SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Modeling Algorithmic Influence On User Behavior In Digital Environments

Anitha M[1], Nivetha N[2], Vikram V. Patel[3], Sanjay Bhatnagar[4], Nilesh Anute[5], Shailesh Tripathi[6]

[1] Assistant Professor, Department of Mathematics, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, India. Email: anitham@maher.ac.in
[2] Assistant Professor, Computer Science, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, India. Email: nivethan@maher.ac.in
[3] Associate Professor, Faculty of Engineering, Gokul Global University, Sidhpur, Gujarat, India. Email: vikrampatel@gokuluniversity.ac.in, ORCID: 0009-0006-4834-6766
[4] Centre for Research Impact & Outcome, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab 140401, India. Email: sanjay.bhatnagar.orp@chitkara.edu.in, ORCID: 0009-0004-7474-1511
[5] Associate Professor, Balaji Institute of Management & Human Resource Development, Sri Balaji University, Pune, Maharashtra, India. Email: nileshanute@gmail.com, ORCID: 0000-0001-6599-813X
[6] Professor, Balaji Institute of Management and Human Resource Development, Sri Balaji University, Pune, Maharashtra, India. Email: motivationtripathi@gmail.com, ORCID: 0009-0001-3009-0460

Abstract

Algorithms have a considerable impact on users' decision making process on online e-commerce websites as recommendation systems affect their buying decisions beyond personal preferences. Traditional methods of Deep Learning (DL) and Machine Learning (ML) are mainly concerned with modeling click-through rate prediction and purchase prediction. However, these methods cannot take into account the behavioral changes due to continuous recommendation exposure. To overcome this problem, an algorithm named the Influence-Aware Proximal Policy Self-Attention Memory Network (IA-PP-SM Net) has been developed for understanding the dynamic relationship between users and recommendation algorithm in online digital commerce context. The IA-PP-SM Net comprises of Proximal Policy Optimization (PPO) with attention-enhanced Long Short-Term Memory (LSTM). The first one is used for learning the sequence of user behavior and the latter one learns recommendation policies. In addition, the algorithm uses an influence-aware reward system to estimate the behavioral change after exposure to recommendations. Moreover, self-attention is introduced to recognize influential past interactions that can affect future user behavior. Min-max normalization was used for data preprocessing for the Kaggle dataset, and Principal Component Analysis (PCA) was utilized to perform characteristic retrieval combined with simplification of data complexity. The proposed approach was modeled through the use of Python programming language utilizing the help of TensorFlow and Keras packages. Experimental outcomes revealed that IA-PP-SM Net model achieves higher performance than other models as it is able to attain a higher accuracy rate of 0.962. This model is able to efficiently identify the effect of dynamic algorithm on recommendation systems for online commerce.

Keywords: Proximal Policy Optimization (PPO), User Behavior Prediction, Long Short-Term Memory (LSTM), Decision Making, E-commerce.

1. Introduction

Emerging research places particular prioritize increasing importance of Artificial Intelligence (AI) in influence branding and purchaser behavior. Leading Electronic-Commerce (e-commerce) websites such as Amazon have been incorporating AI technology not only to optimize their promotional campaigns and sales performance but also to revolutionize the way business is conducted online [1]. AI technology greatly improves competitiveness and the experience of customers for both consumers and organizations. In e-commerce, product selection plays an important role in determining consumer behavior. There are recent trends that investigate the combination of AI and fashion design by using image-generating systems [2]. Post-pandemic there have been significant modifications related to behavior of users, who prefer online purchases because of convenience, access to the Internet, and globalization. Even consumers who are not technologically adept are opting for online purchasing owing to their convenience and payment options [3]. User Behavior is an essential element that affects the functioning of brands through interactions via mobile devices and social networks. E-word-of-mouth (E-WOM) and reviews posted on the Internet contribute as a crucial function in improving trust and purchasing decision-making, usually more than any other marketing effort [4]. The recommendation systems in the online world provide a means to minimize information overload by providing personalized options. Besides technical purposes, they can also affect how people interpret the outcomes of their algorithms [5]. Time and personalized recommendation systems improve making decision based on data in online retailing. These recommendation systems are applied on popular sites such as Amazon and Taobao to minimize the costs involved in searching for information by linking recommendations with changing customer tastes [6].

1.1 Research Aim: This research proposes the development of an Influence-Aware Proximal Policy Self-Attention Memory Network (IA-PP-SM Net) to model user–algorithm interactions in e-commerce systems. The proposed model consists of PPO with attention-enhanced LSTM network for detecting behavioral changes induced by algorithms.

1.2 Research Organization: Section 1 discusses problem statement and the purpose, Section 2 provides the literature review on recommendation systems for e-commerce, Section 3 details the proposed IA-PP-SM Net approach, including data preprocessing and model architecture design, Section 4 gives the experiments and performance evaluation, Section 5 ends with the conclusion of the present model.

2. Literature Review

The research [7] was held to improve e-commerce personalization through the implementation with Model-Driven Development (MDD) and Generative AI. It aimed at improved user satisfaction by 2.3%, 18.6%, and 11.8% for the selected parameters, but suffers from data dependence and scalability issues. Deep Learning (DL), Isolation Forest, and Natural Language Processing (NLP) methods were used to find anomalies in e-commerce data, focusing on improving the platform's security and fraud detection. The model attained an accuracy of 0.83; however, it suffers from several limitations, including dependence on the dataset used [8].The research involved using Spider Monkey Optimized-Deep Neural Networks (SMO-DNN)[9] for predicting customer behavior based on consumer data, clickstream data, and sentiment analysis data. The model showed an improved accuracy; however, it has limitations related to high computational cost and dependence on data.The research [10] investigated the recommendation systems by incorporating e-loyalty variables like purchasing frequency and interaction through various models and datasets for enhanced customization. This method yielded a hit-rate@20 of 0.018 using ds1 with g4 variables. Nonetheless, it is constrained by its reliance on data sets and variable choices.Machine Learning (ML) classification models and Knuth-Morris-Pratt (KMP)-based multivariate pruning methods [11] were conducted using e-commerce transactions for the analysis of website usability and user behavior. The accuracy obtained for the developed models was 94.2%. However, the method suffers from poor interpretability and implementation issues.The research in [12] developed a fake product detection model using image analysis employing Convolutional Neural Network (CNN), integrated via transfer learning and data augmentation techniques to improve classification. The research generated recall 95% and accuracy 90%, demonstrating its effectiveness in identifying counterfeit products. The research utilized linear regression modeling through Diverse Counterfactual Explanations (DiCE) for optimizing the rate structure of mobile phones based on relevant product attributes affecting price. The methodology yields an R^2 score of 96.15%. Nonetheless, its limitations include linear assumptions and dependency on features [13].Ordered Clustering Based Algorithm (OCA) [14] was suggested to enhance the e-commerce recommender system by tackling the obstacles of initial beginning and records limited availability through user clustering based on collaborative filtering. This technique performs better than Hierarchical Clustering Analysis (HCA), where precision, recall, and F-measure scores are better at 10.24%, 14.55%, and 11.23% respectively. The research aimed at analyzing clickstream data through various machine learning (ML) techniques such as Decision Tree (DT), Random Forest (RF), Support Vector Classifier (SVC), and Logistic Regression (LR) to measure the impact of time spent reviewing product details on purchasing activity. Out of the mentioned models, the SVC performed better with 0.9021 test accuracy. But despite its effectiveness, there were certain shortcomings in the model concerning scalability, dependency on features, and capturing the complex behavioral characteristics of users [15]. The research [16] applied techniques of feature engineering and ensembling to forecast purchase intentions in e-commerce. The RF classifier gave an accuracy 0.892 and Area under the Receiver Operating Characteristic curve (AUC-ROC) of 0.934, while Extreme Gradient Boosting (XGBoost) had overall AUC-ROC of 0.952, accuracy of 0.908, and Light Gradient Boosting Machine (LightGBM).

3. Methodology

The model illustrated in Figure 1 presents the architecture of an end-to-end user behavior prediction system for e-commerce applications. The process begins with data loading, followed by data preprocessing using Min–Max normalize the input data. Subsequently, behavioral, aggregational, and temporal features are extracted, and dimensionality reduction is performed using PCA to improve computational efficiency and reduce feature redundancy. During the prediction phase, the attention based LSTM algorithm is employed to identify sequential user behavior patterns, while the optimization process is enhanced through Reinforcement Learning (RL) using the PPO algorithm, where the reward signal serves as an input for adaptive learning and policy improvement.

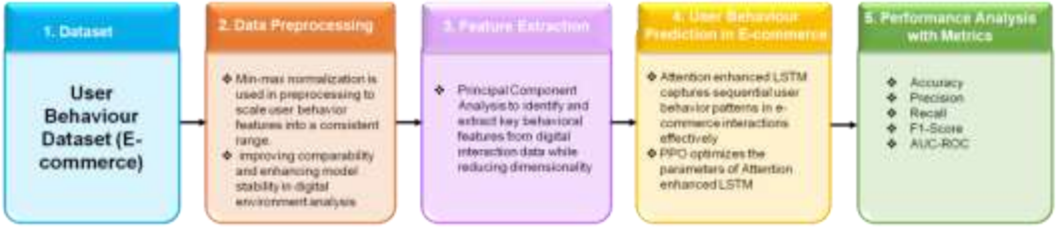


Figure 1: Method flow of developed approach

3.1 Data Collection

Dataset obtained through the Kaggle source (https://www.kaggle.com/datasets/marwa80/userbehavior) is employed in this research to predict the user behavior in e-commerce, where the dataset contains sequential interactions like click, view, and purchase actions. The overall count of records inside the dataset represents approximately 1 million, which makes it adequate for DL experiments. Dataset is organized into 80:20 partition, where 80% is used for training, while 20% is utilized for testing.

3.2 Min-Max Normalization-Based Preprocessing

The Min-Max Scaling is utilized to transform the diverse e-commerce engagement attributes, including clicks, transactions, recommendations shown, and metadata, so that all of the attributes have an equal influence on the modeling process, and no single attribute with large values can overshadow the others. The purpose of the transformation is to map all of the attributes into a common range, usually [0,1].

$$v' = \frac{v-\min(\max_d-\min_d)}{\max-\min}+\min_d \quad (1)$$

In Equation (1), v' refers to the normalized value of the interaction feature,$v$denotes the original value of the user behavior, $\max_d$and $\min_d$ denote themaximum value andminimum value of the feature d, respectively; min and maxare the minimum and maximum values.

3.3 Dimensionality Reduction by Principal Component Analysis (PCA)

To facilitate robust attribute learning in modeling interactions in e-commerce settings, the PCA technique is used after preprocessing data that includes variables such as click rate, purchase rate, recommendation, and product information. The rationale behind this approach is that the features exist on different scales; therefore, they must be standardized to ensure that each variable contributes equally to the model and does not disproportionately influence the learning process.

$$K_{lm} = \frac{(Y_{lm}-\mu_{lm})}{\sigma_{lm}} \quad (2)$$

In Equation (2),K_{lm} denotes the feature value of the standardized interaction, Y_{lm} refers to the feature value of the original behavior,$l$represents the index value of the user interaction, m denotes the identification feature,μ_{lm} refers to the average feature value of the index$l$and feature of identification$m$, σ_{lm} denotes the feature of the standard deviation of the index $l$and feature of identification m, Y represents the original or raw feature value before the standardization process, σ denotes the standard deviation, μ denotes the mean and the K refers to the value of standardization.

3.4 User behavior Prediction in E- commerce platform using Influence-Aware Proximal Policy Self-Attention Memory Network (IA-PP-SM Net) model

The proposed IA-PP-SM Net model combines an encoder–Long Short-Term Memory (LSTM)–attention–decoder architecture with PPO based Reinforcement Learning to predict user behavior and optimize recommendations in e-commerce platforms. The encoder-LSTM captures temporal dependencies in user activities such as clicks, purchases, and browsing patterns. The attention mechanism identifies influential past interactions to improve prediction accuracy, while the decoder predicts future user actions at different time steps. Furthermore, PPO reinforcement learning continuously updates the recommendation policy by balancing exploration and exploitation to deliver adaptive and personalized recommendations.

3.4.1 User behavior Prediction using Self-Attention Memory Network: The proposed e-commerce user behavior prediction model uses an encoder–LSTM–attention–decoder model to process sequential user interaction information like clicks, purchases, recommendation views, and item attributes. LSTM can learn the patterns in the temporal information of user activities and capture their temporal dependencies. Attention is used on top of LSTM hidden states to pinpoint those past events that affect future user decisions. Finally, the decoder leverages the LSTM final hidden state and attention context to make multi-step forecasts of user engagement with one forward pass through the network. Such a model allows for robust predictions of short- and long-term behavioral changes in digital commerce settings.

$$\alpha_j = \mathrm{softmax}_j\left(\frac{r^{S} l_j}{\sqrt{c}}\right) \qquad (3)$$

$$d = \sum_{j=1}^{S_{in}} \alpha_j u_j \qquad (4)$$

In Equation (3& 4), $\alpha_j$refers weight score of the attention, softmaxtransforms raw scores to probabilities,lkey representation at time step, j refers to the index of input elements,l_j denotes the compressed representation, crefers to the scaling factor of the dimensionality of the feature and $\sqrt{c}$ denotes the factor of scaling.rvector representation of the query, S denotes step of the current sequence,udenotes feature of input, u_j value feature vector at the time step, ddenotes the vector representation of the context, S_{in} input sequence length steps,$\sum_{j=1}^{S_{in}}$ combines the value vectors through weighting into one overall context representation incorporating patterns of influential user behaviour.

$$\mathrm{vec}(\hat{Z}) = X_p[yS_{in}\|d] + a_{p,}\hat{Z} \in \mathbb{Q}^{S_{out}\times 2} (5)$$

In Equation (5),$\hat{Z}$is the output matrixof prediction, vec refers to the vectorization, X_p refers to the projection matrixof learning, a_p is a termofbias, $\mathbb{Q}$refers to the set ofrational numbers, $\hat{Z} \in \mathbb{Q}^{S_{out}\times 2}$ refers to the output 2D sequence, and **y**target of theinput.

$$\varkappa = \lambda_1.\frac{1}{S_{out}}\sum_{i=1}^{S_{out}}\left\|\hat{T}_{s+i} - T_{s+i}\right\|_2^2 + \lambda_2.\left|\hat{s}_{stop} - s_{stop}\right|\lambda_3.\frac{1}{S_{out-1}}\sum_{i=1}^{S_{out}-1}\left\|\hat{T}_{s+i+1} - \hat{T}_{s+i}\right\|_2^2 (6)$$

In Equation (6), $\varkappa$denotes the loss value of the overall training,λ_1, λ_2, and$\lambda_3$denotes the weight factor ofsequence error, S_{out} refers to the length steps of the prediction horizon, T_{s+i}refers to the interaction state of the true future, $\hat{T}_{s+i}$denotes the state value of the predicted interaction, $s_{stop}$is the time index of the actual stopping, $\hat{s}_{stop}$denotes the time index of the predicted stopping, $\sum_{i=1}^{S_{out}}$ and $\sum_{i=1}^{S_{out}-1}$ considers the sum of all present and previous prediction errors to calculate overall trajectory error,$\left|\hat{s}_{stop} - s_{stop}\right|$denotes difference ofabsolutestopand $\left\|\hat{T}_{s+i+1} - \hat{T}_{s+i}\right\|_2^2$ refers to the magnitude of squared difference.Srepresents the current time step or present interaction index, idenotes the iteration index used in the summation process. $\hat{T}$represents the predicted interaction state generated by the model. $\| \quad \|_2^2$denotes the squared Euclidean (L2) norm used to measure prediction error magnitude.

$$\hat{j}_{stop} = \min\{i \in [2, S_{out}] | c_i < \varepsilon_{stop}\}, \varepsilon_{stop} = 0.01n \quad (7)$$

In Equation (7), $\hat{j}_{stop}$ denotes the step index of the predicted stopping, c_i denotes the magnitude value of displacement, ε_{stop} denotes the tolerance value of the stopping threshold, min selects the time step of the minimum value,i denotes the index of time step, $i \in [2, S_{out}]$refers to the range sequence index, $c_i < \varepsilon_{stop}$satisfied stop criterion andn is a parameterof scaling.$\hat{j}$ represents the predicted stopping-step index, $\in$ denotes that the index belongs to the specified range.

3.4.2 Hyperparameter tuning with Proximal Policy Optimization (PPO): PPO is adopted to optimize the recommendation policy in e-commerce platforms, where the decisions of users are affected by algorithmic exposure that is not explicitly preferred by the users. The learning process is driven by interaction feedback, such as clicks, purchases, and exposures that help capturing the immediate reactions and behavioral changes in the long term. PPO enhances the stability of the training process through clipping the objective between the current and the previous policy. The advantage function is employed to evaluate the actions for balanced exploration and exploitation.

$$K^{CLIP}(\theta) = \hat{F}_s\left[min\left(q_s(\theta).\hat{B}_s, clip(q_s(\theta), 1-\varepsilon, 1+\varepsilon).\hat{B}_s\right)\right] \quad (8)$$

In Equation (8), $K^{CLIP}(\theta)$ refers to the loss function of clipped policy, θ denotes the vector set of the policy parameters,$\hat{F}_s$denotes the value that are expected over time, $q_s(\theta)$ refers to the ratio of probability in policy**,** $\hat{B}_s$denotes the value of estimated function**,** ε refers to the parameter value of the clipping threshold, min represents the operator function of the minimum selection, $clip$denotes the ratio of bounds within the limits, K denotes the optimization of the clipped policy, F denotes the operator computed expectation, and$B$ denotes the value of the estimated advantage.

$$q_s(\theta) = \frac{\pi_\theta\,(\alpha_s|t_s)}{\pi_{\theta_{old}}(\alpha_s|t_s)} \quad (9)$$

In Equation (9), θ_{old} refers to the parameters of the prior policy,$\pi_\theta\,(\alpha_s|t_s)$refers to the action-taking probability, $\pi_{\theta_{old}}(\alpha_s|t_s)$ refers to the action-taking probability of the same action based on the old policy, $t_s$denotes the time step of the state, andα_s denotes theactionpolicy taken at the step of time.

$$\hat{B}_s = Q_s - U(t_s) \quad (10)$$

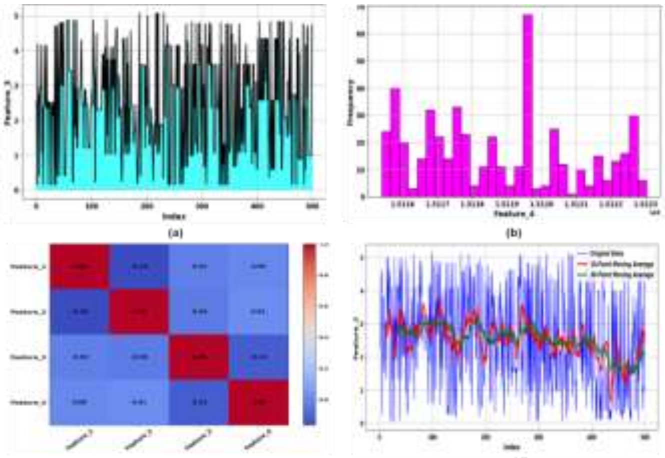
In Equation (10),Q_s refers to the overall reward received from the time step, $U(t_s)$ denotes the function of State Value, Q stands for the good actions that are done in the state, and Ustands for the quality of the state.

3.4.3 Hyperparameter of the proposed model: The IA-PP-SM Net model under consideration involves the application of reinforcement learning based on PPO, self-attention, and a sequential memory mechanism for adaptive behavioral prediction of users. It uses the following parameters: discount factor is equal to 0.99, the learning rate is set to (3 \times 10^{-4}), and PPO clip value is 0.2. Model efficiency is enhanced by multi-head self-attention mechanism, 128-dimensional embedding vectors, usage of Gated Recurrent Unit (GRU)/LSTM memory units, drop-out method, and Gaussian Error Linear Unit (GELU) or the Rectified Linear Unit (ReLU) functions.

4. Results

Result outcomes from experiments supports applicability of the suggested IA-PP-SM Net model in effectively modeling dynamic user behavior and achieving improved recommendations via adaptive learning, attention-based sequential modeling, and PPO-based optimization. The model is developed in Python 3.9/3.10 environment, which uses PyTorch 2.0+, Stable-Baselines3 PPO, NumPy, SciPy, Pandas, and Matplotlib libraries. The experiments were performed on Intel Core i7/AMD Ryzen 7 processors, having NVIDIA RTX 3060/3080 GPU, 16-32 GB RAM, SSD drive, CUDA 11.8/cuDNN, Jupyter notebook, and Visual Studio Code environment.

4.1 Explorative Data Analysis: Figure 2 provides an overall picture of feature analysis on the user behavior prediction model. Figure 2(a) displays how feature values vary with respect to various indices, emphasizing the presence of volatility and dynamic behaviors in the dataset. Figure 2(b) presents the distribution of feature values, suggesting the extent to which samples exist within specific ranges. Figure 2(c) explains the correlation matrix between various features, showing whether the relationship is positive, negative, or weak, and hence affecting the learning process and the prediction of user behavior. Figure 2(d) presents the smoothed trend lines through the moving average methods with a range of 10 points and 30 points.



(a) (b)

(c) (d)

Figure 2: Visualization of (a) Variation of Feature Trend, (b) Analysis of Feature Distribution, (c) correlation matrix, (d) Trends of Moving Average

4.2 Evaluation Metrics: Classification performance measurement standards namely F1-score, AUC-ROC, accuracy, recall, and precision are designed to validate suggested approach aimed at predicting user behavior across e-commerce portals. Accuracy represents the final count of correct detections created by the algorithm, while precision assesses how many of the predicted positives are indeed relevant. Recall represents the capability of the algorithm in identifying all the true samples. F1-score equalizes the recall and precision measures for robust performance evaluation on imbalanced interaction data. AUC-ROC metric indicates how effectively the recommendation influence is distinguished by the model.

4.3 Performance analysis: Table 1 and Figure 3 illustrate the performances of different types of machine learning techniques regarding multiple quantitative measures that include Precision, F1-Score, AUC-ROC, Recall, and Accuracy. The traditional models, especially Naive Bayes (NB) [16], DT [16], along with RF [16], exhibit average performances according to efficiency and accuracy. SVM [16] and logistic regression [16] have balanced results but take more time in training. Advanced algorithms such as XGBoost [16], LightGBM [16], and NN [16] offer a high degree of precision, but they are still outmatched by the proposed IA-PP-SM Net that achieved maximum values of Precision (0.964), F1-Score (0.962), AUC-ROC (0.978), Accuracy (0.962), and Recall (0.960). The major advantage of this technique comes from its ability to utilize self-attention, sequential memory learning, and PPO optimization.

Table 1: Performance Evaluation of Baseline Approaches with Implemented Method

Algorithm	*F1 − Score*	*Recall*	*Precision*	*AUC − ROC*	*Accuracy*
NB [16]	0.757	0.781	0.734	0.841	0.763
DT [16]	0.837	0.851	0.823	0.879	0.847
RF [16]	0.879	0.883	0.876	0.934	0.892
SVM (*Linear*) [16]	0.828	0.839	0.817	0.867	0.829
SVM (*RBF*) [16]	0.858	0.868	0.849	0.901	0.861
LR [16]	0.828	0.845	0.812	0.876	0.834
XGBoost [16]	0.898	0.901	0.895	0.952	0.908
LightGBM [16]	0.893	0.897	0.889	0.948	0.903
NN [16]	0.885	0.887	0.883	0.941	0.896
IA − PP − SM Net [*Proposed*]	0.962	0.960	0.964	0.978	0.962

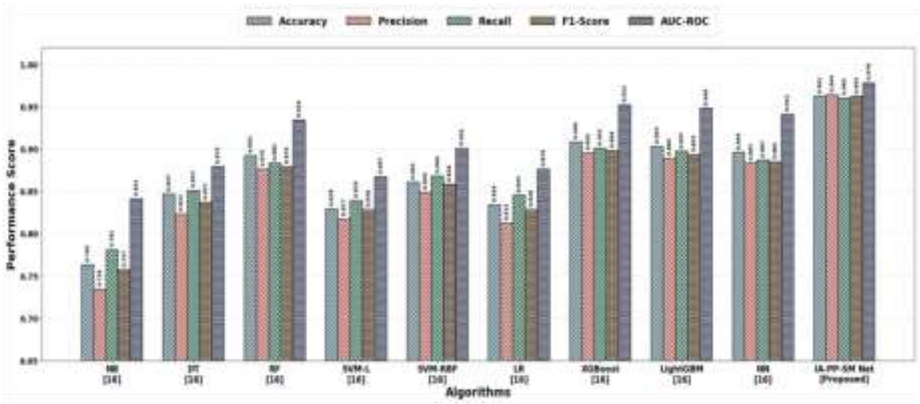


Figure 3: Performance evaluation of Existing and Proposed Models

All existing models have been re-trained based on the Kaggle User Behavior Dataset with identical experimental conditions to perform a fair comparison as presented in Table 2. Similarly, all existing models were analyzed using identical procedures for data pre-processing, i.e., min-max normalization, feature extraction using PCA approach, and evaluation measures. The results of comparative analysis show that IA-PP-SM Net is superior to the other models when it comes to classification, behavior prediction, interaction learning, resource allocation, and intelligent adaptation.

Table 2: Retrained Analysis of Existing and Proposed Models

$Algorithm$	$F1-Score$	$Recall$	$Precision$	$AUC-ROC$	$Accuracy$
NB	0.772	0.793	0.752	0.852	0.781
DT	0.848	0.861	0.836	0.887	0.854
RF	0.888	0.892	0.884	0.941	0.901
$SVM\ (Linear)$	0.836	0.847	0.826	0.874	0.838
$SVM\ (RBF)$	0.867	0.876	0.858	0.913	0.873
LR	0.836	0.851	0.821	0.883	0.842
$XGBoost$	0.906	0.909	0.903	0.958	0.917
$LightGBM$	0.900	0.904	0.897	0.953	0.911
NN	0.893	0.896	0.891	0.947	0.904
$IA-PP-SM\ Net$	0.962	0.960	0.964	0.978	0.962

4.4 Discussion

Classic ML and DL algorithms provide good classification results; yet, their drawbacks include scalability problems, computation efficiency issues, feature dependence, and adaptive learning ability. The NB algorithm [16] enables quick classification based on probabilities yet fails to recognize feature dependences and non-linear feature interrelations. DTs [16] allow interpretability of predictions yet are vulnerable to overfitting and instability. The use of ensemble learning in RFs [16] leads to better prediction quality yet increases memory usage and time needed for training. SVM classifiers based on linear and RBF kernels [16] ensure high precision of classification yet need more computation power when handling big data sets. Both XGBoost [16] and LightGBM [16] improve prediction performance; however, they suffer from sensitivity to hyperparameter setting and problems related to imbalanced features. NN [16], on the other hand, is capable of capturing complicated behavioral patterns but needs considerable amounts of computation and memory when training. On the contrary, the proposed IA-PP-SM Net addresses such weaknesses by offering better adaptability, efficient context learning, resource optimization, enhanced decision intelligence, and generalized capability in predicting user behavior.

5. Conclusion

IA-PP-SM Net models consider the impact of the behavioral pattern by means of attention enhanced LSTM and PPO to capture the sequential behaviors and significant past behaviors along with an adaptive recommendation technique. The model experiments conducted using the Kaggle User Behavior dataset showed good results maintaining an F1-score 0.962, accuracy score of 0.962, recall 0.960, precision 0.964, similarly AUC-ROC 0.978. This makes it better than other conventional ML/DL models in predicting behaviors and adaptively recommending products to users. Limitations of this model include dependency on historical interaction data and computational expenses during RL optimization process. Techniques like Explainable AI, Multimodal Learning, and Instant Recommendation could be developed in future research.

Reference

1. Shabankareh, M., KalantariDaronkola, H., Sarhadi, A., Thaichon, P. and Nazarian, A., 2025. The influence of artificial intelligence quality on business-to-business relationships on E-commerce platforms: unveiling the nexus. International Journal of Advertising, 44(8), pp.1483-1522.https://doi.org/10.1080/02650487.2025.2557053
2. Dumiter, F.C. and Schebesch, K.B., 2025. Using Artificial Intelligence to Determine the Impact of E-Commerce on the Digital Economy. Journal of Theoretical and Applied Electronic Commerce Research, 20(3), p.219.https://doi.org/10.3390/jtaer20030219
3. Baki, R., Ecer, B. and Aktas, A., 2025. A decision framework for supplier selection in digital supply chains of E-commerce platforms using interval-valued intuitionistic fuzzy VIKOR methodology. Journal of Theoretical and Applied Electronic Commerce Research, 20(1), p.23.
4. Figueiredo, N., Ferreira, B.M., Abrantes, J.L. and Martinez, L.F., 2025. The role of digital marketing in online shopping: A bibliometric analysis for decoding consumer behavior. Journal of Theoretical and Applied Electronic Commerce Research, 20(1), p.25.https://doi.org/10.3390/jtaer20010025
5. Choi, J., Kang, S., Moon, J., Jeon, S. and Lim, S., 2026. Algorithmic Transparency and Consumer Trade-Offs in AI-Based Financial E-Commerce Services. Journal of Theoretical and Applied Electronic Commerce Research, 21(3), p.86.https://doi.org/10.3390/jtaer21030086
6. Jiang, S., Chen, T., Tan, Y., Gao, S. and Li, L., 2025. Time-Series Recommendation Quality, Algorithm Aversion, and Data-Driven Decisions: A Temporal Human–AI Interaction Perspective. Mathematics, 13(21), p.3528.https://doi.org/10.3390/math13213528 \
7. Alti, A. and Lakehal, A., 2025. AI-MDD-UX: Revolutionizing E-Commerce User Experience with Generative AI and Model-Driven Development. Future Internet, 17(4), p.180.https://doi.org/10.3390/fi17040180
8. Mah, P.M., Skalna, I. and Pelech-Pilichowski, T., 2025. AI-driven anomaly detection in E-commerce services: a deep learning and NLP approach to the isolation forest algorithm trees. Journal of Theoretical and Applied Electronic Commerce Research, 20(3), p.214. https://doi.org/10.3390/jtaer20030214
9. Wang, Z., 2025. The influence of AI on consumer behavior: Shaping choices and preferences in the digital marketplace. Systems and Soft Computing, p.200397. https://doi.org/10.1016/j.sasc.2025.200397
10. Esmeli, R., Can, A.S., Awad, A. and Bader-El-Den, M., 2025. Understanding customer loyalty-aware recommender systems in E-commerce: an analytical perspective: R. Esmeli et al. Electronic Commerce Research, pp.1-27.https://doi.org/10.1007/s10660-025-09954-6
11. Kumar, B., Roy, S., Sinha, A., Iwendi, C. and Strážovská, Ľ., 2022. E-commerce website usability analysis using the association rule mining and machine learning algorithm. Mathematics, 11(1), p.25.https://doi.org/10.3390/math11010025
12. Premkumar, M., Ashokkumar, S.R., Anupallavi, S., Srinivasan, D., Prema, K. and Kiran, A., 2025. AI Based Strategic Decision Framework for Precise E-Commerce Applications and Ensuring Consumer Trust through Tactical Precision and Decision-Driven Strategies. Results in Engineering, p.108191.https://doi.org/10.1016/j.rineng.2025.108191
13. Oprea, S.V. and Bâra, A., 2025. Diverse counterfactual explanations (DiCE) role in improving sales and e-commerce strategies. Journal of Theoretical and Applied Electronic Commerce Research, 20(2), p.96.https://doi.org/10.3390/jtaer20020096

14. Gulzar, Y., Alwan, A.A., Abdullah, R.M., Abualkishik, A.Z. and Oumrani, M., 2023. OCA: ordered clustering-based algorithm for e-commerce recommendation system. Sustainability, 15(4), p.2947.https://doi.org/10.3390/su15042947
15. Necula, S.C., 2023. Exploring the impact of time spent reading product information on e-commerce websites: A machine learning approach to analyze consumer behavior. Behavioral Sciences, 13(6), p.439.https://doi.org/10.3390/bs13060439
16. Wang, Y. and Zhang, C., 2023. Research on customer purchase intention prediction methods for e-commerce platforms based on user behavior data. Journal of Advanced Computing Systems, 3(10), pp.23-38.https://doi.org/10.69987/JACS.2023.31003