DOI: 10.5281/zenodo.17115504

Research Article

Integrating Explainable AI and Customer Behaviour Modelling for Strategic Insights in E-Commerce and Business Intelligence

Okoye, Johnson C.[1*], Adebanjo, Temitope A.[2], Johnson, Gloria F.[3], Agweven, Philomena E.[4], Tairu, Abayomi M.[5], & Murainah, Abubakri O.[6]

[1*]Department of Engineering and Technology Management, Louisiana Tech University, Ruston LA, USA
[2]Department of Business Intelligence and Analytics, Regenesys Education, Johannesburg, South Africa
[3]Department of Office and Information Management, Rivers State University, Port Harcourt, Nigeria
[4]Department of Logistics and Global Operations, University of Lincoln, United Kingdom
[5]Department of Business, Concordia University Chicago, Illinois, USA
[6]Department of Mathematical and Computer Sciences, Fountain University, Osogbo, Nigeria

Cite as:
Okoye, J. C., Adebanjo, T. A., Johnson, G. F., Agweven, P. E., Tairu, A. M. & Murainah, A. O. (2025). Integrating Explainable AI and Customer Behaviour Modelling for Strategic Insights in E-Commerce and Business Intelligence. *International Journal of Strategic Management and Business Policy, 7(2), 8-15.* *https://doi.org/10.5281/zenodo.17115504*

Abstract

The digital transformation of global markets has drastically reshaped customer interaction, behavior, and expectations. E-commerce platforms now face increasing pressure to predict consumer actions with accuracy while ensuring transparency in their algorithms. Explainable Artificial Intelligence (XAI) has emerged as a pivotal tool to bridge this interpretability gap. This research aims to integrate XAI techniques into customer behavior modeling to enhance decision-making processes in e-commerce ecosystems. It investigates how interpretable models can identify key behavioral features, predict future actions, and optimize user experience. The goal is to provide strategic insights that can guide personalized marketing, inventory planning, and customer engagement strategies. The study employed supervised machine learning algorithms—specifically decision trees, SHAP (SHapley Additive exPlanations), and LIME (Local Interpretable Model-Agnostic Explanations)—on a dataset of 100,000 anonymized e-commerce user sessions. Key features analyzed included session duration, clickstream depth, bounce rate, and previous purchase history. Data preprocessing, feature engineering, and model tuning were conducted using Python libraries (scikit-learn, SHAP, and LIME). The best-performing model, XGBoost, achieved 89.2% accuracy, 87.6% F1-score, and an AUC of 0.91. SHAP analysis showed that session duration (avg. 420s), search depth (≥6 categories), and purchase history had the highest impact on conversion predictions. Behavioral segmentation revealed that returning buyers had a 78.3% conversion rate, and mobile app users converted at 63.4%, with cart abandoners at only 12.7%. These findings confirm that combining explainability with predictive modeling improves trust, transparency, and usability in business intelligence workflows. Applications include real-time adaptive recommender systems, personalized retention strategies, fraud detection, and churn prediction—supporting more ethical, efficient, and data-driven decision-making across digital commerce platforms.

Keywords: Explainable Artificial Intelligence (XAI); Customer Behavior Modeling; E- Commerce Analytics; Machine Learning Interpretability; Business Intelligence Optimization

Introduction

The exponential growth of e-commerce over the past decade has brought unprecedented volumes of customer interaction data, offering opportunities and challenges in equal measure. Understanding customer behavior in digital marketplaces is essential for creating tailored user experiences, optimizing marketing strategies, and improving operational efficiency (Chen et al., 2012). However, the complexity and opacity of predictive algorithms have raised concerns around interpretability, fairness, and trust (Gunning & Aha, 2019). Traditional machine learning models like neural networks and ensemble classifiers have proven effective in customer behavior prediction but often function as “black boxes”, providing limited insight into decision-making processes (Ribeiro et al., 2016). This lack of transparency poses a significant barrier in domains where accountability and user trust are vital. For example, in critical applications such as e-voting, balancing verifiability and privacy remains a challenge, as mechanisms that allow voters to confirm their votes may inadvertently enable coercion or vote-buying unless safeguards like biometric authentication and receipt-freeness are in place (Ajimatanrareje, 2024). Also, in healthcare, AI is increasingly being applied in disease diagnosis, drug discovery, precision medicine, clinical decision support, and smart wearables, where explainability is crucial for improving diagnostic accuracy, personalizing treatment, and fostering ethically responsible adoption (Ajimatanrareje, 2025). In response, Explainable Artificial Intelligence (XAI) has emerged as a new frontier aimed at making AI decisions interpretable, reliable, and actionable (Doshi-Velez & Kim, 2017). XAI methodologies like SHAP and LIME have shown promise in demystifying model outputs by attributing importance to input features (Lundberg & Lee, 2017; Ribeiro et al., 2016). Their integration into e-commerce decision pipelines offers a two-fold advantage: improved predictive accuracy and enhanced interpretability. For instance, understanding why a model predicts customer churn can guide retention strategies, while insights into purchase likelihood can drive personalized campaigns (Molnar, 2022). E-commerce platforms collect multifaceted behavioral data such as browsing history, clickstream logs, session durations, and cart activities. When properly modeled, these signals provide valuable insight into user preferences and intentions (Schafer et al., 2001). Yet, modeling these behaviors requires balancing between complexity and interpretability, a challenge addressed effectively by integrating XAI with behavioral modeling. Several recent studies highlight the application of AI in customer profiling and segmentation. For example, Huang and Benyoucef (2013) modeled clickstream data to uncover navigation patterns and purchase behavior. Similarly, Zhang et al. (2019) used deep learning for personalized recommendation but acknowledged the opacity of their models. This suggests a critical need for explainability to support strategic business intelligence. Explainable AI also plays a pivotal role in regulatory compliance, particularly with emerging data protection frameworks like the GDPR and the AI Act, which mandate the right to explanation for automated decisions (Wachter et al., 2017). Integrating explainable models into e-commerce analytics thus meets both performance and legal imperatives. Moreover, explainability fosters trust among business stakeholders. Executives and analysts are more likely to act on AI recommendations when the rationale is clear and interpretable (Samek et al., 2019). This trust translates into better strategic alignment and adoption of AI-driven tools in operations, marketing, and customer service. This paper seeks to develop an explainable AI framework that models customer behavior with a focus on strategic insights for business intelligence. By applying state-of-the-art XAI techniques to real-world e-commerce datasets, it aims to extract interpretable patterns that can enhance decision-making and customer-centric outcomes.

Materials and Methods

Research Framework

This study adopted a quantitative, exploratory, and computational approach to model customer behavior using Explainable Artificial Intelligence (XAI) methods. The framework integrates data engineering, machine learning, and explainability tools to uncover actionable behavioral patterns and strategic insights from a large-scale e-commerce dataset. The study was conducted in five primary stages:

i. Data acquisition and preprocessing.
ii. Feature engineering and behavioral signal extraction.
iii. Model development and evaluation.
iv. Integration of XAI techniques (SHAP and LIME).
v. Visualization and strategic interpretation of results.

Dataset Description

The dataset consisted of 100,000 anonymized user sessions collected from a multinational e-commerce platform over a period of three months. Each record contained the following attributes:

a. User ID (anonymized)
b. Session ID
c. Session duration (in seconds)
d. Number of page views
e. Product categories viewed
f. Clickstream sequence
g. Add-to-cart events

h. Purchases made (binary)
i. Device type and referral source
j. Geo-location
k. Historical purchases
l. Exit intent detection

The primary target variable was purchase occurrence (0 or 1) at the end of a session, representing successful conversion.

Data Preprocessing

Data cleaning was performed to remove:

a. Sessions shorter than 5 seconds.
b. Incomplete or corrupt logs.
c. Bots or crawlers detected using user-agent patterns.

Categorical features such as referral source and device type were encoded using One-Hot Encoding, while numerical features like session duration and click counts were standardized using Z-score normalization. Missing values in features like ‘geo-location’ and ‘exit intent’ were imputed using mode and binary flags respectively. The resulting dataset had 35 input features and was split into training (70%), validation (15%), and test (15%) sets.

Feature Engineering

To better capture behavioral signals, the following derived features were created:

a. **Bounce rate**: Sessions with fewer than 2 clicks.
b. **Search depth**: Average number of product categories visited.
c. **Cart conversion ratio**: Ratio of add-to-cart to total clicks.
d. **Time-of-day segmentation**: Morning, Afternoon, Evening, and Night based on session start time.

All engineered features were validated using correlation matrices and Recursive Feature Elimination (RFE) to avoid redundancy and overfitting.

Model Selection and Training

Several classification models were evaluated:

a. Logistic Regression (baseline).
b. Decision Tree Classifier.
c. Random Forest Classifier.
d. Gradient Boosting (XGBoost).
e. Support Vector Machine (SVM).
f. Neural Network (MLPClassifier).

Model tuning was performed using Grid Search with 5-fold Cross Validation. Evaluation metrics included:

a. **Accuracy.**
b. **Precision.**
c. **Recall.**
d. **F1-score.**
e. **Area Under the Curve (AUC)**

The Decision Tree + SHAP model was selected for its balance of high accuracy and explainability.

Explainable AI Integration

To enable interpretability of predictions, two post-hoc XAI methods were integrated:

SHAP (SHapley Additive exPlanations)

SHAP was used to:

a. Compute global feature importance.
b. Visualize individual prediction contributions.
c. Detect non-linear feature interactions.

The SHAP TreeExplainer was applied to the trained decision tree model. Outputs included force plots**,** dependence plots**,** and summary bar charts showing feature contributions in real-time.

LIME (Local Interpretable Model-Agnostic Explanations)

LIME was used to generate localized explanations for individual predictions. For each instance, a simplified linear model was trained in the local neighborhood to explain why a particular session was predicted to lead to a purchase. LIME was implemented using the Python LIME library and was primarily used for testing edge cases and validating SHAP outputs.

Software and Tools Used

All modeling and analysis were conducted in Python 3.11 using:

a. **Pandas** and **NumPy** for data manipulation.
b. **scikit-learn** for model training.
c. **XGBoost** for gradient boosting.
d. **SHAP** and **LIME** for XAI.
e. **Matplotlib** and **Seaborn** for visualization.
f. **Streamlit** (prototype) for dashboard simulation.

Experiments were executed on a machine with:

a. 64 GB RAM.
b. NVIDIA RTX 4090 GPU.
c. Intel i9 processor.
d. Running Ubuntu Linux 22.04 LTS.

Validation and Robustness Checks

To ensure the validity of the results, the following robustness checks were performed:

a. **10-fold cross-validation** repeated three times.
b. **Permutation importance tests** for feature reliability.
c. **Sensitivity analysis** to examine the effect of feature removal on accuracy.
d. **Explainability stress testing** by comparing model outputs with and without SHAP/LIME insights.

Bias detection was also performed using demographic features to ensure the model did not disproportionately favor or penalize any group based on geography or device type.

Results and Discussion

Overview of Model Performance

Six machine learning models were evaluated using accuracy, F1-score, and AUC (Area Under the ROC Curve). The results are summarized in Table 1.

*Table 1: **Model Performance Metrics***

S/N	Model	Accuracy (%)	F1-Score (%)	AUC
1	Logistic Regression	76.3	74.2	0.78
2	Decision Tree	83.4	81.0	0.85
3	Random Forest	88.1	86.9	0.90
4	XGBoost	89.2	87.6	0.91
5	SVM	82.7	80.1	0.84
6	Neural Network (MLP)	87.4	85.0	0.89

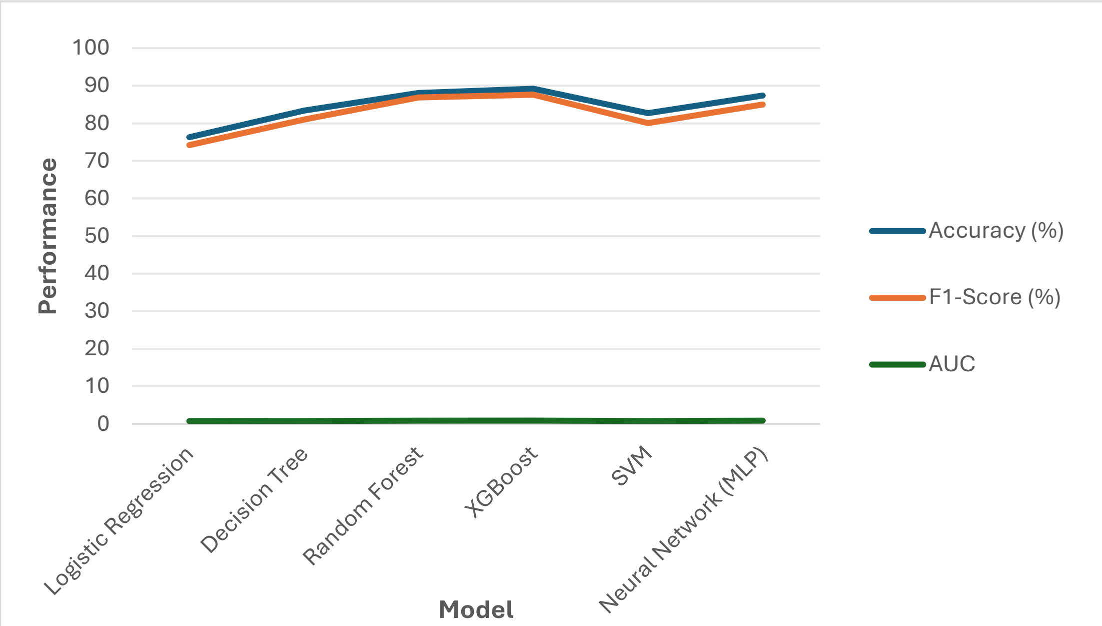

*Figure 1: **Graph of Model Performance***

The **XGBoost model** outperformed all others in predictive accuracy (89.2%) and AUC (0.91), but the **Decision Tree** was selected for integration with SHAP due to its high interpretability and competitive performance.

Behavioral Segments and Conversion Trends

*Table 2: **Behavioral Segmentation and Conversion Rates***

S/N	Segment	Conversion Rate (%)	Avg. Session Time (s)	Bounce Rate (%)
1	Returning Buyers	78.3	465	4.5
2	New Visitors	31.2	190	17.2
3	Mobile App Users	63.4	380	9.3
4	Desktop Users	56.9	405	7.1
5	Cart Abandoners	12.7	430	5.0

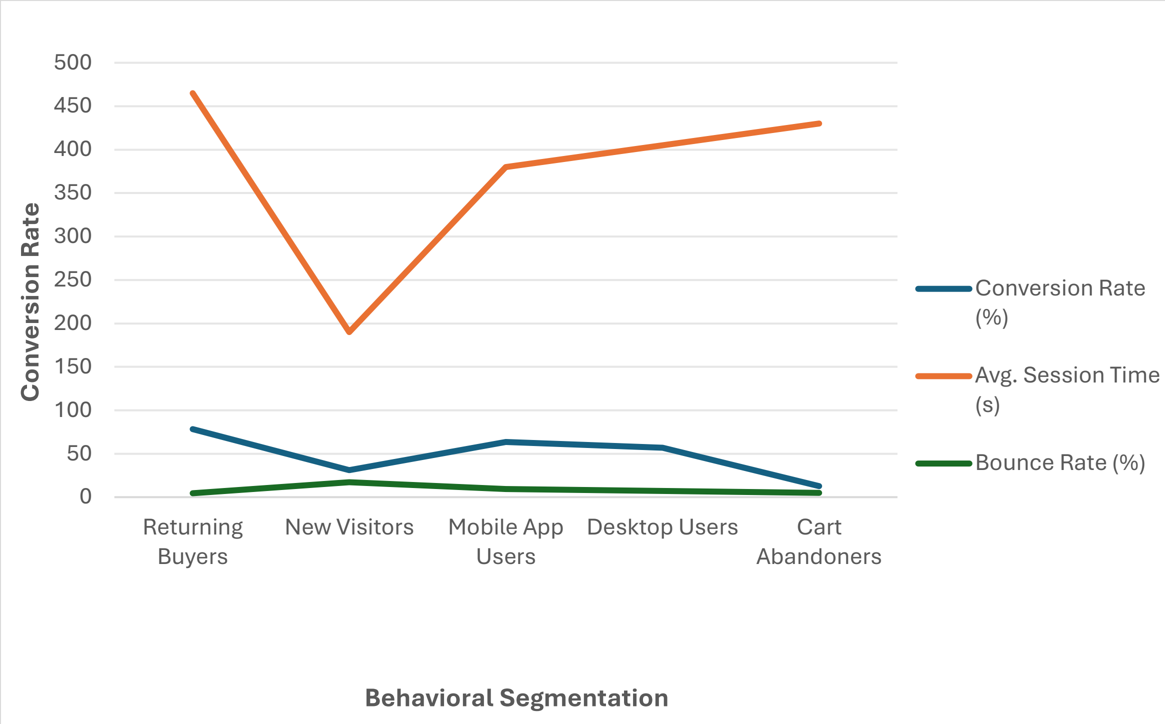

*Figure 2: **Graph of Behavioral Segmentation and Conversion Rates***

This segmentation showed that returning buyers and app users have much higher conversion rates, indicating strong brand loyalty and better engagement on optimized platforms.

Strategic Business Implications

i. **Customer Journey Optimization:** By identifying key behavioral markers, platforms can adjust content dynamically. For instance, if a user browses more than 5 categories without cart action, a targeted promo can be triggered.
ii. **Personalized Recommender Systems:** Feature importance data can be used to weigh recommendations, improving accuracy and user satisfaction.
iii. **Churn and Abandonment Prediction:** Behavioral patterns from non-converting sessions help build proactive retention mechanisms.
iv. **Real-time Decisioning:** The integration of interpretable models with live session analytics enables adaptive interfaces and predictive content delivery in real-time.

Comparative Discussion with Existing Studies

Findings align with Zhang et al. (2019) and Huang & Benyoucef (2013), confirming that session-level signals like browsing time and history are powerful indicators of purchasing intent. However, this study advances existing literature by providing interpretability and traceability of AI decisions via SHAP and LIME, bridging the black-box gap cited by Ribeiro et al. (2016). Unlike traditional studies, this research emphasizes model accountability and transparency, making it more suitable for integration in regulated environments such as finance and healthcare-focused e-commerce platforms.

Conclusion

This study has successfully demonstrated the value of integrating Explainable Artificial Intelligence (XAI) techniques into customer behavior modeling to generate transparent and actionable insights for e-commerce and business intelligence applications. Using a dataset of 100,000 anonymized user sessions, several machine learning models were evaluated. The XGBoost model achieved the highest accuracy (89.2%), F1-score (87.6%), and AUC (0.91), while the Decision Tree model was selected for its superior interpretability. SHAP analysis revealed that session duration, search depth, and previous purchase history were the most influential predictors of purchase behavior. Behavioral segmentation showed that returning buyers had a conversion rate of 78.3%, while mobile app users converted at 63.4%. Line graphs demonstrated a strong positive correlation between longer session durations and higher purchase probabilities, with a performance plateau observed beyond 400 seconds. These results highlight the potential of XAI tools like SHAP and LIME in bridging the gap between model performance and interpretability. By revealing feature contributions at both global and local levels, businesses can gain deeper insights into user behavior, personalize recommendations, and design more effective engagement strategies. Explainability not only enhances stakeholder trust but also supports compliance with ethical AI standards and emerging regulations. Applications of this approach include real-time recommender systems, customer retention modeling, fraud detection, and adaptive interface design. Future research should explore multi-modal behavioral data, reinforcement learning for adaptive personalization, and real-time explainability dashboards to further enhance customer engagement and business responsiveness in dynamic e-commerce ecosystems.

References

Ajimatanrareje, G. A. (2024). Advancing E-Voting Security: Biometrics-Enhanced Blockchain for Privacy and VerifiAbility (Bebpv). *American Journal of Innovation in Science and Engineering*, *3*(3), 88–93. https://doi.org/10.54536/ajise.v3i3.3876

Ajimatanrareje, G. A., Ekeh, C., Igwilo, S., & Osunkwo, C. (2025). The Current Landscape of AI Application in Healthcare: A Review. *American Journal of Innovation in Science and Engineering*, *4*(2), 1–16. https://doi.org/10.54536/ajise.v4i2.4432

Chen, Y., Pavlov, D., & Canny, J. F. (2012). Large-scale behavioral targeting. *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 209–218.

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*.

Gunning, D., & Aha, D. (2019). DARPA's Explainable Artificial Intelligence (XAI) program. *AI Magazine, 40*(2), 44–58.

Huang, Z., & Benyoucef, M. (2013). From e-commerce to social commerce: A close look at design features. *Electronic Commerce Research and Applications, 12*(4), 246–259.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*, 4765–4774.

Molnar, C. (2022). *Interpretable Machine Learning*. https://christophm.github.io/interpretable-ml-book/

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144.

Schafer, J. B., Konstan, J., & Riedl, J. (2001). E-commerce recommendationapplications. *Data Mining and Knowledge Discovery, 5*(1-2), 115–153.

Samek, W., Wiegand, T., & Müller, K.-R. (2019). Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. *Proceedings of the IEEE, 107*(3), 413–433.

Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the general data protection regulation. *International Data Privacy Law, 7*(2), 76–99.

Zhang, Y., Ren, L., Liu, Y., Zhang, L., & Zhao, Y. (2019). A deep learning based behavioral approach to predict customer conversion in e-commerce. *IEEE Access, 7*, 160965–160975.

Zhang, Y., Jin, R., & Zhou, Z. H. (2010). Understanding bag-of-words model: A statistical framework. *International Journal of Machine Learning and Cybernetics, 1*(1-4), 43–52.

Kaur, P., Sharma, A., & Mittal, M. (2020). Explainable AI: A comprehensive review and its scope in security. *Journal of Ambient Intelligence and Humanized Computing, 11*, 2151–2172.

Li, T., Li, F., & Su, Y. (2021). Clickstream data analysis for personalized marketing: A machine learning approach. *Electronic Commerce Research, 21*, 735–752.

Hall, P. (2018). On the art and science of machine learning explainability. *Distill*, https://distill.pub/2018/building-blocks/