of the generated answers. The integration of multiple user preferences enables the generated text content to respond with greater accuracy to users' problems and needs, thereby enhancing the user experience.

Additionally, Zhang et al. [9] developed a pre-trained model for the e-commerce sector, with the objective of enhancing the quality of natural language generation (NLG), particularly in product descriptions and advertising copy. The findings indicate that a pre-trained model tailored to the e-commerce domain can markedly enhance the relevance and naturalness of the generated text, thereby

addressing the need for rapid updates to product content on e-commerce platforms. Furthermore, the study investigates the potential for enhancing product copy with more personalised content by optimising generation strategies to more effectively capture consumer attention and drive sales.

Simanjuntak et al. [10] investigated the influence of value co-creation on the marketing performance of small and medium-sized enterprises (SMEs) operating on e-commerce platforms. Although this article is primarily concerned with marketing strategies and consumer engagement, it is noteworthy that a key conclusion regarding content generation is that the provision of high-quality, personalised content is crucial for enhancing the market performance of small and micro businesses.

Roumeliotis et al. [11] conducted a comparative analysis of the performance of GPT and LLaMA, two prominent large language models, in the assessment of e-commerce product reviews. It has been demonstrated that the two models display disparate characteristics when confronted with a substantial corpus of product reviews. This is particularly evident in the domains of sentiment analysis, review accuracy, and user feedback prediction.

3. Methodologies

3.1. Multimodal Data Fusion

In order to enhance the generative model's understanding of diverse products and services in e-commerce, multimodal data fusion is introduced as a preliminary step. Input information, including images, videos, and text, is considered and jointly modelled in order to facilitate comprehensive analysis of content from multiple sources of information when generating content.

Let the image data be represented by the vector I, with dimensions $\mathbb{R}^{H \times W \times C}$, where H and W are the height and width of the image, and C is the number of colour channels. The video data is represented by $V = \{V_1, V_2, \ldots, V_T\}$, where $V_t \in \mathbb{R}^{H' \times W' \times C'}$ denotes the image data of the video frame t. The text data is represented by $T = \{t_1, t_2, \ldots, t_N\}$, where $t_i \in \mathbb{R}^D$ denotes the word embedding vector. The following process, represented by Equation (1), jointly models multimodal data fusion:

$$X_{fused} = f_{fusion}(I, V, T), \tag{1}$$

where f_{fusion} represents a multimodal fusion function based on the attention mechanism. The image, video, and text are processed through a shared encoder, and a system-feature representation is output fused word $X_{fused} \in \mathbb{R}^d$, which is used for subsequent generation.

Multimodal fusion enables the model to consider a comprehensive range of input data types, enhancing its capacity to understand the interrelationships between product images, videos, and text descriptions. Furthermore, this approach improves the diversity and quality of content generation. The aforementioned method has the potential to generate content that is more accurate and aligns with the needs of users in e-commerce scenarios. The integration of sentiment analysis with a recommendation engine enables the model to adapt its generation strategy in a dynamic manner, based on consumer behaviour and feedback. A sentiment analysis module was designed, which classifies the sentiment of multimodal input through the self-attention mechanism and optimises the generated content based on sentiment labels.

The consumer behaviour data may be represented by the set $B = \{b_1, b_2, \ldots, b_K\}$, where $b_k \in \mathbb{R}^D$, and where each element of B represents a specific action performed by the consumer, such as clicking, viewing, or commenting on a given piece of content. The sentiment analysis network employs the self-attention mechanism to process the aforementioned data, subsequently outputting a sentiment score of $S_{sent} \in \mathbb{R}^{K \times 1}$. This is represented by the following Equation (2).

$$S_{sent} = Attention(Q, K, V), \tag{2}$$

where Q, K, and V represent the query, key, and value vectors of consumer behaviour data, respectively. The following Equations (3), (4) and (5) is also relevant:

$$Q = W_q B, \tag{3}$$

$$K = W_k B, \tag{4}$$

$$V = W_v B, \tag{5}$$

The weight matrices W_q, W_k, and W_v were obtained from the training phase. The sentiment score, S_{sent}, is calculated by determining the degree of similarity between the query and the key. This score is then employed to adjust the sentiment propensity and personalisation of the generated content. In accordance with the sentiment score, the recommendation engine selects the optimal content for display by weighing and fusing the generated text, image, and video features, as illustrated in Equation (6).

$$X_{recommend} = \sum_{k=1}^{K} S_{sent}(k) \cdot X_k, \tag{6}$$

where X_k represents the k-kth content feature, while $S_{sent}(k)$ denotes the corresponding sentiment score. The sentiment analysis module employs self-attention mechanisms to discern consumers' emotional responses, subsequently integrating this information into the content generation process. This approach enhances the relevance and popularity of the generated content. The recommendation engine employs a dynamic selection and adjustment process for generated content, utilising sentiment scores to ensure a high degree of alignment with consumer needs.

3.2. User Feedback Optimization Mechanism

In order to enhance the personalisation and accuracy of content, a real-time optimisation mechanism has been introduced, based on user feedback. In particular, reinforcement learning techniques are employed to modify the generation strategy in response to user interaction data, including metrics such as click-through rates and reviews.

Following the generation of each piece of content, the user feedback data set, represented by $F = \{f_1, f_2, \ldots, f_m\}$, is incorporated into the reinforcement learning model. In this context, f_m represents the user's feedback on a specific generated piece of content, and it is a function of $\mathbb{R}^D$. By optimising the objective function, as defined in Equation (7).

$$\mathcal{L}_{RL} = -\sum_{m=1}^{M} log\pi(a_m|s_m)R_m, \tag{7}$$

In this context, $\pi(a_m|s_m)$ represents the policy network, which is based on the current state s_m. R_m denotes the reward signal that is fed back by the user, whereas a_m refers to the action that is taken. Maximising this loss function enables the model to adapt the generation strategy in accordance with user feedback and thereby optimise the degree of personalisation of the content.

The incorporation of user feedback represents a pivotal aspect of the optimisation of generated content. Reinforcement learning, in particular, facilitates the dynamic adjustment of content generation strategies with the objective of maximising user satisfaction. This approach enables the model to achieve continuous improvement in the quality and user satisfaction of its generated content over time. The aforementioned modules can be integrated to create an end-to-end system, as illustrated in Equation (8).

$$X_{final} = f_{opt}\left(f_{fusion}(I, V, T), Attrntion(B), RL(F)\right), \tag{8}$$

The f_{opt} system employs an optimization function that integrates multimodal fusion, sentiment analysis, a recommendation engine, and reinforcement learning feedback. This enables the generation of content X_{final}, which incorporates text, images, and videos and is tailored to the user's specific requirements.

In the contemporary e-commerce milieu, content generation faces numerous challenges. Primarily, the content is frequently impersonalised and lacks customisation, which has a detrimental effect on the user's purchasing decision and, consequently, the website's conversion rate. Secondly,

the presence of repetitive and substandard content can diminish the user experience, potentially reducing the appeal of the website. Algorithmic bias and data processing issues in the content generation process can also lead to inaccurate or irrelevant recommendations, affecting user satisfaction. It is imperative to address these issues to enhance the user experience and boost business value. The implementation of advanced content generation models, which employ sophisticated analysis of user needs to deliver personalised and engaging content, has been demonstrated to enhance user engagement and foster brand loyalty.

4. Experiments

4.1.Experimental Setups

The publicly available product review dataset from Amazon, comprising millions of authentic user reviews, product descriptions and ratings across a multitude of product categories, including electronics, clothing and home goods, was utilised in this study. The model is fine-tuned based on GPT-4, configured with 1.3 billion parameters, and trained by the Adam optimiser, with a learning rate of 5e-5 and a batch size of 32.

4.2.Experimental Analysis

A comparative analysis was conducted with several existing mainstream content generation models, including GPT-3, BERT, and LLaMA models. The comparison was primarily conducted with the objective of evaluating the quality of the generated content, personalisation, and emotional fit. In particular, the accuracy and fluency of the generated text were evaluated using commonly employed text generation evaluation metrics, namely BLEU, ROUGE, and METEOR. Additionally, actual user behaviour data, such as user satisfaction scores and click-through rates, is employed to further evaluate the personalisation and attractiveness of the generated content. Figure 1 illustrates that the BLEU scores of all models demonstrate a gradual increase with the addition of training steps.

However, the proposed Ours model exhibits a notable advantage. The Ours model demonstrated not only high generation quality at the outset but also a progressive increase in BLEU score, ultimately attaining the highest value in the latter stages of training. Its convergence was also rapid. In contrast, GPT-3, BERT, and LLaMA also demonstrated improvement, albeit with relatively stable performance under the same training steps. Figure 1 illustrates that the Ours model exhibits superior efficiency and quality in e-commerce content generation, capable of producing high-quality and personalized text in a more expeditious manner. This makes it particularly well-suited to content optimization in practical applications.
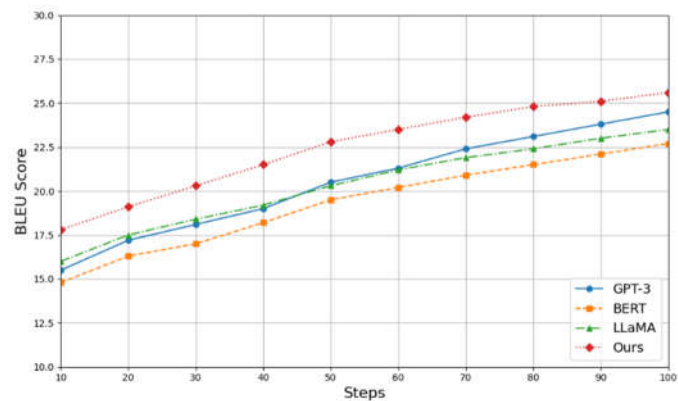

Figure 1. BLEU Score Comparison Across Different Models.

Subsequently, ROUGE and METEOR were employed as the principal evaluation criteria for assessing the performance of diverse generation models in e-commerce content generation tasks. The

Recall-Oriented Understudy for Gisting Evaluation is primarily employed to assess the extent of overlap between generated content and reference content. It is frequently utilized for automatic summarization and quality assessment of text generation. The Metric for Evaluation of Translation with Explicit Ordering is another indicator for evaluating the similarity between generated text and reference text. Particular emphasis is placed on word order and semantic matching.

As you can see in Figure 2, the ROUGE and METEOR scores of all models show a gradual increase as the training steps increase. Nevertheless, the Ours method demonstrated substantial advantages in both metrics, particularly in regard to the ROUGE score, where its growth rate and ultimate level surpassed those of other models. This indicates that our method is more pronounced in terms of the quality and relevance of the content generated. Furthermore, although BERT and LLaMA also demonstrated a relatively consistent improvement, their score increases were less pronounced than those of the Ours method, indicating that they may have inherent limitations in terms of the accuracy and diversity of the content they generate.

Despite the model proposed in this study addressing several issues in e-commerce content generation to a certain extent, there are still some limitations. Primarily, the model's current scope of application is confined to specific product categories and market environments, a limitation that may constrain its generalisability. Secondly, the training of the model is dependent on substantial quantities of high-calibre data; in certain instances, the quality and diversity of data can influence the performance of the model. In light of these limitations, the proposed model enhances content generation efficiency by integrating deep learning and natural language processing technologies.
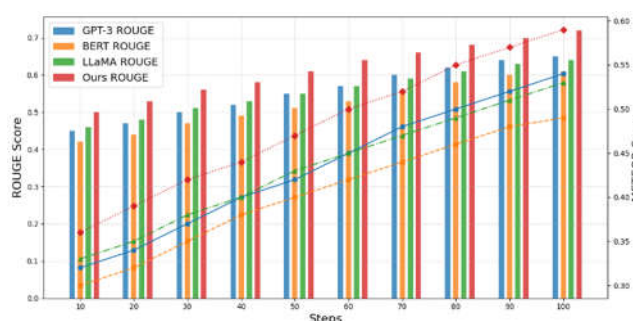

Figure 2. Model Performance Comparison: ROUGE and METEOR Scores.

5. Conclusions

In conclusion, the proposed method significantly outperforms existing models in e-commerce content generation, as evidenced by higher ROUGE and METEOR scores. Ours consistently shows better alignment with reference content, particularly in ROUGE, highlighting its superior performance in both relevance and semantic accuracy. While other models exhibit steady improvement, Ours achieves more substantial gains, demonstrating its potential to generate high-quality, personalized content efficiently for e-commerce applications.

References

1.   Lin, Xiaolin, and Xuequn Wang. "Towards a model of social commerce: improving the effectiveness of e-commerce through leveraging social media tools based on consumers' dual roles." European Journal of Information Systems 32.5 (2023): 782-799.

2.   Farhan, Gusti Muhammad, and Endy Gunanto Marsasi. "The Influence of Information Quality and Perceived Value on Purchase Intention of Game shop E-commerce in Generation Z Based on Framing Theory." Jurnal Pamator: Jurnal Ilmiah Universitas Trunojoyo 16.3 (2023): 620-631.

3.   Ballerini, Jacopo, Dennis Herhausen, and Alberto Ferraris. "How commitment and platform adoption drive the e-commerce performance of SMEs: A mixed-method inquiry into e-commerce affordances." International Journal of Information Management 72 (2023): 102649.

4.   Guo, Xiaojie, et al. "Intelligent online selling point extraction for e-commerce recommendation." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 36. No. 11. 2022.

5.   Seyyedamiri, Nader, and Ladan Tajrobehkar. "Social content marketing, social media and product development process effectiveness in high-tech companies." International Journal of Emerging Markets 16.1 (2021): 75-91.

6.   Sudirjo, Frans, et al. "The Influence of Generation Z Consumer Behavior on Purchase Motivation in E-Commerce Shoppe." Profit: Jurnal Manajemen, Bisnis dan Akuntansi 2.2 (2023): 110-126.

7.   Sudirjo, Frans, et al. "Digital Marketing and Sales Support For Hydroponic MSME Growth Through Mobile Based E-Commerce Design." Jurnal Ekonomi 12.3 (2023): 1750-1756.

8.   Deng, Yang, et al. "Toward personalized answer generation in e-commerce via multi-perspective preference modeling." ACM Transactions on Information Systems (TOIS) 40.4 (2022): 1-28.

9.   Zhang, Xueying, et al. "Automatic product copywriting for e-commerce." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 36. No. 11. 2022.

10.   Simanjuntak, Mariana, Arnaldo M. Sinaga, and Humasak AT Simanjuntak. "The Role of Value Co-Creation in E-Commerce to Improve MSME Marketing Performance." Proceeding International Conference on Information Technology, Multimedia, Architecture, Design, and E-Business. Vol. 2. 2022.

11.   Roumeliotis, Konstantinos I., Nikolaos D. Tselikas, and Dimitrios K. Nasiopoulos. "LLMs in e-commerce: a comparative analysis of GPT and LLaMA models in product review evaluation." Natural Language Processing Journal 6 (2024): 100056.