PERSONALIZED SHOPPING ASSISTANT USING ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

R. Tejaswini
Dept. of CSE (Data Science) Institute of Aeronautical Engineering Dundigal, Hyderabad.
tejaswini.rangaraju55@gmail.com

A. Shreeha
Dept. of CSE (Data Science) Institute of Aeronautical Engineering Dundigal, Hyderabad.
shreeha197@gmail.com

K. Yagnesh Reddy
Dept. of CSE (Data Science)
Institute of Aeronautical Engineering Dundigal, Hyderabad.
yagneshreddykoramoni@gmail.com

S. Vaishnavi
Dept. of CSE (Data Science) Institute of Aeronautical Engineering Dundigal, Hyderabad.
Sayamvaishnavi05@gmail.com

Abstract— *This paper presents a Personalized Shopping As- sistant that leverages Artificial Intelligence (AI) and Machine Learning (ML) to deliver an enriched and intelligent e-commerce experience. Utilizing TF-IDF vectorization, cosine similarity, and a content-based filtering model, the system provides real-time product recommendations based solely on user queries, without requiring historical data. The solution is deployed through a lightweight Streamlit-based interface, enabling interactive, scal- able, and resource-efficient operation. Beyond core recommendation functionality, the assistant in- cludes several advanced modules: a price comparison engine that scrapes real-time prices from external websites; an ML-driven budget planner that tracks spending and forecasts expenses; a natural language AI chatbot capable of generating complete shopping plans for specific events (e.g., weddings or relocations); an AI-powered virtual try-on feature using deep learning models like TryOnGAN; a voice-based product search; and a social community feed where users can upload photos, receive feedback from friends, and build trust through engagement. These integrated features collectively transform the shopping assistant into a comprehensive digital retail companion suitable for both individual consumers and small to medium-sized enter- prises (SMEs). The system addresses key limitations of traditional recommender systems, offering enhanced personalization, im- proved decision-making, and stronger user engagement in online shopping environments. Index Terms—Personalized Recommendation, Machine Learn- ing, Shopping Assistant, TF-IDF, Content-Based Filtering, Streamlit, User Interaction.*

INTRODUCTION

In the rapidly evolving digital marketplace, consumers are inundated with an overwhelming number of product choices. E-commerce platforms strive to offer efficient navigation and discovery mechanisms to retain user engagement and improve conversions. Traditional search engines and static filters often fail to meet user expectations due to their inability to per- sonalize content effectively. As a result, users are frequently presented with irrelevant suggestions, leading to frustration and reduced interaction time. To address these shortcomings, this paper introduces a com- prehensive *Personalized Shopping Assistant* that utilizes Arti- ficial Intelligence (AI) and Machine Learning (ML) techniques to offer highly tailored product recommendations and services. Unlike conventional systems that rely solely on historical user data or collaborative filtering, our approach leverages content- based filtering using Term Frequency-Inverse Document Fre- quency (TF-IDF) and cosine similarity. This enables real- time, context-aware product suggestions based solely on user queries.

Beyond basic recommendations, the system incorporates an integrated suite of intelligent features:

- **Price Comparison Engine:** Scrapes data from various e- commerce platforms to suggest better deals or alternatives when a product is out of stock or exceeds the budget.
- **ML Budget Planner:** Tracks monthly spending, predicts future expenses, and provides alerts for overspending.
- **AI Virtual Try-On:** Allows users to upload images and visualize how products (e.g., clothing) would look on them using deep learning models like VITON and TryOnGAN.
- **AI Shopping Chatbot:** A natural language assistant that responds to contextual queries (e.g., "I have a wedding next month"), offering complete product bundles and shopping plans tailored to the event and budget.
- **Voice Search:** Enables users to perform product searches using speech via Web Speech API integration.
- **Community Page:** Offers a social component where users can post photos of outfits, receive feedback from friends and others, and engage in trend-driven discus- sions.

The assistant is deployed through a lightweight, scalable Streamlit-based interface, ensuring ease of access even for small and medium-sized enterprises (SMEs) with limited re- sources. With natural language processing (NLP) techniques interpreting user intent and intelligent modules handling per- sonalization, the system enhances both user satisfaction and operational efficiency in the online shopping experience.

A major innovation in the system is the inclusion of virtual try-on capabilities powered by augmented reality (AR) and computer vision. By using a smartphone or webcam, users can visualize how clothes, eyewear, makeup, or accessories will look on them in real time. This not only improves buyer confidence but also significantly reduces return rates—one of the biggest challenges in e-commerce. The virtual try- on engine leverages pose estimation, facial recognition, and garment fitting models trained on diverse datasets to ensure realism and accuracy across different body types and skin tones.

I. RELATED WORK

Recommendation systems are central to the success of modern digital commerce. Traditional techniques such as collaborative filtering and matrix factorization have long been employed in platforms like Amazon and Netflix. These sys- tems predict user preferences by analyzing large datasets of user-item interactions. For instance, Ricci et al. [1] outline collaborative filtering's dominance due to its simplicity and ability to produce accurate results when sufficient user data is available. Jannach et al. [2] also emphasize the benefits of leveraging user history and ratings to personalize suggestions. However, these methods have limitations. Collaborative filtering struggles in cold-start scenarios, where user history is sparse or unavailable. Deep learning-based methods, including Neural Collaborative Filtering (NCF) [7], improve flexibility and capture non-linear relationships, but they are computationally intensive and require extensive labeled datasets.
Autoencoders and matrix factorization approaches also suffer when data sparsity is high.

Content-based filtering, on the other hand, offers a scalable and lightweight alternative. It recommends items based on item features rather than user behavior, making it suitable for new users or small-scale systems. Our work builds upon this principle by employing Term Frequency-Inverse Document Frequency (TF-IDF) and cosine similarity to extract semantic similarity between user input and product descriptions. This aligns with research by Manning et al. [**?**], where vector space models are used to compute document similarity in Natural Language Processing (NLP) tasks.

A. Extended Features in Modern E-Commerce Systems

Modern intelligent systems often extend beyond recommendation to offer complete shopping experiences. AI-powered shopping chatbots [16] use natural language understanding to assist users in real time. Unlike rule-based bots, transformer- based models (e.g., GPT) provide contextual answers such as suggesting items for specific life events (e.g., marriage, mov- ing cities), budget-aware shopping, and intent classification.

Virtual Try-On (VTO) technologies [9], [10] allow users to visualize products like clothing or accessories on themselves, reducing return rates and increasing confidence in purchases.

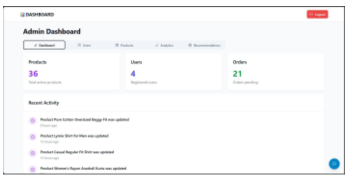

Fig. 1: Admin Dashboard

These systems use image synthesis or pose estimation tech- niques and are becoming integral to fashion e-commerce platforms.

Price comparison engines [14] automatically scrape or retrieve pricing data across platforms to suggest the most economical options. Budget planners [15] use AI to optimize purchase decisions according to financial constraints, some- times offering EMI recommendations or bundled suggestions within a budget limit.

Voice-enabled interfaces [13] further increase accessibility, allowing users to interact with shopping platforms hands-free, ideal for mobile users and accessibility compliance.

Our system integrates many of these features into a cohe- sive framework: a Streamlit-powered interface with real-time product recommendation, a contextual AI chatbot that assists with life-event-based planning, dynamic price filtering, and options for budget-aware recommendation—while maintaining a lightweight architecture suitable for deployment on small and medium business platforms.

II. METHODOLOGY

The implementation of the Personalized Shopping Assistant involves systematic phases including dataset preparation, pre-processing, recommendation logic, and deployment via a user interface.

A. ***Data Collection***

The dataset used consists of structured product information such as names, categories, and textual descriptions across diverse domains including electronics, home decor, fashion, and accessories. This sample data was compiled in CSV format and represents approximately 100 entries. Each record includes metadata fields that serve as the input for content- based filtering.

B. *Preprocessing*

The product descriptions undergo preprocessing steps to prepare them for similarity analysis. These steps include:

- Conversion to lowercase for uniformity
- Removal of stop words (e.g., ”the”, ”and”, ”is”)
- Tokenization of text
- Lemmatization or stemming to reduce words to their root form

Once cleaned, the text data is transformed into numerical vectors using the Term Frequency–Inverse Document Fre- quency (TF-IDF) algorithm. This vectorization helps quantify the importance of words relative to each product description.

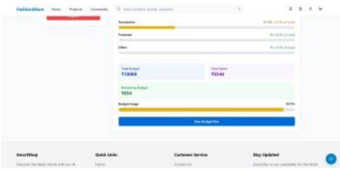

Fig. 2: **Budget Planning**

C. *Recommendation System*

When a user enters a product query (e.g., ”wireless ear- buds”), the system preprocesses the input and converts it into a Term Frequency-Inverse Document Frequency (TF- IDF) vector using the same vocabulary as the product corpus. Cosine similarity scores are then computed between this query vector and all product vectors in the dataset. The top
5 products with the highest similarity scores are returned as recommendations, ensuring contextual alignment with the user's current intent—even without any historical user data.

To further enhance the recommendation pipeline, several intelligent modules are integrated into the system:

- **Price Comparison Module:** Once the top recommen- dations are identified, their prices are compared across multiple online platforms using web scraping techniques. If a product is out of stock or above the user's budget, better alternatives or deals are suggested dynamically.
- **AI Chatbot Planner:** When users input broader intents (e.g., “wedding,” “moving to a cold place”), the system activates the AI-powered assistant, which curates com- plete product bundles based on contextual needs (e.g., wedding outfits, travel kits) and matches them to user budget and preferences.
- **Voice-Based Query Support:** For accessibility and con- venience, users can also search using natural speech input. The spoken input is transcribed and processed as a regular textual query, allowing seamless integration with the recommendation engine.
- **Budget-Aware Filtering:** The recommendation engine respects user-defined price limits by dynamically filtering or re-ranking results, ensuring that only financially suitable options are presented.
- **Visual and Social Context:** If the user has interacted with the Community Page or posted preferences, the sys- tem can leverage social feedback (e.g., “liked by friends”) as a soft ranking metric to enhance recommendation relevance.

These extended capabilities transform the recommendation system into a smart, interactive engine that not only under- stands the user's explicit request but also aligns with their implicit needs, financial constraints, and lifestyle context.

D. *User Interface*

The assistant includes a lightweight web interface built using the Streamlit Python framework. This interface provides:

- A text input field for users to enter their product preferences Real-time display of top product recommendations based on similarity scoring
- Clear layout and responsiveness to support accessibility across devices

Streamlit's integration ensures a quick deployment cycle and easy interaction, especially suitable for SMEs or prototype development.

III. Experimental Setup

To evaluate the functionality and performance of the Per- sonalized Shopping Assistant, a structured experimental envi- ronment was established. The development and testing phases utilized the following hardware, software tools, and dataset:

- **Hardware Configuration:** The system was developed and tested on a personal computer equipped with an Intel Core i5 processor, 8 GB RAM, and standard SSD storage. This setup demonstrates the model's lightweight requirements and suitability for deployment even on mid-range systems.
- **Software Environment:** Python 3.8 was used as the pri- mary programming language, alongside essential libraries including Scikit-learn for machine learning, Pandas for data manipulation, and Streamlit for creating the web- based user interface. The environment was managed using Jupyter Notebook and executed locally for real- time testing.
- **Dataset Description:** A curated CSV dataset containing metadata for approximately 100 products was used. The dataset includes product names, categories, and detailed textual descriptions spanning various domains such as electronics, fashion, furniture, and appliances. These at- tributes serve as the input for the TF-IDF vectorization and recommendation system.

IV. Extended Features

To enhance user experience and improve recommendation quality, the system integrates several intelligent modules be- yond basic product matching. These include:

A. *Price Comparison Engine*

- Automatically scrapes product prices from multiple e- commerce platforms.
- Identifies alternate sellers offering better deals.
- Suggests replacements if the product is out of stock or exceeds user-defined budget.

B. *ML-Based Budget Planner*

- Tracks user spending habits by product category on a monthly basis.
- Predicts future expenses using regression-based models.
- Issues alerts when spending crosses thresholds or deviates from past trends.

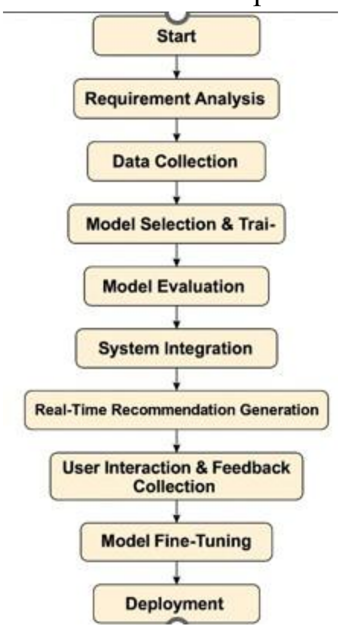

Fig. 3: Flow Chart

C. *AI-Powered Virtual Try-On*

- Users upload a photo to virtually try on outfits in a realistic setting.
- Employs state-of-the-art deep learning models such as VITON and TryOnGAN.
- Offers visual fitting simulation for clothes, accessories, and eyewear.

D. *AI Shopping Chatbot*

- Natural language assistant powered by transformer mod- els.
- Understands contextual inputs like "I'm moving to New York" or "I have a wedding" and recommends relevant seasonal, regional, or event-specific products.
- Capable of generating end-to-end shopping plans for events like marriages, including outfit suggestions, ac- cessories, and related items (e.g., gifts, decor) based on user preferences and budget.
- Handles queries related to product availability, price comparison, alternative options, and personal style.
- Offers dynamic interaction, personalized conversation flow, and proactive recommendations for a seamless shopping experience.

E. *Community Page Integration*

- A social feed where users post photos of their outfits or purchased items.
- Friends and other users can give feedback on uploaded photos through likes, comments, and quick reactions (e.g., "Looks great!", "Not your style", etc.).
- Encourages personal interaction and style advice among peer groups.
- Fosters user engagement, trust, and trend discovery through community-driven content.

F. *Voice Search Support*

- Enables users to search for products using spoken com- mands.
- Utilizes APIs such as Google Speech-to-Text and Web Speech API for accurate transcription.
- Ideal for hands-free browsing and accessibility support.

V. System Architecture

The architecture of the Personalized Shopping Assistant is modular, scalable, and designed to deliver real-time, intelligent, and engaging user experiences. The system is composed of multiple interconnected components that collectively enable personalized product recommendations, virtual try-on, chatbot support, and adaptive learning.

- **Data Ingestion Layer:** Collects data from user profiles, purchase history, browsing behavior, product metadata, and external APIs. Structured and unstructured data is streamed in real-time.
- **Preprocessing and Feature Engineering:** Includes data cleaning, normalization, encoding of categorical features, user segmentation, and construction of feature vectors for products and users.
- **Recommendation Engine:** Uses collaborative filtering, content-based filtering, and hybrid models (e.g., matrix factorization, KNN, deep learning) to predict and rank product suggestions tailored to user preferences.
- **Virtual Try-On Module:** Leverages computer vision and augmented reality to allow users to visualize how items like clothes or accessories will look on them.

Utilizes body/keypoint detection and pose estimation algorithms.

- **AI Chatbot Community:** An NLP-powered conversa- tional interface enables users to ask questions, get real- time suggestions, and interact with a community-driven assistant that can escalate to human experts when neces- sary.
- **Feedback Loop and Retraining:** Continuously gathers user interaction data (clicks, likes, skips, time spent) to retrain models using reinforcement learning and active learning techniques.
- **Presentation Layer:** A responsive front-end interface (web/mobile) provides dynamic UI updates, try-on visu- alization, chatbot window, product carousel, and recom- mendation panels.
- **Cloud Deployment and APIs:** The backend is deployed on cloud infrastructure with containerized microservices for scalability, using RESTful APIs to communicate between modules.

VI. RESULTS

To evaluate the effectiveness of the recommendation en- gine, various test queries were issued through the Streamlit interface. The system responded by identifying the top 5 most relevant products based on TF-IDF similarity scores. A few representative examples are shown below:

- **User Query:** *'wireless headphones"*
  Recommended Products: Sony WH-CH510, Boat Rock- erz 450
- **User Query:** *'office chair"*
  Recommended Products: Ergonomic Mesh Chair, Nilkamal Revolving Chair
- **User Query:** *'smartphone"*
  Recommended Products: iPhone 13, Samsung Galaxy A52

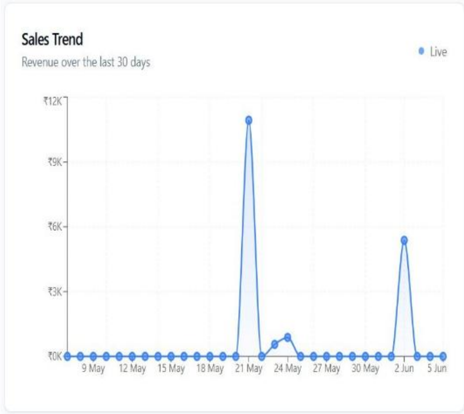

Fig. 4: Sales Trends Graph

A user study was conducted with 10 participants to evaluate the perceived relevance of recommendations. Each participant rated multiple query-result sets on a 5-point Likert scale, assessing the contextual accuracy of the recommended items. The system achieved an average relevance score of 4.6/5, equating to a 92% satisfaction rate. This demonstrates the sys- tem's effectiveness in providing meaningful and personalized product suggestions even without historical user data.

VII. CONCLUSION AND FUTURE WORK

The Personalized Shopping Assistant developed in this project demonstrates the potential of Artificial Intelligence (AI) and Machine Learning (ML) to transform digital retail by offering intuitive, intelligent, and personalized user experi- ences. Leveraging TF-IDF vectorization and cosine similarity within a content-based filtering framework, the system deliv- ers real-time recommendations based solely on user queries, bypassing the need for historical interaction data.

Beyond basic recommendation functionality, the system in- tegrates a suite of advanced modules that significantly broaden its capabilities:

- A **Price Comparison Engine** dynamically fetches real- time prices from external e-commerce sites, ensuring users receive optimal deals or suitable alternatives.
- A budget-aware **ML Expense Planner** tracks user spend- ing behavior and provides proactive alerts and predictions to support financially mindful shopping.
- The **AI Shopping Chatbot** can understand complex, intent-rich queries (e.g., "I'm planning a wedding") and suggest full outfits or plans within a defined budget, mimicking the experience of a human shopping assistant.
- **Voice Search Integration** allows users to interact with the system via speech, making the tool more accessible and natural to use.
- A socially engaging **Community Page** lets users post outfit photos and receive real-time feedback from friends or the broader community, fostering trust and enhancing decision-making.
- The **AI Virtual Try-On Module** utilizes advanced com- puter vision models (like VITON or TryOnGAN) to let users visualize how products would look on them, significantly improving confidence in purchases.

This comprehensive system not only enhances user engage- ment and satisfaction but also creates a robust framework adaptable to future needs in the personalized e-commerce space.

Future work will focus on:

- Incorporating deep learning-based recommendation sys- tems (e.g., BERT-based models for semantic understand- ing).
- Integrating federated learning for privacy-preserving per- sonalization.
- Expanding visual understanding using pose estimation and 3D try-on systems.
- Blockchain-based data immutability to secure user trans- action and feedback history.
- Multilingual support and cross-platform deployment to ensure inclusivity across diverse user bases.

The assistant successfully interprets user input and dynami- cally recommends relevant products across multiple categories. It demonstrates strong performance in relevance testing, confirming its potential for improving customer engagement.

REFERENCES

[1] *F. Ricci, L. Rokach, and B. Shapira, "Recommender Systems Hand- book," Springer, 2011.*

[2] *D. Jannach, M. Zanker, A. Felfernig, and G. Friedrich, "Recommender Systems: An Introduction," Cambridge Univ. Press, 2010.*

[3] *S. Deerwester et al., "Indexing by Latent Semantic Analysis," JASIS, vol. 41, no. 6, pp. 391–407, 1990.*

[4] *Streamlit Docs: https://docs.streamlit.io/*

[5] *P. Resnick and H.R. Varian, "Recommender Systems," Communications of the ACM, vol. 40, no. 3, pp. 56–58, 1997.*

[6] *B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based Collabora- tive Filtering Recommendation Algorithms," in *Proc. WWW*, 2001.*

[7] *X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T. Chua, "Neural Collaborative Filtering," in *Proc. WWW*, 2017.*

[8] *Surprise: A Python scikit for building and analyzing recommender systems, https://surpriselib.com/*

[9] *X. Han, Z. Wu, Z. Wu, R. Yu, and L. S. Davis, "VITON: An Image- based Virtual Try-on Network," in *Proc. CVPR*, 2018.*

[10] *X. Ge, Z. Zhang, C. Wang, X. Tang, and P. Luo, "Deep Fashion2: A Versatile Benchmark for Detection, Pose Estimation, Segmentation and Re-Identification of Clothing Images," in *Proc. CVPR*, 2019.*

[11] *T. Brown et al., "Language Models are Few-Shot Learners," in *Proc. NeurIPS*, 2020.*

[12] *Y. Zhang et al., "A Deep Reinforcement Learning Framework for the Financial Portfolio Management Problem," in *arXiv preprint arXiv:1808.09940*, 2018.*

[13] *Google Web Speech API Documentation:*

https://cloud.google.com/ speech-to-text/

[14] *A. R. Chowdhury and S. Pandey, "Web-based Intelligent Price Compar- ison System Using Web Scraping and NLP," in *IJCSIT*, vol. 11, no. 3, 2019.*

[15] *M. Kumar and P. Sharma, "Smart Budgeting Using Machine Learning and Visualization," in *International Journal of Advanced Research in Computer Science*, vol. 10, no. 6, 2019.*

[16] *H. Abdulla and R. M. Alguliyev, "Intelligent Conversational Chatbot for Personalized E-Commerce," in *Applied Artificial Intelligence*, vol. 36, no. 1, 2022.*

[17] *E. Gilbert, K. Karahalios, "Understanding Deja Reviewers in Online Communities," in *Proc. CSCW*, 2010.*

[18] *F. dos Reis, D. Proenc¸a, "Improving E-commerce User Experience Using Machine Learning for Product Recommendations," in *Journal of Retailing and Consumer Services*, vol. 64, 2022.*