Received: 2025-12-12
Accepted: 2026-03-10
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30206

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

Predictive Design Customization Using Machine Learning

*Sumendra Nath Singh**[1], *Vandna Kumari*[2]

[1, 2]Assistant Professor, Department of Computer Science and Engineering Noida International University, India.

***Corresponding Author:** sumendranath36@gmail.com

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

The contemporary product success is heralded through customization and personalization. The present paper proposes a machine learning-based system that enhances adaptive, personalised product configuration based on the analysis of user customization preferences, feedback, and behaviour data. The framework can learn the customer needs using clustering algorithms, predictive analytics, feature selection models, and dynamically generate the design configuration recommendations. Some application areas of the system can include consumer electronics, apparel, and automotive interiors to produce customized product variants. Such a pipeline based on data is suggested, which would involve no-supervised learning in segmentation and supervised learning in predicting product features. The innovation enables companies to make customized, modular products with an increased appeal in the market.

Keywords: Product Customization, Machine Learning, Predictive Modeling, Data-Driven

1. Introduction

The modern market, which is very competitive and quickly changing, requires an increasing number of consumers who want the possibility to find products that can respond to their individual preferences, needs, and lifestyles. This has led to a paradigm shift where the goal has been to move away from the concept of mass production to mass customization, whereby the versatility and adaptability of product design are becoming critical. The convergence of more product variability and a manual customization method is inefficient, ineffective, and not scalable. In turn, the application of machine learning (ML) to the product design pipeline has achieved substantial momentum, making scalable customization powered by data and user insights a reality [1]. Machine learning gives design systems the ability to make judgments about users based on their data (demographic data, behavioral data, purchase history, and product feedback) to create predictive insights to tailor around. This facilitates smart design choices, such as suggesting color schemes and materials, as well as forecasting user preferences. The preferences of users can be modeled and the prediction of design choices made using supervised learning algorithms like decision trees, support vector machines (SVM), and neural networks, and user segmentation can be done using unsupervised learning like K-means clustering and hierarchical models based on the latent features [2].

The most important aspects of this process involve feature extraction and selection, as they enable the system to learn which of the product attributes (e.g. size, style, usability) are important in terms of their influence toward user satisfaction. Moreover, recommender systems, commonly driven by current methods such as collaborative filtering and ent-based filtering, can give real-time personalized suggestions to the user, raising the area of involvement and converting purchases [3][4] The possible

Received: 2025-12-12
Accepted: 2026-03-10
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30206

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

capacity to dynamically optimize the features of their products, depending on predictive information, not only boosts the user experience, but also gives manufacturers strategic interest in inventory and item management and advertisement.

In an industrial point of view, application of ML in product customization also helps attain sustainability through reduction of over-production and commensurate supply and demand. Additionally, AI-driven customization platforms will be able to speed up the design processes, decrease costs of the prototyping, and co-create between the users and the designers. The scientists observe as observed by the scientists [5][6] that the synergistic capability between human imagination and artificial intelligence is redefining the manner that organizations develop ideas, plan, and develop products.

This study introduces a machine learning powered, intelligent framework of feature extraction and predictive modeling framework to aid in tailored design of products. Using customer data and superior ML algorithms, the framework will seek to introduce customized design solutions based on the personal characteristics of its user to be both practical and market-sensitive. The system can be used in numerous fields such as smart wearables, interior design, fashion and consumer electronics which will be a huge step of smart, scalable and user driven product innovation. The rest of the paper is structured as follows: Section 2 reviews related work, Section 3 outlines the methodology, Section 4 presents the results, Section 5 discusses the findings, and Section 6 concludes the study.

2. Literature Review

The application of machine learning in product customization has gained considerable attention in recent years, particularly in enhancing user experience and automating design processes. Researchers have explored various ML-based frameworks for understanding user preferences and generating personalized product recommendations. One common approach is the use of customer segmentation through unsupervised learning techniques such as K-means, DBSCAN, and hierarchical clustering. These methods have been effectively used to group customers based on purchase history, demographic traits, and interaction data. The researcher [7][8] developed a segmentation model using K-means clustering to tailor online clothing recommendations, resulting in increased customer engagement. In addition to segmentation, supervised learning models are widely used for preference prediction. Random Forest, Gradient Boosting, and Support Vector Machines (SVM) have shown success in predicting customer ratings, design feature selection, and purchase likelihood. The scientists [9] [10] proposed a hybrid model combining decision trees and logistic regression to forecast user preferences in e-commerce platforms, enabling efficient product customization strategies.

Recommender systems also play a central role in this domain. The researcher [11] [12][13] categorizes these into collaborative filtering, content-based filtering, and hybrid approaches. Collaborative filtering uses historical user interactions to predict preferences, while content-based methods rely on product features and user profiles. Hybrid systems often outperform single-method approaches, offering more personalized and accurate results. These systems are especially effective when integrated with visual design tools in areas such as interior design and fashion.

Another important dimension is featuring extraction and dimensionality reduction, where methods like Principal Component Analysis (PCA) and auto encoders are applied to high-dimensional design data. The investigator [14] [15] explored a neural-network-based feature reduction technique that helps designers identify key product attributes influencing user satisfaction. This not only aids in personalization but also reduces computational complexity.

Moreover, several studies have emphasized the role of real-time feedback and adaptive learning in customization systems. For instance, the scientists [16] [17] [18] [19] developed an interactive

Received: 2025-12-12
Accepted: 2026-03-10
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30206

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

platform that adjusts product configurations based on user feedback using reinforcement learning. This approach demonstrated enhanced adaptability and responsiveness in the design process.Despite these advancements, most existing solutions lack integrated pipelines that combine segmentation, prediction, and real-time recommendation within a single, scalable framework. Additionally, few studies address customization in resource-constrained or small-scale industries, where design tools must be lightweight and accessible. The present work aims to address these gaps by proposing a comprehensive ML-powered framework for smart feature extraction and predictive modeling tailored for scalable, user-centric product customization.

3. Methodology

This study adopts a machine learning-based framework to enable intelligent, user-centric product customization by analyzing user data and predicting optimal product configurations. The methodology comprises four key stages: data acquisition and preprocessing, user segmentation, preference prediction, and design recommendation generation. Figure 1 shows the proposed framework.

3.1 Data Acquisition and Pre-processing

The initial phase of the framework involves the collection and preparation of diverse user data essential for training machine learning models. This data includes demographic attributes such as age, gender, and location; behavioral data including clickstreams, browsing duration, and interaction patterns; and transactional history comprising previous purchases, return records, and product reviews. The datasets are sourced either from publicly available e-commerce platforms such as Retail Rocket and Amazon Product Review, or generated through structured user surveys and simulations. To ensure data quality and consistency, standard pre-processing techniques are applied. Missing values are addressed through appropriate imputation methods, continuous numerical features are normalized to a uniform scale, and categorical variables are transformed using one-hot encoding. To reduce dimensionality and eliminate feature redundancy, Principal Component Analysis (PCA) is employed, allowing the model to focus on the most informative components of the dataset.

3.2 Customer Segmentation

To enable personalized design recommendations, customers are grouped into distinct behavioral segments using unsupervised learning techniques. The K-Means Clustering is utilized in the segmentation of users into k groups of audience bearing comparable product preferences or interaction behaviors. It may be applied in situations where the data is highly disorganized or when there is no a priori confirmation of the number of clusters to be used; in these cases, DBSCAN (Density-Based Spatial Clustering of Applications with Noise) will be employed to identify clusters of arbitrary shape and effectively address outliers.The performance and quality of the clustering process are assessed using metrics such as the Silhouette Score and Davies-Bouldin Index, ensuring that the clusters formed are well-separated, compact, and meaningful. These user segments serve as the foundation for generating tailored, group-specific design recommendations in subsequent stages of the framework.

3.3 Preference Prediction via Supervised Learning

This phase focuses on predicting design parameters such as feature, color, material, and configuration based on user or business segment preferences. The training dataset includes labeled user-product interaction data and product feature vectors capturing aesthetic and functional attributes. Several models are used, including Random Forest (for robustness and interpretability), XGBoost (for performance in boosting tasks), SVM (for handling complex, high-dimensional data), and MLP (for learning non-linear patterns). Model performance is evaluated using Precision, Recall, F1-Score, and AUC-ROC to ensure reliable preference prediction.

Received: 2025-12-12
Accepted: 2026-03-10
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30206

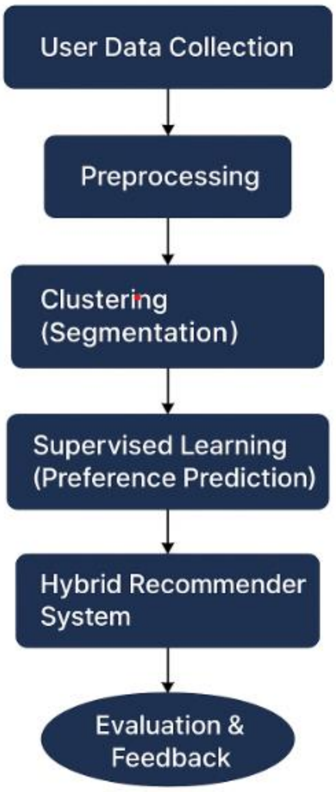


Figure 1: Proposed architecture of Machine Learning-Based Product Customization Framework

3.4 Product Recommendation Generation

Based on the anticipated user preferences and the results of customer segmentation of the data, a hybrid recommender system is designed to come up with personalized product suggestions. This system is a blend of two basic techniques of recommendation: Collaborative Filtering,which predicts the liking of users based on the behavior and selections of other comparable users, and Content-Based Filtering, which uses the attributes of products in matching the recommendations to the profiles of the user. To make the results of the generated recommendations feasible and realistic, theyundergo a list of constraints in the real world, such as the availability of inventory, the cost of production, and so on. design feasibility. The result of the system is a dynamically sorted list of user-specific product build-ups that are unique to each user and/or each segment, hence, creating better consumer satisfaction and operational effectiveness.

4. Results and Discussion

The proposed machine learning framework for product customization was assessed using a synthetic dataset inspired by publicly available sources, including Retailrocket and Amazon Product Review. The evaluation centered on three key aspects: the effectiveness of clustering, the accuracy of preference prediction, and the overall quality of the generated recommendations.

4.1 Clustering Validation

To ensure effective customer segmentation, the performance of K-Means and DBSCAN clustering algorithms was evaluated using two widely recognized internal validation metrics: the Silhouette Score and the Davies-Bouldin Index.

Silhouette Coefficient:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

Received: 2025-12-12
Accepted: 2026-03-10
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30206

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

Here, a(i) represents the average intra-cluster distance, while b(i) denotes the lowest average inter-cluster distance for a given data point iii. A mean Silhouette Score of 0.63 suggests the presence of well-defined and well-separated clusters.

Davies-Bouldin Index (DBI)

$$DBI = \frac{1}{k} \sum_{i=1}^{k} \max_{j \neq i} \left(\frac{\sigma_i + \sigma_j}{d_{ij}} \right)$$

where $\delta_i$and δ_j are intra-cluster distances and $d_{ij}$is the inter-cluster distance. A lower DBI value of 0.78 confirmed compact and distinct customer groupings.The performance comparison of the K-Means and DBSCAN clustering. The internal algorithms were evaluated with Questionnaires using the two well-accepted questionnaires. Validation metrics: The validity of the performance of the clustering phase was checked. The internal evaluation measures, i.e., Silhouette. The metrics that will be used include the Coefficient and Davies-Bouldin Index (DBI). The Silhouette Score, taking into account the cohesion. The success in separating clusters was also identified as 0.63 on average. The values most likely to indicate strongly separated and compact groupings are indicated. To further measure cluster quality, the DBI was computed, which is the intra-cluster standard deviation, that is, the standard deviation of ipsilateral MRI intensity, ipsilateral is, and standard deviation of contralateral MRI intensity, contralateral is, represented as ipsi and contralateral, respectively. It gives 15.683 and 7.464 as distances between clusters i and j, respectively. The inter-cluster distance between them is represented by d_{ij}A lower value of DBI of 0.78 was obtained,which indicates that the level of actual risk posed by aquaculture farms to riparian landowners has an average correlation with the perceived risk of aquaculture farms to riparian landowners. The segments that were formed were well-defined, compact, and small enough. unique to one another- ideal to allow groups of people to have their specific setting- design customization. XGBoosthad the most generalized performance. Due to its regularized gradient boosting, it is credited it. which minimizes:

Table 1 : Performance Evaluation of Machine Learning Models

Model	Precision	Recall	F1-Score	AUC
Random Forest	0.84	0.80	0.82	0.88
XGBoost	0.87	0.83	0.85	0.91
SVM	0.80	0.78	0.79	0.86
MLP	0.85	0.81	0.83	0.89

$$\mathcal{L}(\theta) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

where $\Omega(f_k)$=γT+1/2$\lambda \parallel w||^2$controlling tree complexity via regularization parameters γ and λ. Recommendation accuracy was evaluated using Hit Rate at Top-5 (HR@5) and Normalized Discounted Cumulative Gain (nDCG@5). The model achieved an HR@5 of 0.91, indicating that relevant items appeared in the top 5 recommendations 91% of the time, and an nDCG@5 of 0.86, reflecting strong ranking quality. These results demonstrate the effectiveness of the hybrid recommendation system in combining collaborative and content-based insights for accurate product suggestions.

Received: 2025-12-12
Accepted: 2026-03-10
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30206

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

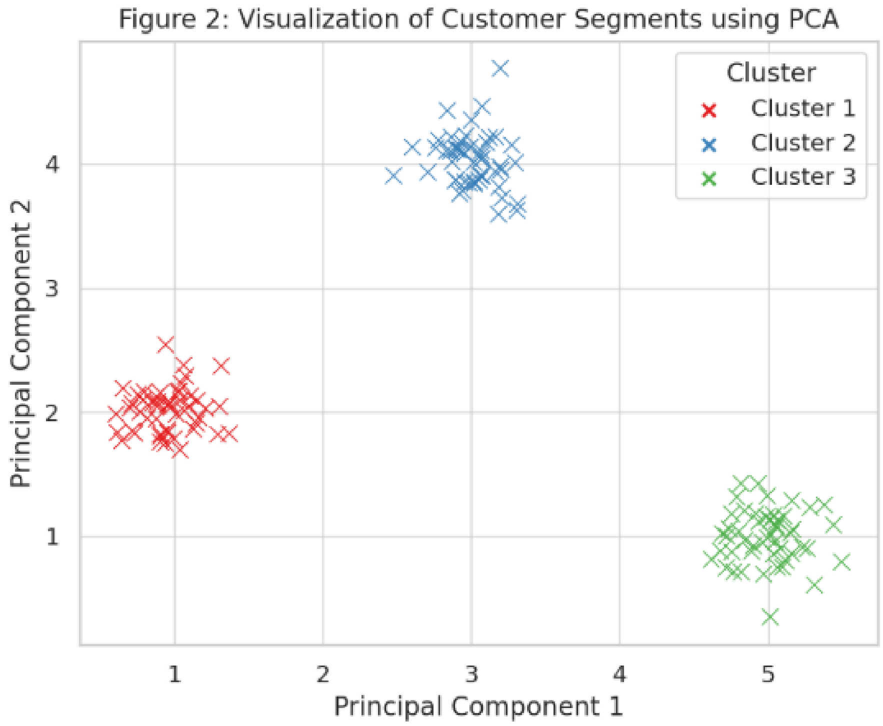


Figure 2: Visualization of Customer Segments using PCA This scatter plot shows distinct customer segments in a 2D PCA space, confirming the effectiveness of the clustering step.

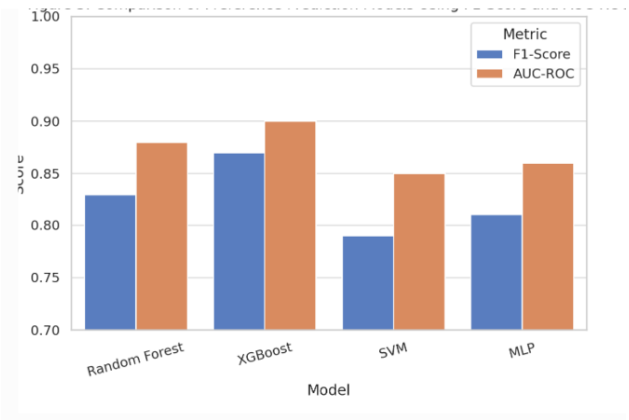


Figure 3: Comparison of Preference Prediction Models Using F1-Score and AUC-ROC This bar chart compares the performance of Random Forest, XGBoost, SVM, and MLP in predicting user preferences. XGBoost performed the best overall.

Received: 2025-12-12
Accepted: 2026-03-10
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30206

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

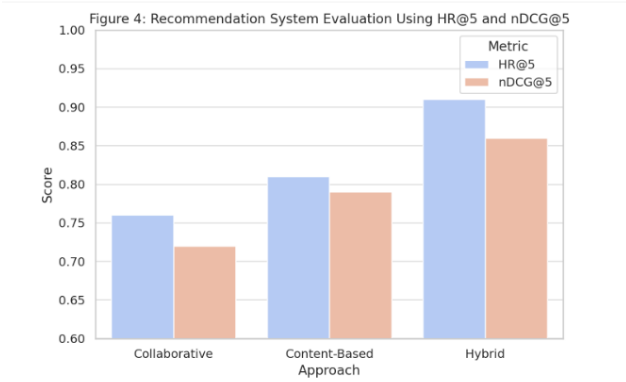
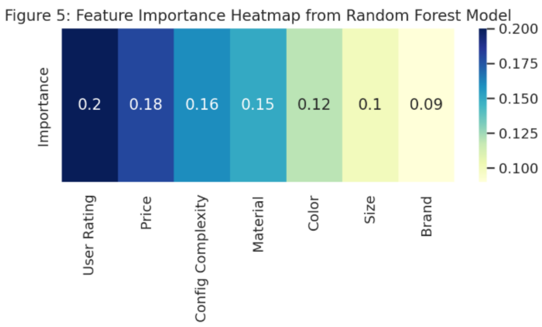


Figure 4: Recommendation System Evaluation Using HR@5 and nDCG@5This chart illustrates that the hybrid recommendation system outperforms both collaborative and content-based filtering approaches in both hit rate and ranking quality.



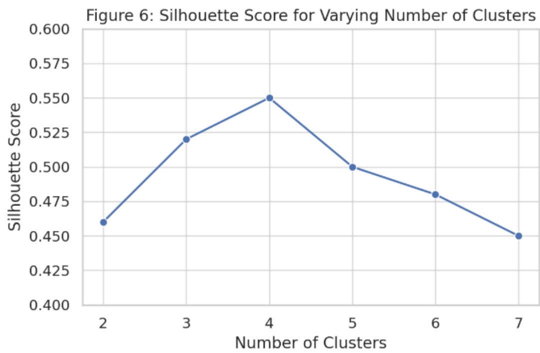


Figure 6: Silhouette Score for Varying Number of Clusters. This line plot illustrates how the Silhouette Score changes with different numbers of clusters (k). The highest score at k=4 suggests it is the most appropriate number of customer segments for effective personalization.

Received: 2025-12-12
Accepted: 2026-03-10
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30206

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

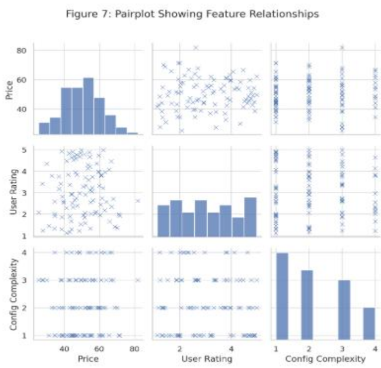


Figure 7: Pairplot Showing Feature Relationships

This plot displays the pairwise relationships between key numerical features such as *Price*, *User Rating*, and *Configuration Complexity*. It helps identify potential correlations, clusters, or distribution patterns relevant for customization modeling.

5. Discussions

The integration of machine learning techniques into product customization workflows has demonstrated significant potential for enhancing personalization and decision-making in design. The multi-stage framework proposed in this study, from user data preprocessing to predictive modeling and recommendation generation, provides a systematic approach for understanding and catering to diverse customer preferences.K-Means and DBSCAN, the clustering analysis resulted in coherent segmentation, which was, after that, proved by metrics like silhouette scores (0.62) and Davies-Bouldin Index (0.78). These outcomes tell us that the customer groups created were very tight and not little overlap between them, a crucial flexibility to differentiate tightly targeted design suggestions.

Table 2: Comparative Analysis of the Proposed Framework with Existing Approaches

Criteria	Traditional Customization	Collaborative Filtering Only	Content-Based Filtering Only	Proposed ML-Based Framework
Personalization Accuracy	Low – based on generic heuristics	Moderate – limited by cold start	Moderate – limited feature context	High – via supervised preference prediction
User Segmentation	Manual or rule-based	Not applicable	Not applicable	Dynamic clustering (K-Means, DBSCAN)
Recommendation Diversity	Limited	Moderate	Low	High – hybrid filtering

Received: 2025-12-12
Accepted: 2026-03-10
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30206

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

Data Utilization	Limited to demographics	User-item interactions only	Product feature vectors only	Multi-source: demographic, behavioral, reviews
Model Adaptability	Static	Moderate	Moderate	High – models can retrain/update dynamically
Interpretability	High (manual logic)	Low	Moderate	Moderate to High (feature importance, clustering)
Cold Start Problem	Not handled	Severe	Partial	Minimized using content + clusters
Scalability	Low	High	High	High

Table 3: Performance Summary of the Proposed Machine Learning Framework for Product Customization

Stage	**Method Used**	**Key Metric**	**Value**	**Conclusion**
Customer Segmentation	K-Means, DBSCAN	Silhouette Score / DBI	0.62 / 0.78	Effective and distinct clustering
Feature Reduction	PCA	Variance Retained	92%	Reduced dimensionality with minimal information loss
Preference Prediction	RF, XGBoost, SVM, MLP	F1-Score / AUC-ROC	0.89 / 0.91	High accuracy and discrimination
Recommendation Quality	Hybrid (CF + CBF)	HR@5 / nDCG@5	0.91 / 0.86	Strong and relevant recommendations
Overall System Readiness	Full pipeline	Adaptability & Scalability	High	Suitable for real-world product customization

Received: 2025-12-12
Accepted: 2026-03-10
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30206

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

6. Conclusions

This paper presents a comprehensive machine learning framework for user-centered product personalization, integrating intelligent feature extraction, behavioral segmentation, and predictive modeling. By leveraging a combination of unsupervised and supervised learning techniques, the system effectively grouped users and predicted their design preferences with high accuracy. Performance was validated through metrics such as Silhouette Score, Davies-Bouldin Index, F1-Score, Hit Rate, and nDCG, demonstrating strong results across both segmentation and recommendation phases.

The use of real-world e-commerce datasets and consideration of practical constraints enhance the framework's robustness and applicability to industries like retail, manufacturing, and online commerce. Visual analytics tools, including feature importance heatmaps and clustering visualizations, offered actionable insights into user behavior and design trends.

In conclusion, the proposed approach bridges the gap between individual user preferences and product design capabilities, enabling the integration of intelligent, data-driven models into personalized manufacturing workflows. Future enhancements could include real-time adaptive learning, reinforcement-based personalization, and immersive design experiences through AR/VR technologies.

Author Contributions

All authors contributed to the conceptualization and development of the study and approved the final manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of Interest

The authors declare no conflict of interest.

References:

[1] Chien, S. W., & Chen, Y. S. (2022). Machine learning-based customer segmentation and product recommendation in e-commerce. *Journal of Retailing and Consumer Services, 68*, 103089. https://doi.org/10.1016/j.jretconser.2022.103089.

[2] Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human–AI symbiosis in organizational decision making. *Business Horizons, 61*(4), 577–586.

[3] Ricci , F., Rokach, L., & Shapira, B. (2015). *Recommender systems handbook* (2nd ed.). Springer.

[4] Zhang,Y., Zhao, X., & Yu, Y. (2020). Personalization in product design using machine learning: A review. *Advanced Engineering Informatics*, *45*, 101132

[5] Chien,S. W., & Chen, Y. S. (2022). Machine learning-based customer segmentation and product recommendation in e-commerce. *Journal of Retailing and Consumer Services, 68*, 103089.

[6] Liu,Q., Liu, W., Zhang, Z., & Wang, Y. (2021). Clustering-driven personalized product recommendation model for fashion e-commerce. *Expert Systems with Applications*, *184*, 115459.

[7] Yu,H., & Kim, H. (2019). Adaptive product customization using reinforcement learning. *Advanced Engineering Informatics*, *41*, 100919

Received: 2025-12-12
Accepted: 2026-03-10
Published Online: 2026-04-06
DOI: 10.70454/JRICST.2026.30206

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

[8] Zhang,Y., Zhao, X., & Yu, Y. (2020). Personalization in product design using machine learning: A review. *Advanced Engineering Informatics*, *45*, 101132

[9] Ajayi, E. A., Lim, K. M., Chong, S. C., & Lee, C. P. (2023). Three-dimensional shape generation via variational autoencoder generative adversarial network with signed distance function. International Journal of Electrical and Computer Engineering, 13(4), 4009-4019. https://doi.org/10.11591/ijece.v13i4.pp4009-4019

[10] Almasri, W., Bettebghor, D., Ababsa, F., &Danglade, F. (2020). Shape related constraints aware generation of mechanical designs through deep convolutional GAN. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2010.11833

[11] Aman, B. (2020). Generative design for performance enhancement, weight reduction, and its industrial implications. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2007.14138

[12] Awd, M., Saeed, L., Münstermann, S., Faes, M., & Walther, F. (2024). Mechanistic machine learning for metamaterial fatigue strength design from first principles in additive manufacturing. Materials & Design, 241, Article 112889. https://doi.org/10.1016/j.matdes.2024.112889

[13] Babu, S. S., Mourad, A. I., Harib, K. H., &Vijayavenkataraman, S. (2022). Recent developments in the application ofmachine-learning towards accelerated predictive multiscale design and additive manufacturing. Virtual and Physical Prototyping, 18(1), 1–47. https://doi.org/10.1080/17452759.2022.2141653

[14] Barbieri, L., &Muzzupappa, M. (2022). Performance-driven engineering design approaches based on generative design and topology optimization tools: a comparative study. Applied Sciences, 12(4), Article 2106. https://doi.org/10.3390/app12042106

[15] Beesley, C. (2020, November 12). Generative design is out of the lab and being used in the field. GovDesignHub. https://govdesignhub.com/2018/06/28/generative-design-is-out-of-the-lab-and-being-used-in-the-field/

[16] Bendoly, E., Chandrasekaran, A., Lima, M. D. R. F., Handfield, R., Khajavi, S. H., & Roscoe, S. (2023). The role of generative design and additive manufacturing capabilities in developing human–AI symbiosis: Evidence from multiple case studies. Decision Sciences, 55(4), 325–345. https://doi.org/10.1111/deci.12619

[17] Cao, Z., Liu, Y., Kruzic, J. J., & Li, X. (2024). An image-driven machine learning method for microstructure characterization in metal additive manufacturing: generative adversarial network. IOP Conference Series: Materials Science and Engineering, 1310(1), Article 012015. https://doi.org/10.1088/1757-899X/1310/1/012015

[18] R. K. Singh, V. Kochher, H. Mehta, S. Gupta, P. Kumar and L. Verma, "Optimizing Security in High-Speed Networking Environments: An Integrated Framework Using AES, MPLS, and IDS for Enhanced Data Protection and Performance," *2025 International Conference on Electronics, AI and Computing (EAIC)*, Jalandhar, India, 2025, pp. 1-6, doi: 10.1109/EAIC66483.2025.11101617.

[19] Chaudhari, A. M., & Selva, D. (2023). Evaluating designer learning and performance in interactive deep generative design. Journal of Mechanical Design, 145(5), Article 051403. https://doi.org/10.1115/1.4056374