

Online Purchasing Analytics and Recommendation System Using K-Medoids Clustering

Mary Anne B. Manandeg, Mary Jean Jayobo, Richard Morris Santos, Zharina Pascual, Cyrille Anne Nery

College of Computer Studies Quezon City University

DOI: <https://doi.org/10.51583/IJLTEMAS.2025.1409000071>

Received: 25 Sep 2025; Accepted: 02 Oct. 2025; Published: 09 October

Abstract: The rise of online shopping has transformed consumer behavior, generating large volumes of data related to preferences, purchases, and browsing activities. Businesses now rely heavily on analytics and recommendation systems to improve customer satisfaction, drive sales, and deliver personalized shopping experiences. Among different recommendation techniques, clustering methods like K-Medoids have gained attention for their reliability in grouping users and products. Unlike K-Means, K-Medoids uses actual data points as cluster centers, making it less sensitive to outliers and more stable in identifying purchasing patterns.

The proposed Online Purchasing Analytics and Recommendation System using K-Medoids clustering was designed to analyze user behavior, identify trends, and generate accurate recommendations. Its scope includes collecting and processing purchase history, browsing activities, product details, and demographics. A web-based dashboard was developed to visualize analytics and user clusters, while a recommendation engine provides personalized suggestions. The system, however, is limited to offline model training and excludes deep learning, advanced NLP, or real-time data integration. Experimental research was employed to measure the system's effectiveness, supported by ISO 25010 quality model-based evaluations. Both qualitative and quantitative tools, including Likert scale questionnaires, ensured systematic, reliable, and meaningful results on attributes like usability, reliability, and functional suitability.

The development process followed Agile (SCRUM) methodology, where tasks such as data preprocessing, feature extraction, clustering, and recommendation logic were iteratively built and refined. Regular sprint reviews and team collaboration ensured flexibility and continuous improvement. For pilot testing, Barangay 432 in Sampaloc, Manila, with 2,136 residents, was selected due to its active online shopping culture. Out of this population, 1,856 participants tested the system, alongside businesses, IT students, and faculty members, bringing the total respondents to 1,976. Evaluation results showed high usability (4.13), functional suitability (4.11), and security (4.07), while reliability (3.96) and efficiency (3.94) also performed positively. The overall weighted mean of 4.04 confirmed that users agreed the system met quality expectations, validating its readiness for real-world deployment.

In conclusion, the study demonstrates that K-Medoids-based recommendation systems significantly enhance online retail by providing accurate, personalized, and meaningful product suggestions. Retail Insight not only improves user satisfaction and shopping efficiency but also supports better stock management and eco-friendly product promotion. The study recommends expanding the system with hybrid recommendation models, dynamic dashboards, real-time sentiment analysis, and integration with sustainability databases for even greater personalization and environmental impact tracking.

Keywords: Online purchasing System, Analytics, Recommendation and K-Medoids Clustering

I. Introduction

In the digital era, online shopping has revolutionized the way consumers interact with products and services. E-commerce platforms generate vast amounts of data, capturing user preferences, purchase history, and browsing behavior. Analyzing this data is crucial for businesses aiming to enhance customer satisfaction and increase sales. Data-driven insights enable companies to provide personalized experiences that align with individual needs. Consequently, effective analytics and recommendation systems have become central to modern online retail strategies.

Recommendation systems serve as a bridge between consumer preferences and product offerings. By predicting what products a user may be interested in, these systems improve engagement and encourage repeat purchases. Several techniques, including collaborative filtering, content-based filtering, and clustering algorithms, are applied to model user behavior. Among clustering methods, K-Medoids offers a robust approach by identifying representative data points that minimize dissimilarity. This ensures more reliable grouping of users or products, leading to better recommendations.

According to Wu, Z. (2022), clustering techniques like K-Medoids help segment customers based on purchasing behavior, improving the accuracy of recommendations. Similarly, Bellini, P. (2022). demonstrated that K-Medoids clustering outperforms K-Means in handling noisy and outlier-prone data in e-commerce applications. These studies confirm that selecting representative medoids reduces computational complexity while maintaining recommendation quality. Hence, clustering-based recommendation systems can significantly enhance user experience and operational efficiency.

In accordance to Yıldız, Şen, & Işık((2023) research presents a differentially private user-based collaborative filtering recommendation system based on K-Means clustering, aiming to avoid performance degradation caused by differential privacy. Additionally, Kumar & Daruvurifound (2025) that segmenting users into meaningful clusters allowed personalized marketing

strategies that increased conversion rates. Such findings reinforce the potential of K-Medoids in understanding consumer patterns and predicting future purchases. Collectively, these studies support the adoption of clustering-based analytics for modern e-commerce environments.

K-Medoids clustering works by partitioning data into clusters, where each cluster is represented by a medoid that minimizes dissimilarity with other data points. Unlike K-Means, K-Medoids selects actual data points as centers, making it less sensitive to outliers and skewed distributions. In online shopping analytics, users or products with similar characteristics are grouped together to identify trends and preferences. These clusters can then be used to recommend items that a user is most likely to purchase. Consequently, K-Medoids provides a more stable and interpretable framework for recommendation systems.

Implementing a recommendation system with K-Medoids clustering involves several stages, including data collection, preprocessing, clustering, and recommendation generation. User interaction data, such as clicks, purchase history, and ratings, is first collected and cleaned. The algorithm then identifies clusters and selects representative medoids that capture the characteristics of each group. Finally, the system generates recommendations based on cluster similarity and individual user preferences. This process ensures that users receive relevant suggestions, increasing satisfaction and engagement on the platform.

Barangay 432 in Geronimo St., Sampaloc, Manila was chosen as the pilot testing community for the system. The barangay has a total voter population of 2,136 based on census data. Living in the community for 11 years, I was observed that residents frequently purchase products online, as delivery riders are seen daily in different households. This trend highlighted the growing reliance of the community on e-commerce. For this reason, an Online Purchasing Analytics and Recommendation System using K-Medoids Clustering was pilot tested in the area.

The primary goal of this study is to develop an online purchasing analytics and recommendation system that leverages K-Medoids clustering for accurate product suggestions. The system aims to analyze user behavior, group similar users, and provide personalized recommendations that enhance shopping experiences. It also seeks to reduce irrelevant suggestions, thereby improving the efficiency of marketing efforts. Additionally, the system can help businesses identify trends and make data-driven decisions.

Theoretical Framework

The study is grounded in data-driven decision-making and customer behavior analysis in e-commerce. Online purchasing generates large volumes of transactional and behavioral data, which can be leveraged to understand consumer preferences. Clustering algorithms, specifically K-Medoids, are applied to segment users based on similarity in purchase patterns, browsing history, and product interactions. Unlike K-Means, K-Medoids selects actual data points as cluster centers, making it more robust to outliers and noisy data. By grouping similar users or products, the system can generate personalized recommendations that match user interests and improve shopping experiences. The framework integrates analytics, clustering, and recommendation algorithms to provide insights that support both business decision-making and enhanced customer engagement. Overall, this study relies on the principles of unsupervised learning and behavioral modeling to create a system that predicts and recommends products effectively.

Conceptual Framework

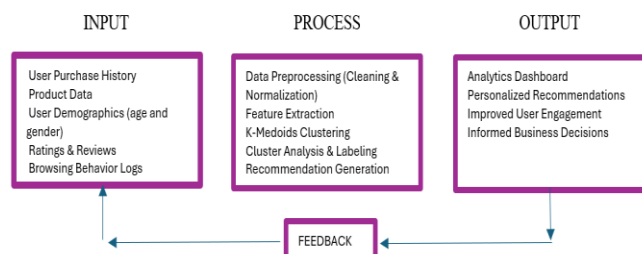


Figure 1: IPO of Online purchasing Analytics and Recommendation System using K-Medoids Clustering

Input: The system gathers various forms of data such as user purchase history, product information, and demographic details like age and gender. It also collects user feedback through ratings and reviews, along with browsing behavior logs. These data sources provide valuable insights into customer preferences and interaction patterns.

Process: The collected data is first cleaned and normalized to ensure accuracy and consistency. Key features are extracted and then processed using the K-Medoids clustering algorithm to group users based on similar behaviors and preferences. Each cluster is analyzed and labeled to generate meaningful insights and personalized product recommendations.

Output: The system outputs an interactive analytics dashboard that displays customer segmentation and behavior trends. It provides personalized product recommendations tailored to each user group, enhancing the shopping experience. These outputs help businesses make informed decisions and improve overall user engagement and satisfaction.

Feedback: feedback in the form of ratings and reviews is collected alongside purchase history, product data, demographics, and browsing logs. This feedback provides direct insights into user satisfaction and product performance. It serves as a critical input for understanding user preferences and improving recommendation accuracy.

Scope and Limitations

The Scope of this system encompasses the analysis of user purchasing behavior and the generation of personalized product recommendations through clustering techniques.

It covers the collection and processing of various data inputs such as user purchase history, browsing behavior, product details, and user demographics. The system utilizes K-Medoids clustering to segment users into behavior-based groups, enabling the identification of purchasing patterns and preferences.

It includes the development of a web-based dashboard to visualize analytics and user clusters, and a recommendation engine that provides tailored product suggestions based on cluster affiliation.

Limitation: It does not include real-time user interaction tracking or integration with external recommendation APIs. Furthermore, the project focuses solely on offline model training and analysis and does not incorporate deep learning or advanced natural language processing.

The Respondents of This Study

1) Community at the barangay 432 in Geronimo St., Sampaloc, Manila of an online shopping platform: Customers benefit from the system through personalized product recommendations that match their preferences and behavior. This enhances their shopping experience by making it more convenient, relevant, and engaging.

2) Administrators or system managers: Administrators gain access to insightful analytics dashboards that display user behavior, product trends, and cluster profiles. These insights help in optimizing marketing strategies, inventory management, and overall platform performance.

3) Future researchers: Future researchers can build upon this system by exploring advanced clustering techniques or integrating real-time feedback mechanisms. The current framework provides a solid foundation for developing more intelligent and adaptive recommendation systems.

Hypothesis:

The implementation of K-Medoids clustering in an online purchasing analytics and recommendation system significantly improves the accuracy and relevance of product recommendations compared to traditional non-clustering methods.

You can also frame it in null and alternative forms for research purposes:

Null Hypothesis (H_0): The use of K-Medoids clustering does not significantly affect the accuracy or relevance of product recommendations in online purchasing systems.

Alternative Hypothesis (H_1): The use of K-Medoids clustering significantly improves the accuracy and relevance of product recommendations in online purchasing systems.

Related Studies and Literature

According to Deng et al. (2019) The study proposes an enhanced K-medoids algorithm utilizing Kullback-Leibler divergence to improve recommendation accuracy in collaborative filtering. The approach effectively distinguishes items by considering asymmetric relationships, enhancing the quality of recommendations.

As stated by Salunke & Nichite (2022) that role of recommender systems in e-commerce, emphasizing their importance in personalizing user experiences and driving sales. It provides insights into various algorithms, including clustering methods, for effective recommendation system design.

Based on the findings of Widiyaningtyas et al. (2021) that K-means clustering with user behavior score-based similarity to enhance recommendation accuracy. Experimental results demonstrate improved performance metrics, showcasing the effectiveness of clustering in recommendation systems.

As highlighted by Widiyaningtyas et al. (2021) The study combines K-means clustering with user behavior score-based similarity to enhance recommendation accuracy. Experimental results demonstrate improved performance metrics, showcasing the effectiveness of clustering in recommendation systems.

In the study by Zhang et al. (2024) that the use of K-means clustering algorithm to enhance commodity information management in e-commerce. The approach facilitates better organization and management of product data, leading to more accurate recommendations.

According to Chen et al. (2008) that the application of data mining techniques, including clustering, to offer personalized recommendation services. It highlights the potential of clustering methods in enhancing the relevance of recommendations in various domains.

Synthesis

The reviewed literature and studies collectively highlight the effectiveness of clustering algorithms, particularly K-Medoids, in enhancing e-commerce recommendation systems. By grouping users or products based on behavioral and transactional data, K-Medoids provides accurate segmentation that improves personalized recommendations. Several studies demonstrate that integrating user behavior analytics with clustering methods increases recommendation relevance, conversion rates, and customer satisfaction. Moreover, K-Medoids' robustness to outliers and noise ensures stable and reliable performance in dynamic online shopping environments. Overall, these findings support the adoption of clustering-based analytics as a critical tool for data-driven decision-making and optimized recommendation strategies in e-commerce platforms.

II. Research Methodology

Experimental research was used for this system because it measures the effectiveness of K-Medoids in clustering customers or products based on purchasing behavior and it helps businesses improve customer segmentation, product recommendations, and targeted marketing strategies.

Tests and assessments, along with structured research methods, played a vital role in developing and validating the ISO model-based questionnaires used in this study. These tools ensured that the evaluation process was systematic, reliable, and aligned with recognized quality standards. The research incorporated both qualitative and quantitative approaches to gain comprehensive insights into user perceptions and system performance. Rating scales were utilized to measure specific attributes such as usability, functionality, and user satisfaction. In particular, the Likert Scale was employed to capture the degree of agreement or disagreement with various statements, providing a nuanced understanding of user feedback. Overall, these methods contributed to producing accurate, consistent, and actionable results that supported the effectiveness of the Online Purchasing Analytics and Recommendation System.



Figure 2: Agile Scrum methodology software

The development of the Online Purchasing Analytics and Recommendation System using K-Medoids Clustering followed the Agile (SCRUM) methodology to ensure flexibility and continuous improvement. The project began with a product backlog, where core features such as data collection modules, clustering algorithms, and recommendation logic were defined. During each sprint, specific tasks were selected and implemented, including data preprocessing, feature extraction, and integration of the K-Medoids algorithm. Daily stand-up meetings were conducted to monitor progress, identify obstacles, and encourage team collaboration. After each sprint, the team held a sprint review and retrospective to evaluate the results, gather feedback, and refine upcoming tasks. This iterative and collaborative approach ensured that the system was user-focused, adaptable to changes, and developed in manageable, functional increments.

The *ISO 25010 software* quality model was deployed to evaluate the Online Purchasing Analytics and Recommendation System using K-Medoids Clustering, focusing on key quality attributes. Functionality was assessed to ensure the system met its intended purpose, accurately performing data clustering and generating personalized recommendations. Reliability measured the system's ability to maintain consistent performance under various conditions without failures or data loss. The system's usability was evaluated based on how easily users, both customers and administrators, could navigate and interact with its features. Efficiency looked at the system's resource usage and response time, ensuring it handled large datasets and user interactions quickly and effectively. Security was also considered, particularly in protecting sensitive user information and ensuring secure access to the system. Overall, using ISO 25010 provided a structured and comprehensive evaluation that validated the system's quality, user satisfaction, and readiness for deployment.

The System

The system discussion provides a detailed examination of the Online Purchasing Analytics and Recommendation System using K-Medoids clustering. It highlights the system's core features, including data collection, clustering algorithms, and recommendation generation tailored to user preferences.

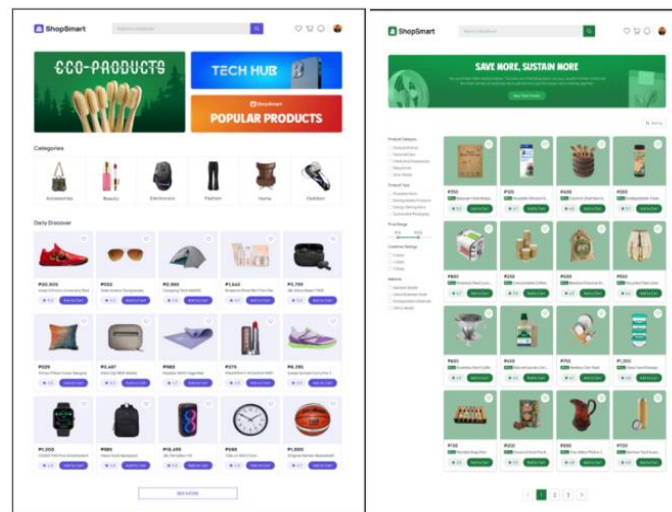


Figure 3: Landing Page(left) & Eco-Product Page(right)

After signing in, the Retail Insight system landing page highlights eco-products, a tech hub, and popular sections to capture user attention. A well-structured navigation bar allows users to browse categories such as fashion, electronics, and sustainable items with ease. The "daily discover" section uses K-means clustering to provide personalized product recommendations based on past behavior and preferences. On the Eco-Products page, users can explore sustainable items with advanced filter and sorting options for category, type, price, ratings, and relevance.

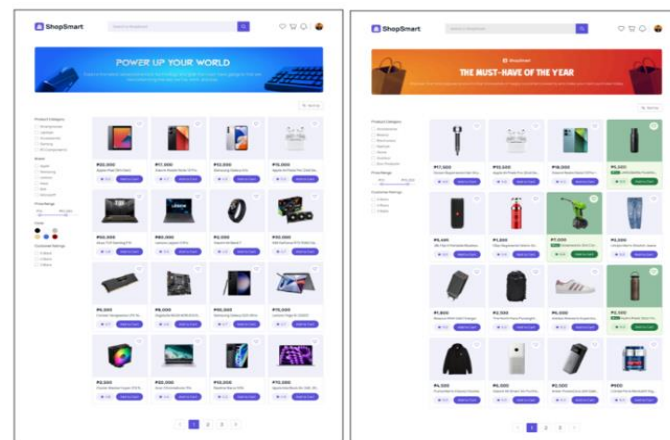


Figure 4: Tech Hub Page(left) & Popular Products Page(right)

The Tech Hub page of the Retail Insight system showcases the latest gadgets, smart devices, and accessories with filter and sort options for easy navigation. Meanwhile, the Popular Products page highlights trending items, bestsellers, and top-rated products, offering filters and personalized recommendations based on user preferences.

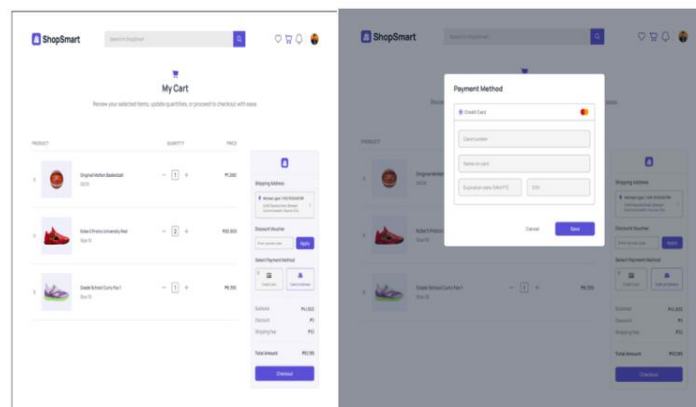


Figure 4: My Cart/Add to Cart(left) & Payment Method(right)

The My Cart/Add to Cart page lets users review selected items with details like name, image, quantity, price, and total cost, while offering options to update, remove, or save products before checkout. A Payment Method pop-up then prompts users to enter or save credit card details, ensuring a smooth and convenient transaction process.

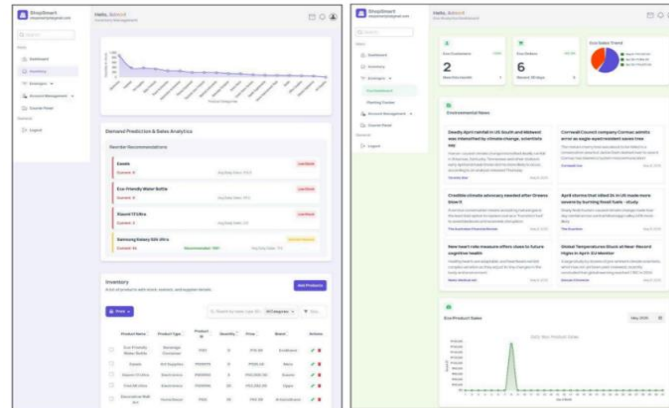


Figure 5: Admin Inventory Page(left) & Eco Admin Page (right)

The Admin Inventory page provides tools for managing stock, tracking product performance trends, and performing actions like adding, editing, deleting, or printing product lists, supported by demand prediction and sales analytics. Meanwhile, the Eco Admin Page focuses on managing eco-friendly operations by showing insights into eco customers, orders, and sales, along with sustainability news and detailed reports on eco products and inventory.

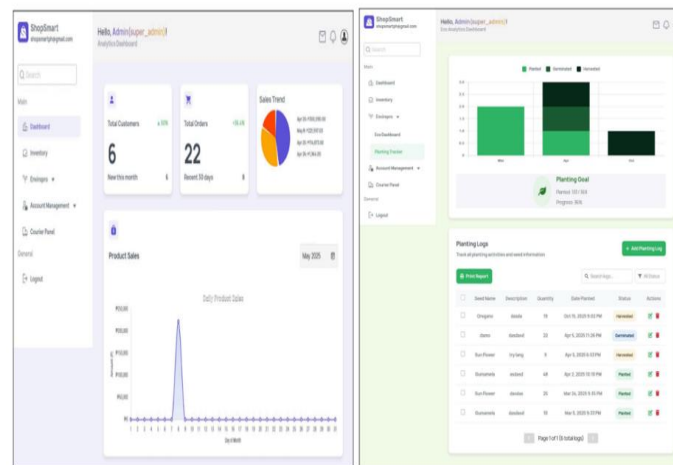


Figure 6: Admin Dashboard(left) & Plant Tracker(right)

The Admin Dashboard page offers a centralized view of platform performance with key metrics like customers, orders, sales, inventory, and predictive analytics, alongside quick access to major admin sections for efficient management. The Plant Tracker page supports sustainability efforts by displaying tree statistics, providing planting logs for updates, and offering a print option for offline reporting and analysis.

Ethical Considerations

- The system must ensure user data privacy by collecting and processing personal and transactional information securely and with explicit consent.
- Recommendations generated by the system should avoid bias or discrimination, providing fair and equal exposure to all products and users.
- Transparency in data usage and algorithmic processes is essential to maintain user trust and accountability in decision-making within the recommendation system.

III. Result and Discussion

Barangay 432 in Geronimo St., Sampaloc, Manila served as the pilot testing community for the system. The barangay has a total voter population of 2,136 based on census data. Out of this number, 1,856 participants were evaluated the system. The evaluation was conducted using the ISO 25010 quality model as the framework. The study specifically assessed the system across six key

characteristics: functional suitability, performance efficiency, reliability, usability, and security. Respondents came from different age brackets, ensuring a more diverse and representative feedback. This pilot testing provided valuable insights into how the system performs under real community conditions. In addition to the participants from Barangay 432, other groups were also included in the evaluation of the system. These groups consisted of 10 businesses, 93 fourth-year IT students from various schools, and 15 faculty members. Their inclusion helped provide more diverse perspectives in assessing the system's quality. Altogether, a total of 1,976 respondents participated in the evaluation process.

Table 1: Interpretation of Data using the ISO 25010

Criteria	Weighted Mean	Verbal Interpretation	Rank
Functional Suitability	4.11	Agree	2
Reliability	3.96	Agree	4
Usability	4.13	Agree	1
Performance Efficiency	3.94	Agree	5
Security	4.07	Agree	3
Average Weighted Mean	4.04	Agree	-

The table shown that the evaluation of the Online Purchasing Analytics and Recommendation System using K-Medoids Clustering showed strong performance across ISO 25010 quality attributes, with an overall weighted mean of 4.04, indicating that users generally agreed the system meets quality standards. Usability scored the highest at 4.13, proving that users found the interface intuitive and easy to use, while functional suitability also scored high at 4.11, confirming effective task performance. Reliability (3.96) and performance efficiency (3.94) were slightly lower but still positive, suggesting stable operations with room for minor optimizations. Security was rated 4.07, reflecting user confidence in the protection of data and system access. With 1,976 diverse respondents—including residents of Barangay 432, businesses, students, and faculty—the evaluation provided comprehensive insights, validating that the system is reliable, user-centered, and ready for real-world deployment.

IV. Conclusion

Participants in the survey acknowledged that typical e-commerce websites could not give meaningful product recommendations, segment customers appropriately, and respond dynamically to user behavior. Deployment of Online purchasing Analytics and Recommendation System using K-Medoids Clustering customer purchase history and browsing habits to yield smart, meaningful, and focused product recommendations.

By using Retail Insight, companies can achieve the best possible customer satisfaction, improve stock management, and aid environmentally friendly product promotion. K-Medoids Clustering and refreshed recommendations dramatically enhanced shopping by boosting shopping interaction, effectiveness, and individualization for single customers.

V. Recommendations

The proponents recommend introducing newer visualization techniques involving dynamic charts and interactive dashboards to further advance customer trends and behavior.

The proponents recommend enriching the recommendation engine with hybrid models combining collaborative filtering and K-Medoids clustering to further improve accuracy in personalization.

The proponents suggest incorporating real-time sentiment analysis through customer reviews to offer yet more contextually appropriate product recommendations.

The proponents suggest investigating the use of Retail Insight in conjunction with green product databases and third-party sustainability APIs to enhance the environmental impact tracking capacity of the system.

References.

1. Chen Et Al. (2008), Using Data Mining To Provide Recommendation Service, ACM Digital Library
2. Emre Yıldız, Ceyda Güngör Şen, Eyüp Ensar Işık (2023) A Hyper-Personalized Product Recommendation System Focused On Customer Segmentation: An Application In The Fashion Retail Industry, Journal Of Theoretical And Applied Electronic Commerce Research
3. Jiangzhou Deng, Junpeng Guo, Yong Wang (2019), A Novel K-Medoids Clustering Recommendation Algorithm Based On Probability Distribution For Collaborative Filtering, Knowledge-Based Systems (Elsevier)
4. Kiran Kumar & Rajesh Daruvuri (2025), Enhancing Online Retail Insights: K-Means Clustering And PCA For Customer Segmentation

5. Pierfrancesco Bellini, Luciano Alessandro Ipsaro Palesi, Paolo Nesi, Gianni Pantaleo(2022)Multi Clustering Recommendation System For Fashion Retail, Multimedia Tools And Applications Are Provided Here Courtesy Of Nature Publishing Group
6. Salunke & Nichite (2022), Recommender Systems In E-Commerce
7. Widiyaningtyas Et Al. (2021), Recommendation Algorithm Using Clustering-Based User Behavior Score Similarity, Computers
8. Zengyuan Wu, Lingmin Jin, Jiali Zhao, Lizheng Jing, Liang Chen(2022), Research On Segmenting E-Commerce Customer Through An Improved K-Medoids Clustering Algorithm, Computational Intelligence And Neuroscience Wiles Pulication.
9. Zhang Et Al. (2024), E-Commerce Recommender System Based On Improved K-Means Clustering Algorithm , Pubmed Central