iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VI **Month of publication:** June 2026

DOI: https://doi.org/10.22214/ijraset.2026.83701

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

ShopFusion: A Hybrid Retail Analytics and Recommendation System

Rameshwar Mane[1], Sarvesh Asawa[2], Omkar Kad[3], Sakshi Hudge[4], Vandana Dixit[5]
Progressive Education Society's Modern College of Engineering (PESMCOE), Pune, Maharashtra, India

Abstract: ***The rapid expansion of retail commerce—both online and offline—has generated vast volumes of customer transaction data that most retailers are ill-equipped to leverage effectively. Conventional decision-support systems rely on manual analysis or simplistic rule-based methods, leading to missed opportunities in product bundling, personalized recommendations, and inventory planning. The proposed ShopFusion system unifies the entire retail analytics workflow by integrating Market Basket Analysis (MBA), Content-Based Filtering (CBF), and Collaborative Filtering (CF) to transform raw purchase data into actionable business insights.***

The system employs the FP-Growth algorithm for efficient association rule mining, combines multiple recommendation paradigms through a configurable fusion layer, and delivers results through an interactive, retailer-friendly visualization dashboard. Built on a three-tier architecture using React.js for the frontend, Node.js with Express for the backend API, MongoDB as the database, and Python as the machine learning engine, ShopFusion provides dedicated modules for retailers and shoppers with secure JWT-based Role-Based Access Control (RBAC).

Experimental evaluation on a real-world grocery transaction dataset demonstrated that the system accurately identifies product co-purchase patterns, generates relevant personalized recommendations, and presents insights through intuitive visual dashboards. ShopFusion represents a scalable, open-source, and cost-effective solution that transforms retail decision-making from fragmented guesswork into an automated, data-driven ecosystem.

Keywords: Market Basket Analysis (MBA), Collaborative Filtering, Content-Based Filtering, FP-Growth algorithm, association rule mining, retail analytics, recommendation system, visualization dashboard, hybrid recommendation, data mining.

I. INTRODUCTION

The rapid growth of the retail industry and digital commerce has resulted in the generation of massive volumes of transactional and customer behavior data every day. Every purchase, browsing activity, product interaction, and customer preference contributes to an ever-expanding repository of valuable information. Despite the availability of such datasets, many retailers struggle to transform raw transactional data into meaningful business intelligence that can support strategic decision-making. This gap between data collection and data utilization remains one of the major challenges in the modern retail ecosystem.

Large-scale e-commerce platforms such as Amazon, Flipkart, and Walmart successfully utilize advanced recommendation systems and data mining technologies to improve customer engagement, optimize inventory management, and increase sales through personalized experiences. However, these sophisticated analytical capabilities are often inaccessible to small and medium-sized retailers due to high implementation costs, infrastructure limitations, lack of technical expertise, and the complexity of existing enterprise-level solutions. As a result, many retailers continue to depend on traditional methods such as manual auditing, static reports, spreadsheets, and basic Point-of-Sale (POS) analytics, which provide only limited insight into customer purchasing behavior.

Conventional retail analysis methods are insufficient for identifying hidden relationships among products, predicting customer interests, or understanding evolving buying trends. Retailers frequently face challenges such as inefficient inventory planning, missed cross-selling opportunities, poor product placement strategies, and inability to provide personalized product recommendations. Furthermore, fragmented analytical tools often fail to deliver an integrated view of customer behavior, making it difficult for businesses to make data-driven decisions in real time.

To address these limitations, intelligent recommendation systems and association rule mining techniques have emerged as powerful solutions in the field of retail analytics. Market Basket Analysis (MBA), one of the most widely used data mining approaches, helps identify relationships between products that are frequently purchased together. By analyzing customer transaction patterns, retailers can discover valuable associations useful for product bundling, promotional campaigns, shelf arrangement optimization, and targeted marketing strategies.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

ShopFusion is proposed as an integrated smart retail analytics and recommendation platform designed to bridge the gap between advanced data analytics and practical retail usability. The system combines Market Basket Analysis using the FP-Growth algorithm with hybrid recommendation techniques integrating Collaborative Filtering (CF) and Content-Based Filtering (CBF). The platform is further enhanced with an interactive visualization dashboard that enables retailers to explore sales trends, co-purchase relationships, customer preferences, and business insights through intuitive graphical representations.

II. METHODOLOGY

A. Requirement Analysis

- Identify user requirements for retailers, shoppers, and administrators.
- Study existing manual retail analytics processes and their limitations.
- Define functional and non-functional requirements such as accuracy, security, and real-time performance.

B. System Design

Architecture Design:

- Three-tier architecture: React.js frontend, Node.js/Express backend, Python ML engine, MongoDB database.
- JWT-based RBAC authentication module for retailer, shopper, and admin roles.
- Database schema for transaction records, product catalog, user profiles, and association rules.

Data Flow Diagrams:

- Level 0, Level 1, and Level 2 DFDs illustrate the flow of data between users, the system, and databases.
- ML engine modules for MBA, CBF, CF, and fusion layer are represented for visualization.

C. Module Development

a. User Authentication Module

- Retailer and Shopper login with JWT-secured session management.
- Role-based access control limiting dashboard features per user type.

b. Data Upload & Preprocessing Module

- Retailers upload CSV/Excel transaction datasets.
- System performs deduplication, normalization, and one-hot encoding.

c. Market Basket Analysis Module

- FP-Growth algorithm extracts frequent itemsets and association rules.
- Configurable support, confidence, and lift thresholds.

d. Hybrid Recommendation Engine

- Content-Based Filtering computes cosine similarity between product attribute vectors.
- Collaborative Filtering constructs a user-item matrix to find behaviorally similar customers.
- Fusion layer combines MBA, CBF, and CF signals with adjustable weights α, β, γ.

e. Visualization Dashboard Module

- Interactive charts using Plotly and Chart.js for bundling, cross-selling, and demand trends.
- Retailer-friendly UI requiring no data science expertise.

D. Implementation

- Frontend: React.js for responsive and interactive UI.
- Backend: Node.js with Express.js for RESTful API development.
- Database: MongoDB for flexible document-oriented data storage.
- ML Engine: Python with mlxtend, scikit-learn, pandas, and numpy.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

E. Testing

Perform unit testing, integration testing, and system testing.

F. Maintenance and Future Enhancements

- Regularly update product catalog and transaction datasets.
- Improve hybrid recommendation accuracy with larger, more diverse datasets.
- Integrate deep learning-based demand prediction in future iterations.

III. FLOW DIAGRAM AND SYSTEM ARCHITECTURE

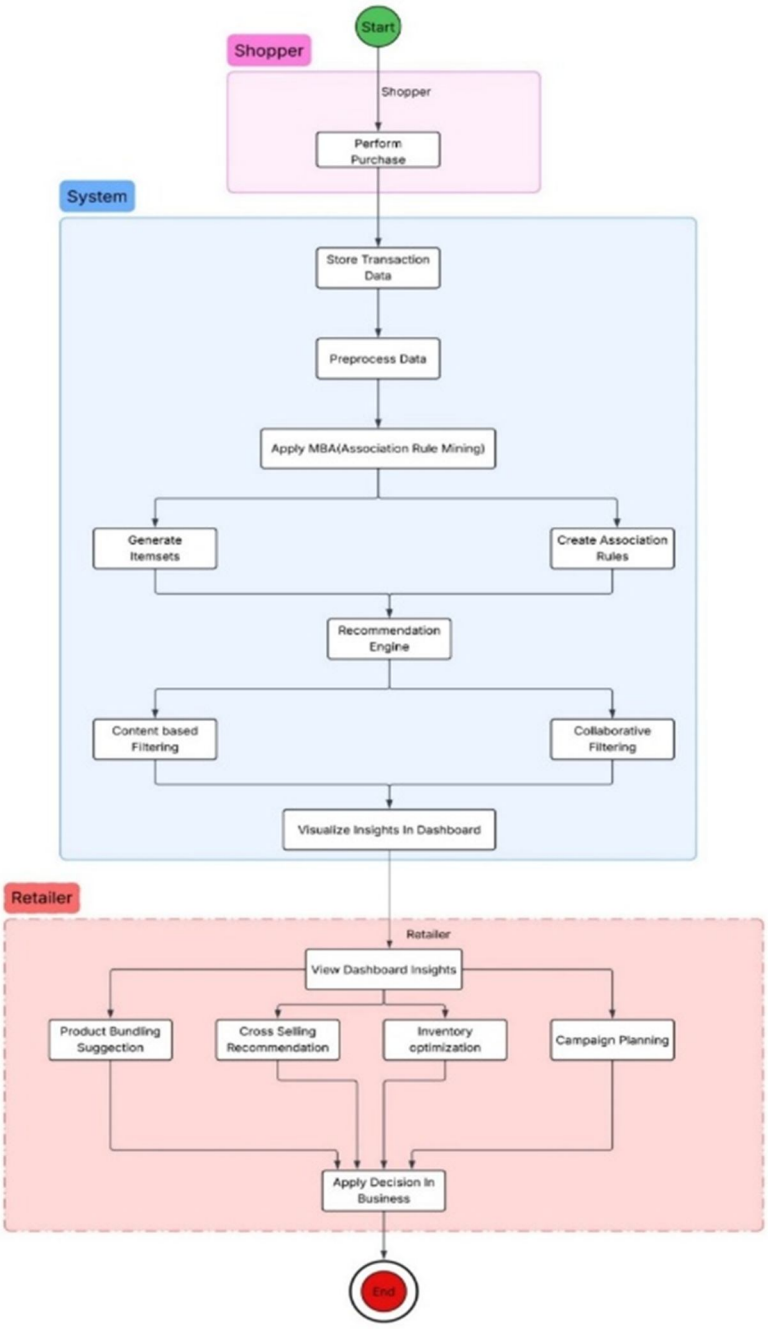


Figure 1: System Flow Diagram of ShopFusion

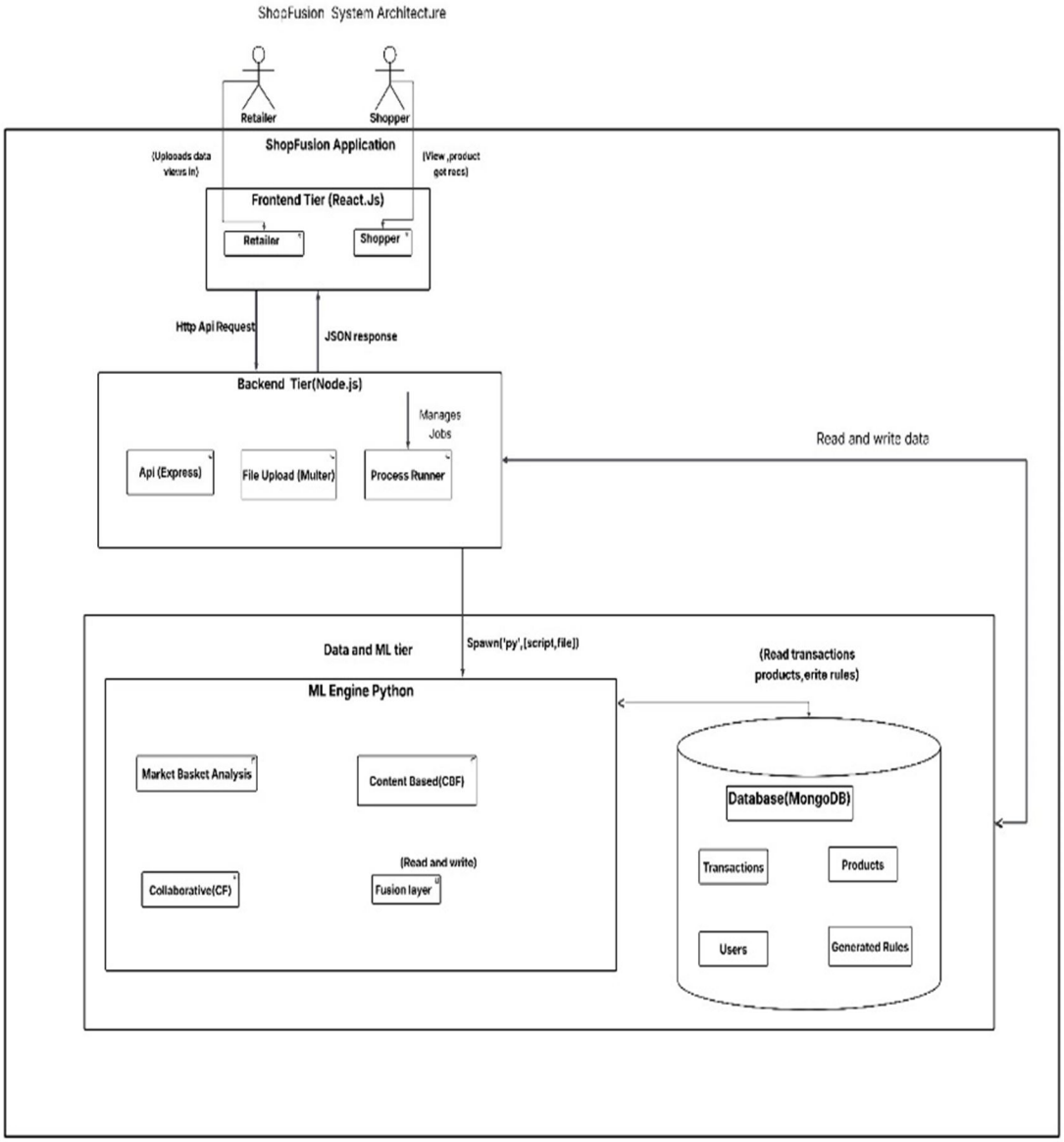


Figure 2: System Architecture of ShopFusion

IV. ALGORITHMS

A. FP-Growth Algorithm for Market Basket Analysis

The FP-Growth (Frequent Pattern Growth) algorithm is employed to extract frequent itemsets from the transaction dataset without generating candidate itemsets, making it significantly more efficient than the traditional Apriori approach for large retail datasets. The algorithm constructs a compact FP-Tree data structure from the transaction database, then mines the tree for frequent patterns directly. Association rules are generated from these patterns based on configurable minimum support, confidence, and lift thresholds, revealing products that are frequently purchased together.

B. Pseudocode (FP-Growth)

```
L1 = Find frequent 1-itemsets with support >= minsup
k = 2
while L(k-1) is not empty:
   Ck = Generate candidates from L(k-1)
   For each transaction T:
      For each candidate C in Ck:
         if C ⊆ T: increment count(C)
   Lk = {C ∈ Ck | count(C) >= minsup}
   k = k + 1
Return all Lk (Frequent Itemsets)
Generate Association Rules using Confidence and Lift thresholds
```

C. Hybrid Recommendation Fusion Layer

The fusion layer combines signals from three recommendation paradigms to produce ranked, explainable product suggestions. The final recommendation score is computed as:

$$Final\ Score = \alpha(MBA\ Score) + \beta(CBF\ Score) + \gamma(CF\ Score)$$

where α, β, and γ are adjustable weight parameters that enable domain-specific customization. This formulation ensures the system can leverage association rules, product attribute similarity, and behavioral similarity simultaneously.

D. Content-Based Filtering (CBF)

CBF converts product metadata into numerical feature vectors and computes pairwise cosine similarity between products. Products with the highest similarity scores to items in a shopper's current basket are recommended, ensuring relevance based on product attributes such as category, brand, and price range.

E. Collaborative Filtering (CF)

CF constructs a user-item purchase matrix to identify customers with similar purchase histories. Items purchased by behaviorally similar customers but not yet acquired by the target user are recommended, capturing latent preference patterns that content-based approaches alone cannot detect.

V. DATA FLOW DIAGRAMS (DFD)

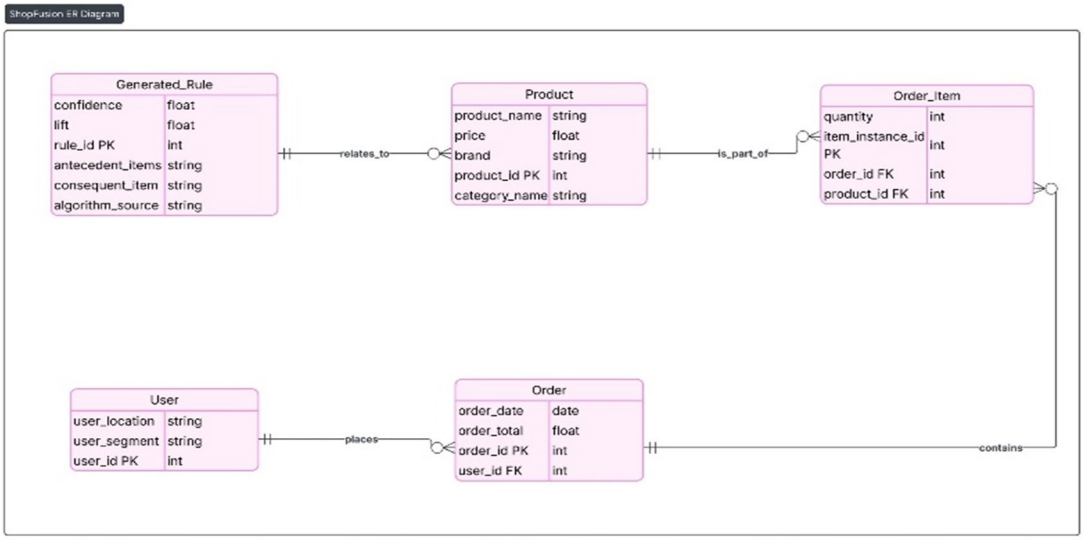


Figure 7: Entity Relationship (ER) Diagram

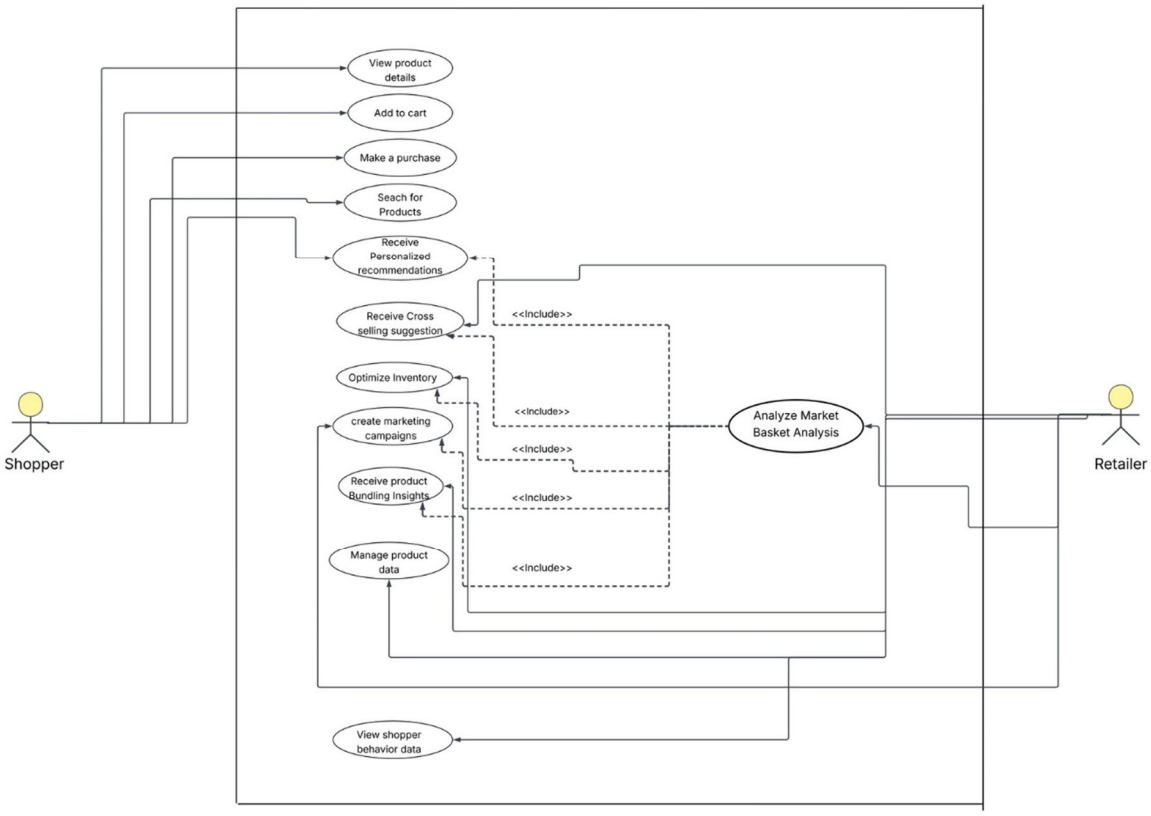


Figure 5: DFD Level 2 – Detailed Process Flow \

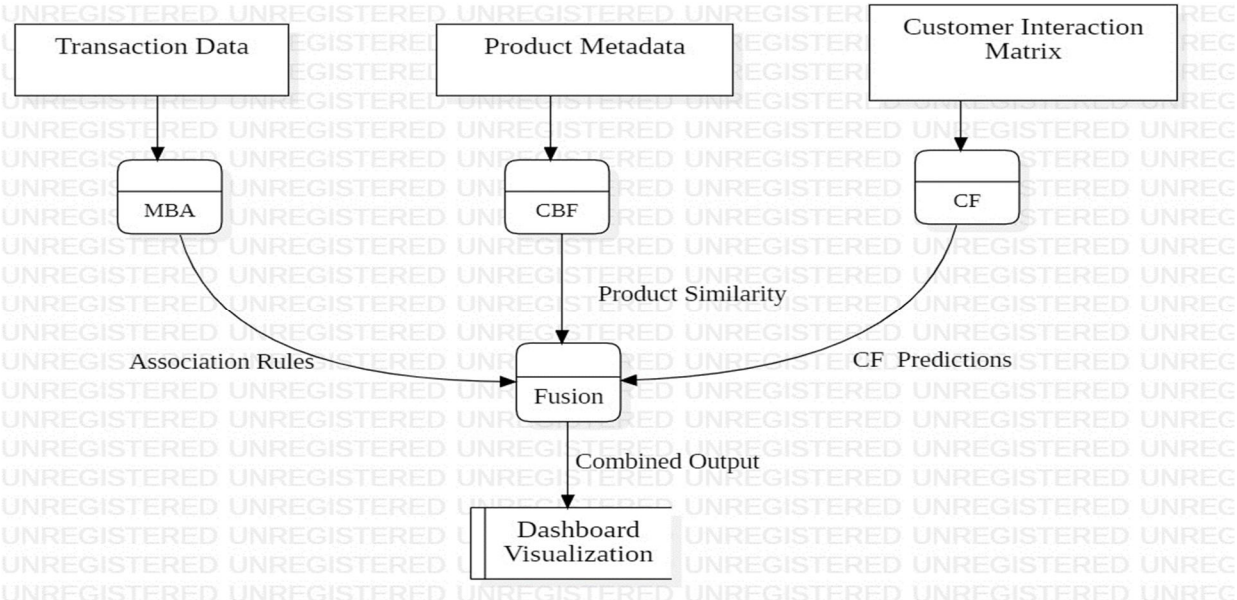


Figure 6: Use Case Diagram

VI. LIMITATIONS

1) Data Quality Dependency: The accuracy of association rules and recommendations depends heavily on the quality, completeness, and size of the uploaded transaction dataset. Sparse or noisy data may produce unreliable recommendation outputs.
2) Cold Start Problem: New shoppers with no purchase history cannot benefit from Collaborative Filtering-based recommendations until sufficient interaction data is collected, requiring fallback to content-based suggestions.
3) Computational Overhead for Large Datasets: FP-Growth performance may degrade for extremely large datasets (>1M transactions) without distributed computing infrastructure such as Apache Spark.
4) Static Weight Parameters: The fusion layer weight parameters α, β, and γ are manually configured and do not auto-optimize, which may require domain expertise to tune for optimal recommendation accuracy.
5) Internet Connectivity: The dashboard-based interface requires a stable internet connection for real-time visualization rendering when deployed in cloud environments.
6) No Real-Time Transaction Streaming: The current implementation processes uploaded batch datasets rather than live POS transaction streams, limiting the immediacy of insights for fast-moving retail environments.

VII. CONCLUSION

The proposed ShopFusion system provides an effective and technology-driven solution to enhance retail intelligence and streamline product recommendation for modern retail environments. By integrating FP-Growth-based Market Basket Analysis, Content-Based Filtering, and Collaborative Filtering into a configurable hybrid fusion engine, the system delivers accurate, transparent, and explainable product recommendations to both retailers and shoppers.

This project successfully addresses the limitations of traditional manual retail analytics by automating the entire intelligence pipeline—from data upload and preprocessing to association rule mining, recommendation generation, and interactive dashboard visualization. Retailers receive actionable insights about product bundling, cross-selling opportunities, and demand forecasting trends without requiring data science expertise, while shoppers benefit from personalized product suggestions tailored to their purchase behavior.

Overall, ShopFusion establishes a scalable, open-source, and cost-effective foundation for retail digital transformation. With future enhancements such as deep learning-based demand prediction, real-time streaming data support, and multi-language dashboard localization, the project has strong potential to evolve into a comprehensive smart retail intelligence ecosystem applicable to businesses of all sizes.

VIII. FUTURE SCOPE

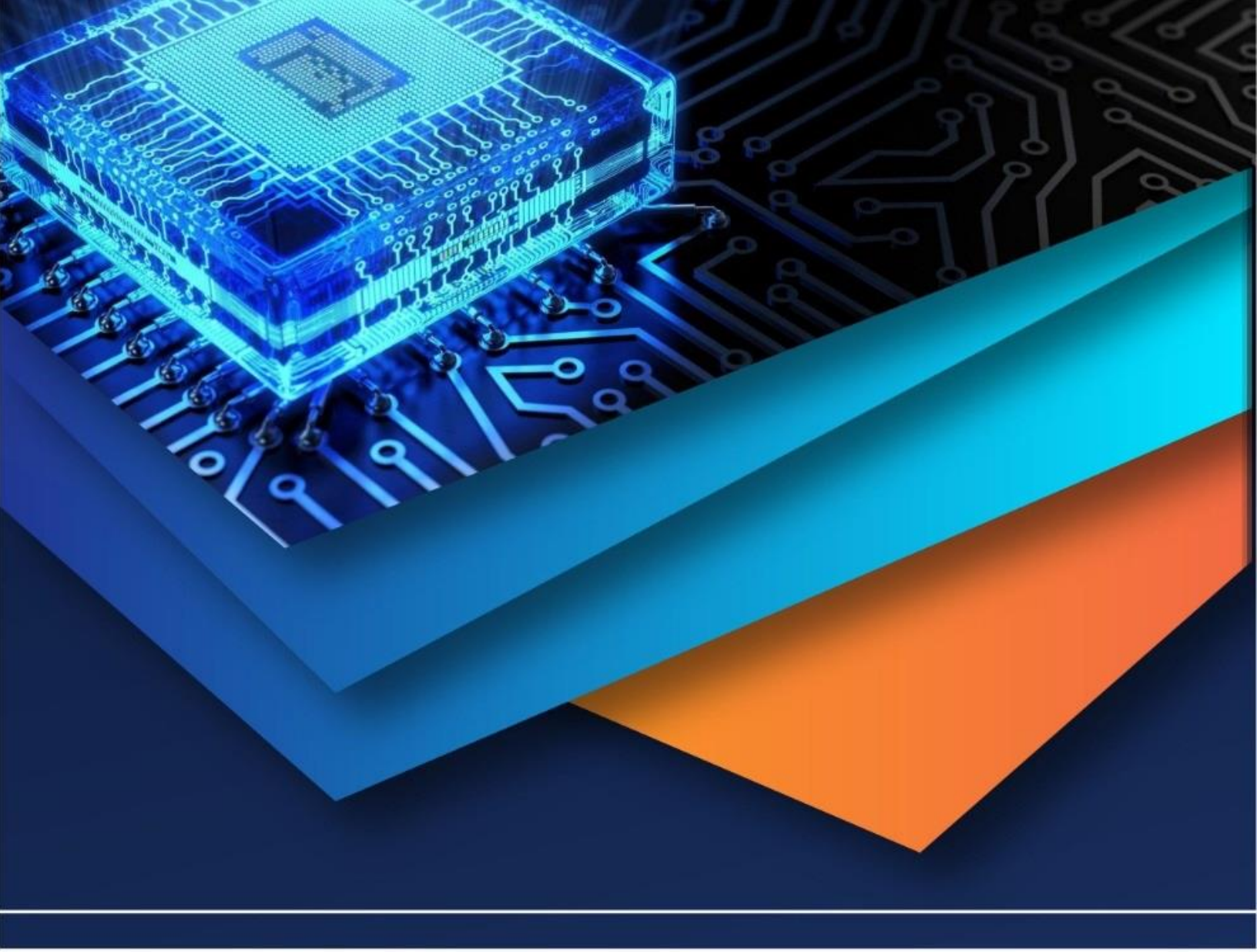
1) Deep Learning-Based Recommendation: Replace or augment the collaborative filtering module with neural collaborative filtering (NCF) or transformer-based sequential recommendation models to capture complex, long-range purchase sequence patterns.
2) Real-Time Transaction Streaming: Integrate Apache Kafka or Spark Streaming to process live POS transaction data, enabling instant recommendation updates as customers shop, rather than batch dataset uploads.
3) Mobile Application for Retailers and Shoppers: A dedicated mobile app can deliver instant bundling alerts, recommendation notifications, and inventory forecasting summaries, making the system more accessible for on-the-go retail management.
4) Multi-Language Dashboard Support: Adding support for regional languages will make the analytics interface accessible to diverse retail operators across different geographies.
5) Auto-Tuning of Fusion Weights: Implement Bayesian optimization or reinforcement learning to automatically tune the recommendation fusion weights α, β, γ based on continuously measured click-through and conversion performance.
6) Integration with ERP and POS Systems: Direct API integration with existing retail ERP platforms and point-of-sale systems will enable seamless two-way data synchronization, eliminating the need for manual CSV uploads.
7) Cloud-Based Multi-Tenant Deployment: Migrating to a Kubernetes-orchestrated cloud environment will enable multi-tenant deployment, allowing multiple retail organizations to use the platform concurrently with isolated data and customized configurations.
8) Sentiment Analysis for Product Enrichment: Incorporating NLP-based sentiment analysis on customer reviews will enrich product attribute vectors used in content-based filtering, improving recommendation relevance.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

REFERENCES

[1] M. A. Chatti, M. Guesmi, and A. Muslim, "Visualization for Recommendation Explainability: A Survey and New Perspectives," IEEE Access, vol. 11, pp. 12834–12852, 2023.
[2] I. Fister, D. Fister, and M. Mernik, "A Comprehensive Review of Visualization Methods for Association Rule Mining," IEEE Access, vol. 11, pp. 21145–21162, 2023.
[3] Kholod and Mokrenko, "Market Basket Analysis Using Rule-Based Algorithms and Data Mining Techniques," IEEE International Conference on Computational Intelligence and Data Science (ICCIDS), 2024.
[4] I. Sajwan and R. Tripathi, "Unveiling Consumer Behavior Patterns: A Comprehensive Market Basket Analysis for Strategic Insights," IEEE Xplore, 2024.
[5] S. M. Darshan, N. Raj, and B. Jaisachin, "Integrating Data Mining and Predictive Modeling Techniques for Retail Analytics," IEEE International Conference on Innovative Computing and Communication (ICICC), 2024.
[6] J. S. Breese, D. Heckerman, and C. Kadie, "Empirical Analysis of Predictive Algorithms for Collaborative Filtering," Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence, 1998.
[7] R. Agrawal and R. Srikant, "Fast Algorithms for Mining Association Rules in Large Databases," Proceedings of the 20th International Conference on Very Large Data Bases (VLDB), pp. 487–499, 1994.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)