

LoRA 를 이용한 4 비트 양자화 LLM 미세 조정을 통한 효율적인 패션 추천

테레 로시다트 올루와부콜라¹, 김정백¹

¹ 전남대학교 AI 융합학과

roshidatdere93@jnu.ac.kr¹, Kyungbaekkim@jnu.ac.kr¹

Efficient Fashion Complementary Prediction via 4-bit Quantized LLM Fine-Tuning with LoRA

Dere Roshidat Oluwabukola¹, Kyungbaek Kim²

¹Dept. of AI Convergence, Chonnam National University

²Dept. of AI Convergence, Chonnam National University

Abstract

Fashion is an integral aspect of daily life that significantly influences both the health and social lifestyle of individuals. While traditional fashion industries have long relied on costly designs to fulfill tasks such as fashion recommendations, artificial intelligence has emerged as a promising tool in this domain. Although current large language models (LLMs) offer potential for enhancing fashion recommendation systems, they often compromise user privacy through data exploitation. To address these challenges, we propose an innovative framework that utilizes pre-trained fashion contrastive language-image models for zero-shot recommendations based on input images. These recommendations are then employed to fine-tune a Llama-8B model with 4-bit quantization specifically for complementary prediction tasks. Our experimental results demonstrate significant improvements in key performance metrics: accuracy (0.85), F1 score (0.86), and AUC (0.94). These findings suggest that the proposed approach could pave the way for leveraging low-bit LLMs in fashion-related tasks while effectively protecting user privacy. This study underscores the potential of integrating advanced AI techniques with privacy-preserving mechanisms to enhance the sustainability and ethical practices of the fashion industry.

1. Introduction

Fashion serves as a multifaceted medium through which individuals express themselves and reveal aspects of their personality. It functions not merely as attire worn around the body but as a means of personal expression, cultural identity, and communication with others. The fashion industry has been shown to influence an individual's self-image and can have a significant impact on various psychological aspects, including emotional well-being, physical health, and phonetics [1]. Research indicates that a poor fashion sense can negatively affect these areas, thereby causing

considerable investment in the field [2]. This emphasis on style is particularly evident among celebrities, who frequently grace runways and red carpets at major events. Fashion designers traditionally play a crucial role by creating garments that complement an individual's aesthetic preferences while also aligning with broader cultural trends. However, the cost of traditional fashion complementary can be astronomical for individuals without access to financial resources.

Artificial Intelligence (AI) has catalyzed transformative changes across traditional industries, with the fashion sector

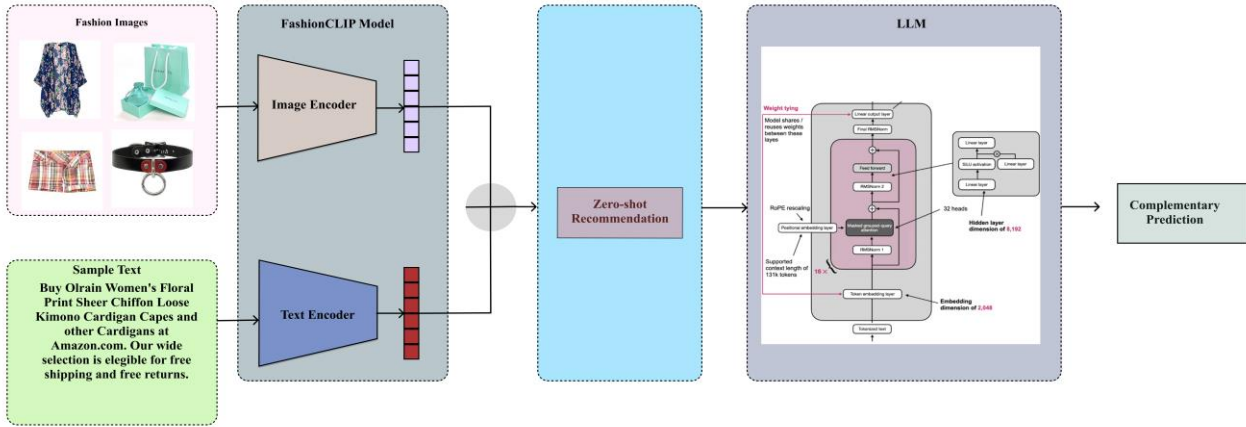


Figure 1 Our Proposed Fashion Complementary Framework

being no exception [3]. Large language models (LLMs), a subset of AI technology, have emerged as invaluable resources across various sectors, including fashion, healthcare, education, and beyond. In the fashion industry, generative AI tools are revolutionizing creative processes by enhancing design exploration, enabling virtual try-ons for immersive styling experiences, identifying outfit compatibilities, and providing personalized recommendations. However, despite these advancements, significant challenges remain that require critical attention and further research to address fully.

Although adapting commercial LLMs for fashion complementary prediction (CP) presents potential benefits, there are significant challenges. Privacy concerns personal data exposure, risk misuse and ethical issues. Additionally, current LLMs are computationally intensive hindering deployment at the edge, compounded by the fact that many existing models aren't specifically designed for fashion tasks, leading to inaccuracies or hallucinations in such contexts.

In this study, we propose an efficient lightweight model for fashion CP that employs a contrastive approach to learn fashion representations alongside the application of LLMs in CP (Figure 1). To the best of our knowledge this is the first time this has been proposed for fashion CP.

2. Methodology

2.1 Dataset

In this study, we employed the open-access Polyvore dataset for evaluation [4]. The dataset includes basic descriptions of fashion images retrieved from the internet

along with their corresponding categories. This dataset comprises over 50,000 outfits, which are categorized into disjoint and non-disjoint versions.

2.2 Preprocessing

Given that the Polyvore dataset crawled from the internet, some basic preprocessing steps were necessary. We filtered the dataset by retaining only those images that had corresponding descriptions. After preprocessing, the train, validation, and test datasets were: 2995 (training), 1000 (validation), and 1145 (test). Preprocessing steps were essential to ensure the integrity of our experimental setup.

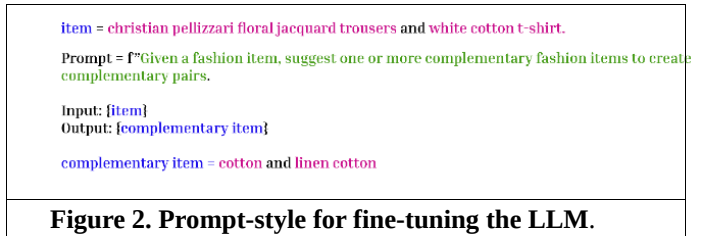


Figure 2. Prompt-style for fine-tuning the LLM.

2.3 Method

The images from Polyvore were used as input into an instruction-prompt fine-tuning process of a pre-trained fashion contrastive language image model (FashionCLIP)[5]. This approach leverages FashionCLIP's robust pre-trained fashion representation to perform zero-shot recommendation, meaning it can generate recommendations without specific training data for those categories. The output from this zero-shot recommendation was subsequently used as input for the fine-tuning process of a LLM to enhance its prompting capabilities in the context of fashion.

We adopted a Llama-8B model LLM [6], quantizing to 4-bit. Utilizing the Low-Ranked Adaptation (LoRA) and Unsloth, we fine-tuned the LLM using the instruction prompt style. A sample of the prompt style is shown in Figure 2. Our experimental setup and hyperparameter used to fine-

tuned the LLM are reported in Table 1.

Hyperparameter	Value
Batch size	32
Dropout	0.2
Optimizer	Adam_8 bit
Learning rate	0.0002
Weight decay	0.01
Max length	2048

validation was set to monitor performance, followed by inference using FastLanguageModel for generating complementary item recommendation

3. Results and Conclusion

3.1 Results

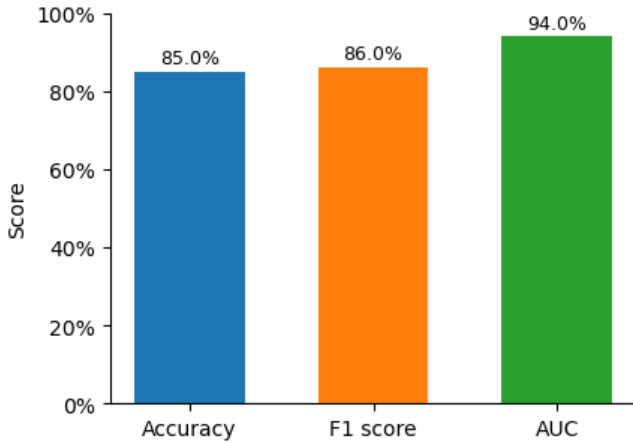


Figure 3 Performance of Our proposed framework

Figure 3. indicate that our proposed leveraging 4-bit quantization and optimized inference techniques yielded successful complementary item recommendations achieving and accuracy of 0.85%, F1 score of 0.86%, and AUC of 0.94%, demonstrating significant advancement beyond image-based systems and offering a viable pathway for smaller businesses to implement AI-driven prediction systems.

3.2 Conclusion

This research demonstrates the potential of LLM-based models for efficient and effective fashion complementary prediction, paving the way for broader accessibility within the retail industry. We plan to finetune more models in future for generalization across different

cultural outfits.

Acknowledgements

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation(IITP)-Innovative Human Resource Development for Local Intellectualization program grant funded by the Korea government(MSIT)(IITP-2025-RS-2022-00156287, 50%). This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) under the Artificial Intelligence Convergence Innovation Human Resources Development (IITP-2025-RS-2023-00256629, 50%) grant funded by the Korea government(MSIT)

References

- [1] Leung, A. C., Yee, R. W., & Lo, E. S. (2015). Psychological and social factors of fashion consciousness: An empirical study in the luxury fashion market. *Research Journal of Textile and Apparel*, 19(3), 58-69.
- [2] Hossain, M. J., Chang, H. J. J., & Jones, R. P. (2025). I bought it and I feel good! An examination of fit factors and self-evaluation related to confident clothing decisions and psychological well-being. *Journal of Retailing and Consumer Services*, 84, 104167.
- [3] Shi, W., Wong, W., & Zou, X. (2025). Generative AI in Fashion: Overview. *ACM Transactions on Intelligent Systems and Technology*, 16(4), 1-73.
- [4] Vasileva, M. I., Plummer, B. A., Dusad, K., Rajpal, S., Kumar, R., & Forsyth, D. (2018). Learning type-aware embeddings for fashion compatibility. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 390-405).
- [5] Chia, P. J., Attanasio, G., Bianchi, F., Terragni, S., Magalhaes, A. R., Goncalves, D., ... & Tagliabue, J. (2022). Contrastive language and vision learning of general fashion concepts. *Scientific Reports*, 12(1), 18958.
- [6] Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., Letman, A., ... & Ganapathy, R. (2024). The llama 3 herd of models. *arXiv e-prints*, arXiv-2407.