Personalized AI Shopping Copilot: A Conversational Multimodal Assistant for Enhanced E-Commerce Recommendations

Abstract

Current e-commerce recommendation engines such as those in Amazon and Flipkart primarily rely on collaborative filtering, content-based filtering, and keyword-driven search queries. While efficient, these systems often fail to deliver dynamic and conversational shopping experiences, leading to limited personalization and lack of transparency in recommendations. This paper proposes a Personalized AI Shopping Copilot that integrates multimodal inputs (text + image) with a conversational chatbot interface. By leveraging advanced models such as CLIP for image-text alignment, Sentence-BERT for semantic similarity, and LLaMA/DialoGPT for dialogue, the system aims to provide explainable recommendations, product comparisons, and bundle suggestions. The copilot enhances user trust, improves transparency, and shifts the user experience from a static "search-and-filter" to an engaging "conversational discovery" journey.

I. Introduction

E-commerce platforms have become the backbone of modern retail. Giants like Amazon and Flipkart use large-scale recommendation systems powered by big data pipelines (Kafka, Hadoop, Spark) and machine learning models. Their engines analyze user clicks, purchases, and browsing history to generate recommendations. While effective in suggesting products, these models are mostly one-way systems: users enter a keyword, the system ranks items, and results are displayed as static product lists. Few systems allow for natural, human-like conversation, or the integration of multimodal queries such as uploading an image of an item along with a price constraint. Studies show that trust, empathy, and transparency in AI chatbots greatly influence customer satisfaction. Thus, there is a strong motivation to build a Personalized AI Shopping Copilot capable of interactive, explainable, and multimodal recommendations.

II. Problem Statement

Despite advancements in recommender systems, the following gaps exist in current systems (Amazon, Flipkart):

- Keyword dependency: Users must rely on exact keywords for product searches.
- Static product lists: No conversational refinement; users must re-search.
- Limited personalization: Based only on past clicks and purchases.
- Opaque recommendations: No explanation for "why" a product was shown.
- Lack of multimodal support: Cannot combine image + text queries. User Experience Issue: Current apps focus on "find and purchase" rather than a "guided shopping journey."

III. Objectives

The goals of the project include:

1. Build a conversational AI assistant capable of understanding natural queries.
2. Integrate multimodal search (text + image) for richer personalization.
3. Design a recommendation engine supporting product bundles and comparisons.
4. Ensure explainability and transparency in recommendations.
5. Provide real-time personalization adapting to user preferences in conversation.

I. Existing Work and Drawbacks

- **Amazon & Flipkart:** Rely on collaborative filtering, content-based filtering, and deep learning (VisNet for images). Their system uses a data pipeline (Kafka, Spark, Hadoop) and learning-to-rank algorithms for large-scale ranking.

- **Drawbacks:** No natural conversation, no multimodal queries, and recommendations often lack explainability.

- **Prior Research Papers:**

Many focus on chatbots with fixed menus (not natural language). Reverse image search exists but is not integrated with dialogue systems. Anthropomorphic chatbots increase trust, but excessive human-likeness may reduce credibility.

Comparison Between Existing Recommendation Models and Proposed AI Shopping Copilot:

Feature	Existing Models	Proposed Copilot
Input	Text-based search only	Multimodal (Text + Image)
Interaction	One-way listing	Conversational chatbot
Personalization	History-based filtering	Real-time semantic understanding
Transparency	Opaque (“black box”)	Explainable recommendations
Recommendations	Single product	Bundles + Comparisons
Adaptability	Static filters	Context-aware & dynamic
Tech Foundation	Basic filtering/NLP	CLIP + Sentence-BERT + LLMs
User Experience	Functional purchase	Engaging, human-like shopping

II. Proposed Solution (Our Approach)

• **Conversational AI Layer:** The system integrates advanced dialogue models like LLaMA or DialoGPT to power natural, human-like conversations. This layer is responsible for understanding user queries, detecting intent behind the requests, and refining queries through

back-and-forth interaction, ensuring that the shopping assistant provides accurate and context-aware responses.

• **Multimodal Understanding:** To bridge the gap between visual and textual information, the copilot leverages **CLIP** for creating joint embeddings of product images and textual descriptions, while **Sentence-BERT** handles semantic embeddings for natural language queries. This combination allows the assistant to understand both "show me something like this" (image-based) and "I want a casual blue shirt" (text-based) queries seamlessly.

• **Recommendation Logic:** The recommendation engine is designed not only to suggest relevant products but also to add intelligence to the process. It can generate product comparisons (e.g., comparing two shirts on price and features), create bundled suggestions (such as pairing a shirt with a matching watch), and provide transparent explain-why reasoning to build user trust in its suggestions.

• **Database & Storage:** All product-related information, including metadata, images, and embeddings, is stored securely in **Firebase Firestore**. This ensures fast retrieval and scalability, allowing the assistant to handle large inventories and provide real-time responses to user queries without performance issues.

• **Deployment & Infrastructure:** The system is deployed with a modern, scalable stack. The **frontend** is built using React and Tailwind for a smooth and responsive shopping experience. The **backend** is powered by FastAPI or Firebase Functions to handle AI model requests and logic processing efficiently. For hosting and scalability, **Vercel** is used for frontend deployment, while **Google Cloud** manages backend services and data infrastructure.

Figure 1: Current Apps vs Proposed Copilot (Features)

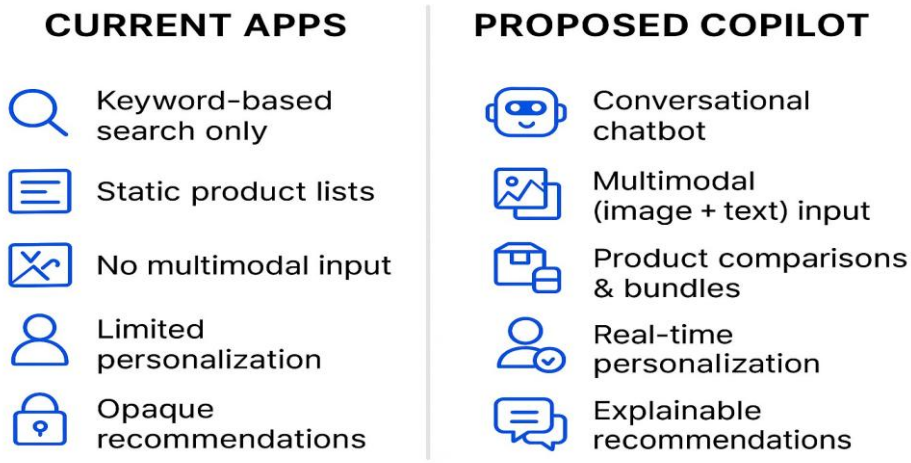
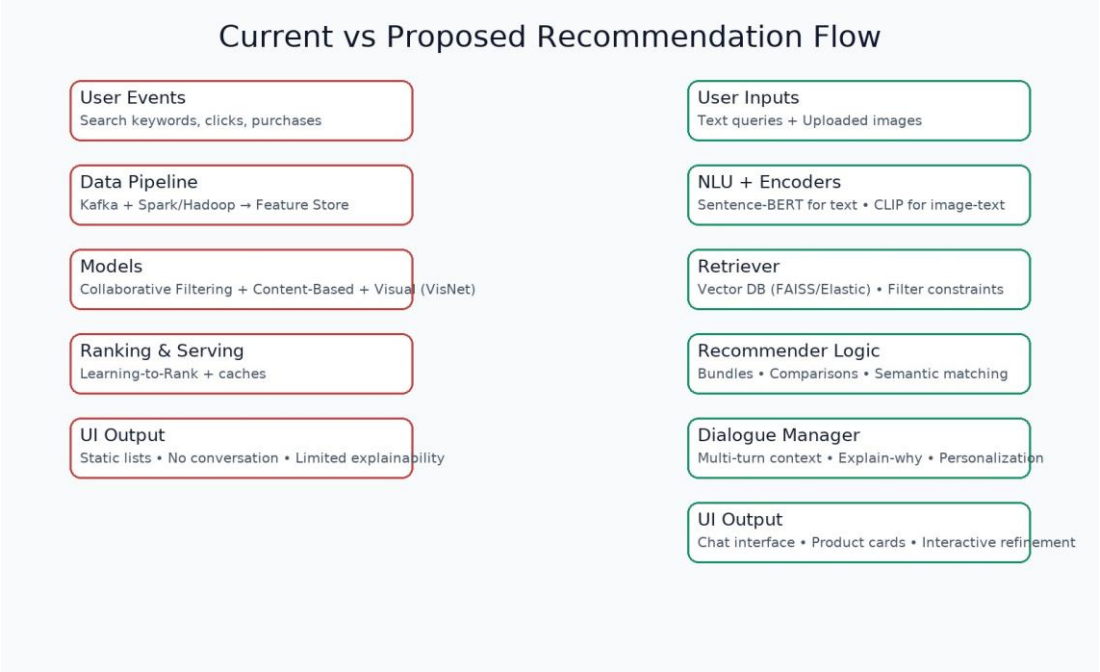


Figure 2: Current vs Proposed Recommendation Flow



III. Innovation / Novel Contributions

- Multimodal integration: Unlike Flipkart/Amazon, users can query with both image + text.

- Conversational refinement: Multi-turn chat instead of re-searching.

- Explainable recommendations: "This shoe matches your photo and fits under 2000."

- Dynamic bundling & comparisons: AI suggests complementary items (shirt + accessories).

- Real-time personalization: Adaptive to mood, tone, and shopping context.

IV. Conclusion

The proposed Personalized AI Shopping Copilot shifts the paradigm of e-commerce from static search to conversational discovery. By combining multimodal AI models (CLIP, BERT) with conversational LLMs (LLaMA, DialoGPT), the system enhances personalization, trust, and transparency. Future work involves building a prototype and evaluating its performance against existing e-commerce systems in terms of accuracy, user trust, and satisfaction.

VIII. References

[1] Jindi Fu, Samar Mouakket, and Yuan Sun, "The role of chatbots' human-like characteristics in online shopping", Electronic Commerce Research and Applications, vol. 61, p. 101304, August 2023. doi:10.1016/j.elerap.2023.101304. Link.

[2] Sidlauskiene, Justina, Yannick Joye, and Vilte Auruskeviciene, "AI-based chatbots in conversational commerce and their effects on product and price perceptions", Electronic Markets, vol. 33, no. 1, May 2023. doi: 10.1007/s12525-023-00633-8. Available: ResearchGate.

[3] Badave P, Bhomaj B, Bindu B, Shivarkar R, Prof Dhavase N,“E-commerce website with recommendation system including chatbot and reverse image search”, International Journal for Research in Applied Science & Engineering Technology (IJRASET), vol. 10, no. 9, pp. 1663-80., Sep, 2022. Available: Link.

[4] M. Rahevar, Maharshi, and S. Darji, “The adoption of AI-driven chatbots into a recommendation for e-commerce systems to targeted customer in the selection of product”, International Journal of Management, Economics and Commerce, vol: 1, no. 2, 128-137, 2024. Available: Link.

[5] Valencia-Arias, Alejandro, Hernán Uribe-Bedoya, Juan David González-Ruiz, Gustavo Sánchez Santos, Edgard Chapoñan Ramírez, and Ezequiel Martínez Rojas. "Artificial intelligence and recommender systems in e-commerce. Trends and research agenda." Intelligent Systems with Applications 24, p.200435, Dec 2024. doi: 10.1016/j.iswa.2024.200435. DOI

[6] Priya , and Dr Nidhi Bhagat, “The impact of AI-powered chatbots on shopper experience in e-commerce,” International Journal of Creative Research Thoughts (IJCRT), vol. 13, April 2025. Available: Link

[7] Dwivedi, Rohit, Abhineet Anand, Prashant Johri, Arpit Banerji, and N.K Gaur. "Product based recommendation system on amazon data." Int J Creat Res Thoughts–IJCRT, June 2020. Available: Link

[8] Illescas-Manzano, María, Sergio Martínez-Puertas, Paulo Ribeiro Cardoso, and Cristina Segovia-López. "Use of Online Shop Chatbots: How Trust in Seller Moderates Brand Preference and Purchase Intention." In International Conference on Advanced Marketing Practice, pp. 151-171. Cham: Springer Nature Switzerland, Nov 2024. DO**I**