Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 13, No. 2, 2025

Exploring the Depths of User Interest: How LLMs are Revolutionizing Personalized Shopping

Xin Guan *

International School, Beijing University of Posts and Telecommunications, Beijing, 100876, China

*** Corresponding author Email:** guanxin2021@bupt.edu.cn

Abstract: This paper explores the transformative impact of Large Language Models (LLMs) on personalized shopping, highlighting their revolutionary capabilities in natural language understanding, dynamic interest modeling, and multimodal interaction. It discusses how LLMs address the limitations of traditional recommendation systems by enabling semantic deconstruction of user queries, real-time strategy updates, and scenario-based solutions. The study reviews the evolution of LLMs and personalized recommendation systems, emphasizing breakthroughs in contextual reasoning and cross-modal feature fusion. Practical implementation strategies are proposed, including lightweight technology deployment, cross-departmental collaboration, and user-centric experience design. The paper concludes by underscoring the dual technical and commercial value of LLMs, advocating for future research in causal reasoning and multimodal evaluation to enhance algorithmic fairness and sustainability.

Keywords: Large Language Models (LLMs); Personalized Recommendation; Intelligent E-Commerce; Natural Language Understanding; Multimodal Interaction; Dynamic Interest Modeling.

1. Introduction

The emergence of Large Language Models (LLMs) marks a revolution in the field of artificial intelligence. These models have achieved astonishing performance, especially in aspects such as understanding, language synthesis, and common - sense reasoning, almost comparable to human capabilities. This leap in general AI capabilities will profoundly transform the implementation of personalized services.

In the era of the rapid development of large language models, the ideas of personalized marketing are undergoing a profound transformation from "weak intelligence" to "strong intelligence". The driving force behind this transformation stems from the challenges to traditional personalized marketing methods, and the rise of large language models injects a more intelligent and flexible impetus into personalized marketing. It will greatly expand the scope of personalized services, evolving them into services that provide more complex functions. Traditional personalized marketing methods often rely on pre - defined rules and algorithms, with relatively fixed decision - making logic, making it difficult to flexibly meet the diverse needs of users. The emergence of large language models has brought about a disruptive change to personalized marketing. By using large language models as a general interface, personalized systems can compile user requests into plans, call external tools to execute these plans, and integrate the outputs, achieving end - to - end personalized tasks. This makes personalized promotion more flexible, enabling a better understanding and interpretation of users' true needs, marking a significant step forward in the field of personalized marketing. Against this backdrop, this study aims to deeply explore the underlying logic of personalized promotion, systematically sort out the development context of changes in the underlying logic of personalized promotion, examine the challenges faced by personalized services, and use large language models to address them. It also elaborates on specific application scenarios in detail and proposes ideas for personalized marketing in the era of large language models.

2. Review of Domestic and Foreign Research Status

2.1. Large Language Models (LLMs)

In recent years, researchers have found that by increasing the number of parameters and the amount of data in pre - trained language models, large language models can significantly improve their performance and demonstrate many special capabilities that small models do not possess (such as context learning ability, step - by - step reasoning ability, etc.) [1]. In 2018, Google proposed the BERT model, which achieved astonishing performance on a large - scale corpus through pre - training [2]. The introduction of the Transformer model makes the model more effective in handling long - distance dependencies. Recently, OpenAI's GPT series (Generative Pre - trained Transformer) has rapidly promoted the development of large language models. GPT - 3, with 175 billion parameters, set a new record in model scale, achieving great success in understanding and generating natural language. In 2024, GPT - 4o released by OpenAI showed outstanding performance in multimodal interaction, greatly enhancing the ability to understand and generate complex long texts, and further improving controllability, attracting widespread attention from the global technology community [3].

With the introduction of the GPT series, Chinese researchers have actively participated in the development of large language models. Some research institutions have made localized improvements based on GPT, making these models more suitable for the Chinese language environment and have actively launched their own large language models. For example, Baidu's "Wenxin Yiyan", 360's large language model, Alibaba's "Tongyi Qianwen", and SenseTime's "Shangliang" were all launched. They are committed to applying large language models to practical scenarios such as search engines and speech recognition and have achieved remarkable results. At the same time, the DeepSeek model has emerged. Based on the Transformer architecture, it independently develops deep neural networks and has outstanding capabilities in many aspects such as intelligent dialogue, text generation, and semantic understanding. DeepSeek - R1 contains 671 billion parameters and is trained on a dataset of 2 trillion tokens, covering Chinese and English. It performs excellently in reasoning, coding, mathematics, and Chinese understanding, surpassing some similar models.

In general, the emergence of large language models represents a major breakthrough in the field of artificial intelligence. Their enhanced capabilities in understanding, language analysis, and common - sense reasoning open up new possibilities for personalization.

2.2. Personalized Recommendation

Personalized recommendation systems are the product of the development of the Internet and e - commerce. They were proposed to solve the problem of Information Overload through information retrieval and filtering. The first collaborative filtering system was published by scholars such as David Goldberg and David Nichols (1992), aiming to solve the problem of information overload in users' emails [4]. Later, Resnick and Varian (1997) gave a clear definition of recommendation systems: using e - commerce websites to provide customers with product information and suggestions, helping users decide what products to buy, simulating the process of salespeople helping customers complete purchases [5].

With the rise of deep learning technology, foreign researchers began to introduce deep learning into the field of personalized recommendation. They proposed a series of recommendation algorithms based on neural networks, such as models based on RNN and CNN, achieving significant performance improvements. When big data technology became more mature, foreign researchers proposed more accurate recommendation algorithms through the analysis of large - scale user behavior data. These algorithms can better understand users' interests and behavior patterns, significantly improving the personalization level of recommendation systems.

China entered the initial stage of research on personalized recommendation systems only after the 21st century. Around 2010, domestic researchers mainly focused on traditional collaborative filtering methods and gradually introduced machine - learning - based recommendation technologies, striving to solve the problems of user cold start and data sparsity to improve the accuracy and coverage of recommendation systems [6]. With the increasing emphasis on deep learning technology in China, research on personalized recommendation has gradually incorporated deep learning. Some domestic research institutions and enterprises began to propose recommendation algorithms based on neural networks and initially achieved some results. With the advent of the big data era, domestic research on personalized recommendation has been able to conduct more in - depth research in user behavior mining and pattern recognition with the help of big data analysis technology, to improve the personalization level of recommendation systems [7].

3. Overall Construction Ideas of Intelligent E - commerce Empowered by Large Language Models

Currently, research on "personalized recommendation" mainly focuses on technical principles, platform development and applications, and communication effects. However, regarding the ultimate service object of personalized recommendation - the users themselves, as users are key participants in the information demand market and personalized recommendation exists to better meet users' needs, this paper argues that the research direction of "personalized recommendation" should mainly focus on the following aspects:

3.1. Deep Breakthrough in Natural Language Understanding

Large language models are different from traditional rule - based or simple statistical - based recommendation systems. They can not only analyze users' historical behaviors but also provide more accurate personalized recommendations by learning users' natural language inputs in real - time.

In the aspect of natural language understanding, LLMs achieve high - order semantic parsing through the self - attention mechanism of the Transformer architecture and the pre - training - fine - tuning paradigm. For example, users' vague requirements (such as "lightweight backpacks suitable for summer commuting") can be decomposed into multi - dimensional features such as "seasonal attributes", "usage scenarios", and "functional requirements", and combined with product descriptions to generate accurate recommendations. Traditional rule - based systems based on keyword matching have difficulty capturing such complex semantic relationships. In 2024, Walmart launched a generative AI search function. By integrating proprietary data with Azure OpenAI models, it achieved in-depth parsing of user intentions, transforming search into an end-to-end match between 'intentions' and 'products', significantly enhancing recommendation relevance [8]. In addition, the zero - shot classification ability of LLMs (such as inferring new product labels) and sentiment analysis (such as identifying the polarity of reviews) further optimize the granularity of user profiles, solving the cold start problem of traditional rule - based systems.

In terms of product recommendation, large language models can not only understand users' explicit needs (such as when users clearly search for a certain product) but also mine users' implicit needs (such as inferring their potential interests by analyzing users' browsing and purchase history). The application of this technology makes product recommendation no longer limited to simple rule matching but can dynamically generate personalized recommendation lists according to users' personalized preferences, thus significantly improving the relevance and click - through rate of recommendations.

In content generation, the introduction of large language models solves the content production bottleneck faced by traditional e - commerce platforms. Through large language models, platforms can automatically generate high - quality product copy, advertising creativity, and summaries of user reviews, greatly improving the production efficiency of marketing content. This automated content generation ability not only reduces the operating costs of enterprises but also enhances users' trust and satisfaction with the platform's content.

3.2. Returning to the Users Themselves

Currently, research on large models in the e - commerce field has made important progress at the technical level, but there is relatively insufficient attention to the ultimate service object of personalized recommendation - the users themselves. Consumers still need frequent manual intervention to solve their needs.

Data from iiMedia Research shows that currently, consumers have a relatively high demand for manual intervention functions. 45.5% of consumers use it occasionally, and 33.7% of consumers use it when needed, indicating that users have certain expectations for AI e - commerce. In addition, 33.3% of consumers believe that the manual intervention function has improved the shopping experience of AI e - commerce, indicating that AI e - commerce still needs to improve its problem - solving ability in complex scenarios, and its development still requires human assistance.

From the data, we can see users' positive feedback on manual intervention, which provides an optimization direction for personalized recommendation systems. When designing recommendation systems, e - commerce enterprises should introduce a user feedback mechanism, allowing users to express their needs and opinions during the recommendation process to promote a better personalized experience.

Users' demand for and feedback on manual intervention indicate that personalized recommendation does not only rely on automated AI shopping guides or customer service but needs to truly understand users' shopping behaviors and psychology. The application of large models in the field of dynamic interest capture can help analyze users' historical data, behavior patterns, and emotional responses, thus formulating recommendation strategies that better meet users' needs.

The innovation of LLMs in the field of dynamic interest capture is reflected in their co - evolution with behavior sequence modeling. The combination of user behavior sequence analysis based on LSTM/GRU and the active prompting technology of LLMs can perceive interest drift in real - time. For example, when a user shifts from "browsing sports shoes" to "searching for hiking gear", the system dynamically adjusts the recommendation strategy from "exploring style diversity" to "focusing on in - depth recommendations of a certain style" through the time - series analysis of historical behaviors (such as the LDA topic model) and real - time dialogue interaction [9].

3.3. Multimodal Interaction Experience

Modal interaction reconstruction promotes scenario - based implementation through cross - modal feature fusion. For example, in the practical application of intelligent fitting mirrors, the system uses image recognition technology to capture users' body shape data (such as height, shoulder width, waist circumference, etc.) in real - time, and combines voice dialogue interaction to analyze users' style preferences (such as "wanting a casual style" or "needing outfits for formal occasions"). Then, it inputs the visual and language modal data into a multimodal large language model. The model integrates users' body shape features and semantic requirements through cross - modal alignment algorithms, dynamically generating high - precision 3D virtual fitting effects, and recommending suitable matching schemes based on users' historical behaviors and the product attribute database (such as recommending styles according to body shapes and matching colors according to preferences). Taobao's AI shopping assistant integrates visual search (such as extracting product image features) and natural language processing, encoding text and image information into a unified vector space, solving the problem of information sparsity in traditional single - modal recommendations.

Such technologies not only solve the problems of lagging user feedback and single - dimensional matching suggestions in traditional fitting scenarios but also enhance user engagement through real - time interaction, significantly improving the efficiency of purchase decision - making and customer satisfaction. At the same time, they provide merchants with insights into user preferences to optimize product selection strategies.

4. Suggestions on the Implementation Strategies for the Construction of Intelligent E - commerce Platforms

4.1. Exploring the Path of Lightweight Technology Implementation

During the process of large language models empowering e - commerce platforms, due to the high computing resources, professional technologies, and development time required for model training and deployment, enterprises often face huge technical barriers and cost pressures during the introduction and implementation of models. Therefore, in order to complete the deployment and application of large language models in e - commerce platforms in a short time, small and medium - sized enterprises can adopt the "API + low - code" mode to quickly deploy the capabilities of LLMs. For example, by calling API services or deploying lightweight models, they can embed conversational shopping functions in existing APPs, avoiding the high - cost investment of building their own models to achieve basic functions such as intelligent search and conversational recommendation.

On this basis, the introduction of lightweight model optimization technologies (such as model quantization, pruning, and knowledge distillation) can further reduce the model size and computational requirements. For example, 8 - bit quantization can reduce the number of model parameters to 1/4 of the original scale while maintaining more than 90% of the performance, making local deployment possible in environments with limited computing power. In addition, the maturity of open - source lightweight frameworks (such as Hugging Face's Transformer.js and TensorFlow Lite) provides enterprises with flexible toolchain support, enabling them to dynamically adjust the model architecture according to business scenarios. For example, embedding a lightweight dialogue engine in mobile applications to achieve real - time interaction and low - latency response. The core advantage of this technical path lies in balancing costs and performance, allowing small and medium - sized enterprises to achieve intelligent upgrades through modular and scalable technical solutions without building large - scale computing infrastructure.

4.2. Promoting Cross - Departmental Collaboration Mechanisms

The advantages of large language models in multimodal data analysis and cross - domain knowledge fusion provide a technical basis for breaking down data silos within enterprises and building collaborative business chains. Enterprises need to make full use of the multimodal data analysis capabilities of LLMs to break down data silos and establish collaborative links among product, technology, and operation teams.

To achieve this goal, enterprises need to establish cross - departmental data sharing and collaboration mechanisms. First, by building a unified data center, standardize and integrate the supply chain data of the product department (such as inventory dynamics, supplier information), the user behavior logs of the operation department (such as click - through rates, conversion paths), and the external environment data of the marketing department (such as social media trends, competitor dynamics), forming a global data asset pool. On this basis, use the semantic understanding and correlation analysis capabilities of large language models to extract the potential features of cross - domain data. For example, combine designers' fashion forecasts with users' historical purchase records to generate seasonal product recommendation strategies.

To achieve efficient collaboration, enterprises need to optimize internal processes simultaneously. For example, use automated workflow tools to trigger data updates and model iterations in real - time, and clarify the roles and responsibilities of each department in data annotation, model feedback, and strategy adjustment through a collaborative platform. In addition, the integration of privacy protection technologies can ensure the security of cross - departmental data flow and avoid the leakage of sensitive information. The core of this mechanism is to transform scattered business units into a collaborative network centered around data and user needs through technology - driven and process reengineering, thereby maximizing the value of large language models in business decision - making.

4.3. Building Scenario - Based User Experience Maps

Building a multimodal and highly responsive scenario - based interaction system around the entire user lifecycle journey (search, decision - making, after - sales) is the core direction for large language models to enhance the user experience of e - commerce platforms. In the search stage, it is necessary to deeply integrate natural language processing and computer vision technologies. For example, support dialect and fuzzy semantic input through a speech recognition engine, and at the same time, combine a visual search model to achieve the "search - by - image" function, enabling users to trigger accurate product matching through diverse input methods.

The key in the decision - making stage lies in dynamic interest modeling and personalized content generation. For example, use the context reasoning ability of large language models to analyze users' browsing trajectories and real - time feedback (such as page dwell time, price - comparison behavior) in real - time, generate comparative analysis reports (such as product parameter comparisons, user review summaries) to assist in decision - making, and automatically create customized marketing content (such as limited - time discount reminders, matching suggestions) through generative technology.

In the customer service stage, the boundaries of services can be extended through augmented reality (AR) and virtual reality (VR) technologies. For example, in the after - sales scenario of home appliances, the real - time guidance function based on AR glasses can help users complete equipment installation or troubleshooting, and at the same time, recommend related accessories or extended warranty services to increase the repurchase rate. The construction of user experience maps needs to focus on the data loop, that is, collect users' behavior data at each link through data tracking, and continuously iterate the model strategy to ensure that the intelligence and humanization of interactive scenarios always evolve in line with users' needs.

5. Conclusion

This study systematically demonstrates the reconstruction logic and practical paths of large language models (LLMs) for the personalized shopping ecosystem, revealing their dual subversiveness in technical breakthroughs and commercial value. The core conclusions are as follows:

First, LLMs solve the fundamental limitations of traditional recommendation systems through three technical leaps: semantic deconstruction, dynamic response, and scenario integration. The natural language understanding ability transforms vague requirements into structured features, dynamic interest modeling enables minute - level strategy updates, and multimodal interaction promotes the upgrade of shopping behavior from "single - product transactions" to "scenario - based solutions". These technical breakthroughs verify the irreplaceability of LLMs in terms of the depth of demand insight and response speed.

Second, the implementation of LLMs needs to replace pure technical thinking with business value orientation. Successful cases show that enterprises should focus on three dimensions: first, building cross - departmental data collaboration links; second, designing end - to - end scenario - based experiences; third, establishing a "lightweight + compliant" technical path. Technical investment must be directly linked to quantifiable business indicators to avoid falling into the trap of "AI for the sake of AI".

Future research can further explore two directions: on the one hand, use causal reasoning models to identify the true associations between recommendation strategies and user behaviors (such as excluding the confounding factors of price promotions); on the other hand, construct a multimodal evaluation system, incorporating indicators such as algorithmic fairness and carbon footprint tracking into the decision - making framework for technical iterations. It is worth emphasizing that the value of LLMs does not lie in replacing humans, but in enhancing merchants' "cognitive bandwidth" and users' "decision - making freedom", reshaping the symbiotic relationship of "people - goods - environment". When technology becomes invisible and needs are met instantly, personalized shopping can truly achieve the ideal state of "both precise and unobtrusive".

References

[1] Chen, J., Liu, Z., Huang, X., et al. (2023). When large language models meet personalization: Perspectives of challenges and opportunities. World Wide Web, 27: 42.

[2] Editorial Office. (2018). Google open-sources the source code of the BERT model. Data Analysis and Knowledge Discovery, 2(11): 18.

[3] Bubeck, S., Chandrasekaran, V., Eldan, R., et al (2023). Sparks of artificial general intelligence: Early experiments with GPT-4. ArXiv, abs/2303.12712.

[4] Goldberg, D., Nichols, D., Oki, B.M., & Terry, D. (1992). Using collaborative filtering to weave an information tapestry. Communications of the ACM, 35(12): 61-70.

[5] Resnick, P., & Varian, H.R. (1997). Recommender systems. Communications of the ACM, 40(3): 56–58.

[6] Lu, W. (2007). Collaborative filtering algorithm and its application in personalized recommendation system. Beijing University of Posts and Telecommunications, 5.

[7] Zhou, F.Y., Jin, L.P., & Dong, J. (2017). A review of convolutional neural networks. Chinese Journal of Computers, 40(6): 1229–1251.

[8] Zhang, Y., Wu, J., Liu, J., et al. (2024) Towards efficient and scalable sharpness-aware minimization. arXiv. https://arxiv. org/ abs/2505.07105.

[9] Harte, J., et al. (2023). Leveraging large language models for sequential recommendation. RecSys '23: Proceedings of the 17th ACM Conference on Recommender Systems, 1096-1102.