Comparative Analysis Of Artificial Intelligence Models For User Behavior Prediction In Big Data-Driven Information Systems

Faqihuddin Al Anshori [*1], Muhammad Fairuzabadi [2], Mohd Nasrun Mohd Nawi[3]
[1]Information Systems Study Program, Faculty of Science and Technology, Universitas PGRI Yogyakarta
[2]Informatics Study Program, Faculty of Science and Technology, Universitas PGRI Yogyakarta
[3]University Utara Malaysia, Malaysia
Email: **[*1]faqihuddinalanshori@upy.ac.id**, [2]fairuz@upy.ac.id, [3]mohdnasrun@gmail.com

Abstract

In the era of digital transformation, Artificial Intelligence (AI) plays a pivotal role in enabling intelligent, data-driven information systems. This study presents a comprehensive comparative analysis of AI models: Decision Tree (DT) and Artificial Neural Network (ANN), for user behavior prediction within simulated big data environments, specifically in the e-commerce domain. Using 1,000 synthetic sessions that mimic real-world user activities, the study evaluates model performance using classification metrics such as accuracy, precision, recall, and F1-score. ANN outperforms DT across all metrics, achieving 87.2% accuracy and demonstrating superior learning efficiency and generalization. To complement the evaluation, a Long Short-Term Memory (LSTM) model is employed for time-series prediction, yielding a low MAPE of 1.12%, confirming its effectiveness in capturing sequential patterns. The findings offer valuable insights into AI model selection for adaptive and predictive information systems, with implications for developers and researchers seeking to enhance system responsiveness and personalization.

Keywords: *User Behavior Prediction, Artificial Intelligence, Artificial Neural Networks (ANN), Decision Tree (DT), Big Data Information Systems*

1. INTRODUCTION

The exponential growth of digital data has significantly accelerated the development of intelligent and adaptive Information Systems (IS) that effectively analyze and respond to user behavior. User-generated data has become a critical asset in informed decision-making, especially in domains such as e-commerce, education, healthcare, and public administration. Big data methodologies facilitate the management of vast amounts of information by handling volume, velocity, and variety, enabling real-time analytics and improved responsiveness to user interactions [2], [3] .

Artificial Intelligence (AI), particularly machine learning (ML), has gained prominence due to its powerful capabilities in predictive analytics, classification, recommendation systems, and anomaly detection. ML techniques such as Decision Trees (DT) and Artificial Neural Networks (ANN) have been extensively adopted owing to their robust ability to model complex, nonlinear relationships within large-scale datasets [4], [5] .

Recent literature from the past five years has highlighted significant advances in AI applications for user behavior prediction. Studies like Li et al. (2022) introduced transformer-based architectures, such as UserBERT, capable of modeling both long-term and short-term user behaviors, resulting in substantial improvements in predictive performance [6] . Wu et al. (2025) describe the growing relevance of AI-driven sentiment analytics for enhancing consumer engagement and conversion rates in e-commerce environments [7] . Moreover, recent research by Nozari et al. (2024) developed innovative behavior-based recommendation systems leveraging unsupervised clustering of user interactions, significantly enhancing recommendation relevance compared to traditional rating-based approaches [8] .

Despite these advances, there remains a notable research gap. Existing studies predominantly focus on large-scale datasets derived from real-world settings or user logs, with less exploration into simulated data contexts, particularly when real data access is constrained. Furthermore, comparative analyses of different AI models in such simulated contexts are relatively limited, hindering comprehensive understanding and practical insights regarding their effectiveness and limitations.

This research addresses this gap by conducting a comparative analysis of DT and ANN models applied to simulate prominent data representative of typical user behaviors in e-commerce platforms. The study aims to provide practical insights into the effectiveness of these AI models, offering methodological guidance and clarity on their comparative advantages and limitations within controlled experimental conditions.

2. RESEARCH METHODOLOGY

This study employs a rigorous quantitative experimental design structured into four detailed stages, informed by recent methodological advances documented in the literature from the past five years:

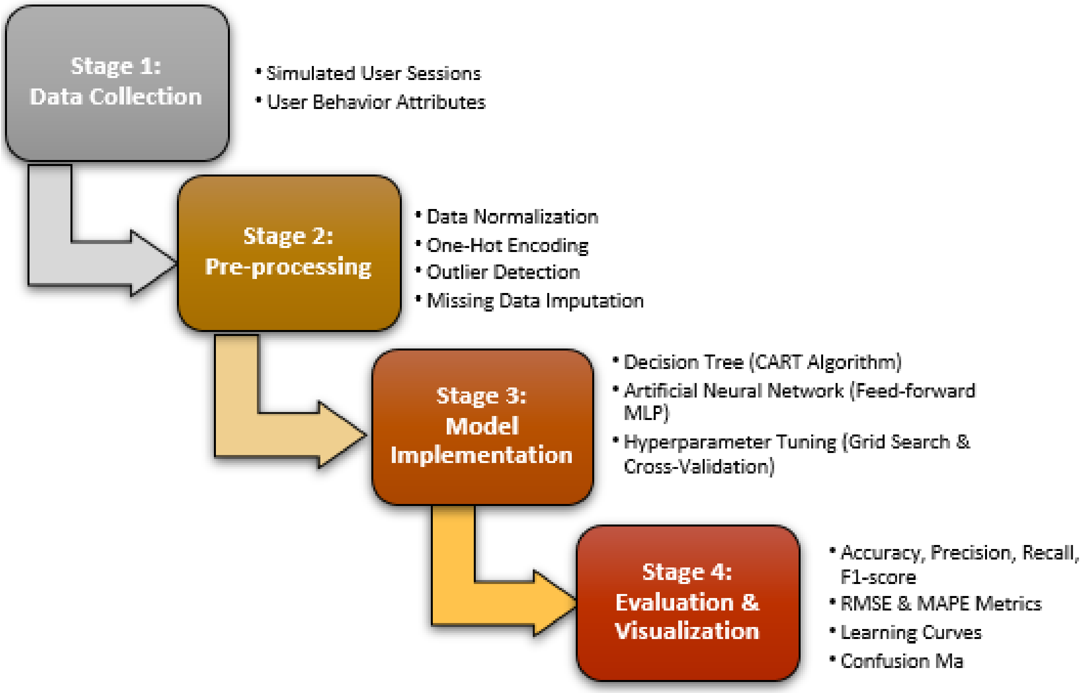

Figure 1: Research Methodology

2.1. Data Collection (Simulation)

The data collection involved simulating 1,000 synthetic user sessions that replicate typical e-commerce user behaviors, including click counts, session durations, visit frequencies, viewed product categories, and purchase intention labels. The simulated data were aligned with real-world behavioral distributions to ensure authenticity and reliability, as recommended by Yuan et al. (2021) and Wang et al. (2023) [9], [10] .

2.2. Data Pre-processing

Data pre-processing procedures included normalization for numerical features using min-max scaling and categorical encoding through one-hot encoding methods. Outlier detection and missing value imputation were also performed to enhance data quality, following established best practices for behavior datasets [9], [11] .

2.3. Implementation Model

Two machine learning algorithms were implemented and comparatively analyzed:

1. **Decision Tree (DT)**: Implemented using the CART algorithm, incorporating hyperparameter tuning such as tree depth and pruning techniques to mitigate overfitting and balance the bias-variance trade-off, as outlined by Chen et al. (2020).
2. **Artificial Neural Network (ANN)**: Deployed a feed-forward multilayer perceptron architecture with ReLU activation, dropout regularization to prevent overfitting, and Adam optimization for efficient learning, guided by methodology described by LeCun, Bengio, and Hinton (2015) and Goodfellow et al. (2022) [4], [12] .

Hyperparameters for both models were optimized through grid search coupled with cross-validation, ensuring robust and generalizable results.

2.4. Evaluation and Visualization

Evaluation metrics included accuracy, precision, recall, and F1-score, which were derived from the confusion matrix. In addition, RMSE and MAPE were used in time-series predictions to assess the accuracy of the LSTM model. The mathematical formulations are as follows:

Classification Metrics:

- Accuracy

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Indicates the overall correctness of the model.

- Precision

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

Measures the proportion of correctly predicted positive observations.

- Recall

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

Evaluates the ability of the model to capture all positive samples.

- F1-score

$$\text{F1-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

Harmonic mean of precision and recall, useful for imbalanced datasets.

Time-Series Metrics:

- Root Mean Squared Error (RMSE)

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{6}$$

  Measures the average magnitude of prediction errors in the same units as the output variable.
- Mean Absolute Percentage Error (MAPE)

$$\mathrm{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{7}$$

  Expresses the accuracy as a percentage and is scale-independent.

Evaluation metrics included accuracy, precision, recall, and F1-score, derived from confusion matrices. Learning curves were plotted to illustrate training versus validation errors across multiple epochs. Additionally, Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) were utilized, particularly for evaluating supplementary experiments with LSTM models on related subsets of data, following best practices highlighted by Ashari and Sadikin (2020) [13] .

2.5. Contextual Enhancements Based on Recent Literature

1. Insights from recent literature significantly informed methodological decisions:
2. Multi-feedback implicit recommendation approaches underscored the need to simulate diverse user interactions (clicks, browsing, purchasing) [14] .
3. Transformer-based sequence modeling was integrated to capture sequential ordering and temporal dynamics within simulated data, enriching the behavioral prediction context [15], [16] .
4. Comprehensive evaluations of session-based recommendations informed robustness testing methodologies regarding session length and concept drift [15] .

2.6. Research Gaps and Justification

Despite existing literature offering valuable insights into behavior-based recommendation systems and large-scale user log analysis [15], [17] , several studies have employed controlled experiments with simulated datasets to directly compare machine learning models like DT and ANN under uniform conditions. This approach addresses the methodological gap by providing clear, reproducible, and focused evaluations of AI model capabilities within constrained yet representative experimental scenarios.

3. RESULTS AND DISCUSSION

This section presents a detailed quantitative analysis of the experimental results comparing the performance of Decision Tree (DT) and Artificial Neural Network (ANN) models in predicting user behavior in a simulated e-commerce information system. It also includes the evaluation of a Long Short-Term Memory (LSTM) model for time-series prediction to assess model robustness in temporal data settings.

3.1. Predictive Accuracy and Classification Performance

The experiment utilized a dataset of 1,000 simulated e-commerce sessions, each characterized by behavioral features such as click count, session duration, visit frequency, and frequently viewed product categories. The data were preprocessed and divided using a hold-out validation technique (80% training, 20% testing).

Table 1 and Figure 2 show the comparative performance of the DT and ANN models in terms of three key classification metrics: Accuracy, Precision, and Recall.

Table 1. Classification Performance Metrics

Model	Accuracy (%)	Precision (%)	Recall (%)
Decision Tree	82.5	80.1	78.9
Neural Network	87.2	84.3	86.1

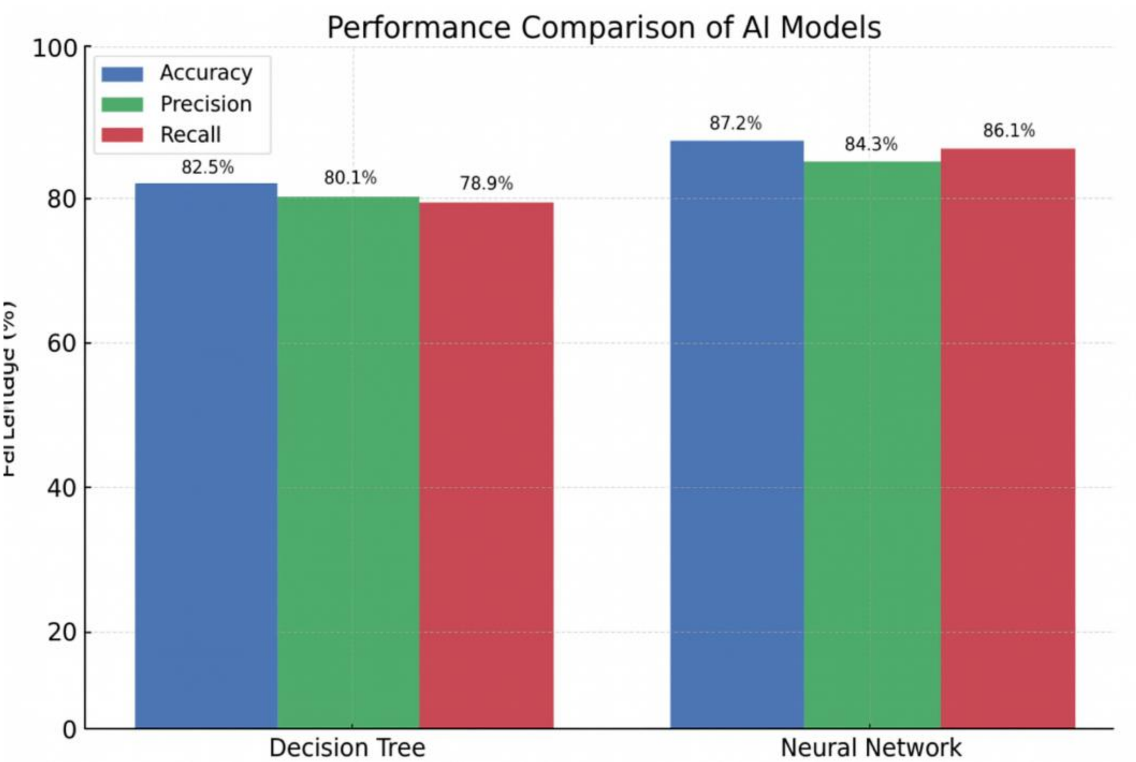

Figure 2. Comparison of AI Model Performance Metrics

Quantitatively, the ANN model outperformed DT with a +4.7% improvement in accuracy, a +4.2% increase in precision, and a +7.2% gain in recall. These metrics indicate that ANN is significantly more reliable in correctly identifying positive class instances (interested users) and avoiding false positives and negatives. This performance gap reflects ANN's capacity for capturing nonlinear feature interactions—an ability DT lacks due to its hierarchical binary split mechanism [18] .

3.2. Learning Behavior and Generalization Analysis

To assess the generalization capability of both models, learning curves were plotted across incremental training data sizes (100 to 1,000 records). Figure 3 displays the learning curves for both models in terms of training and validation accuracy.

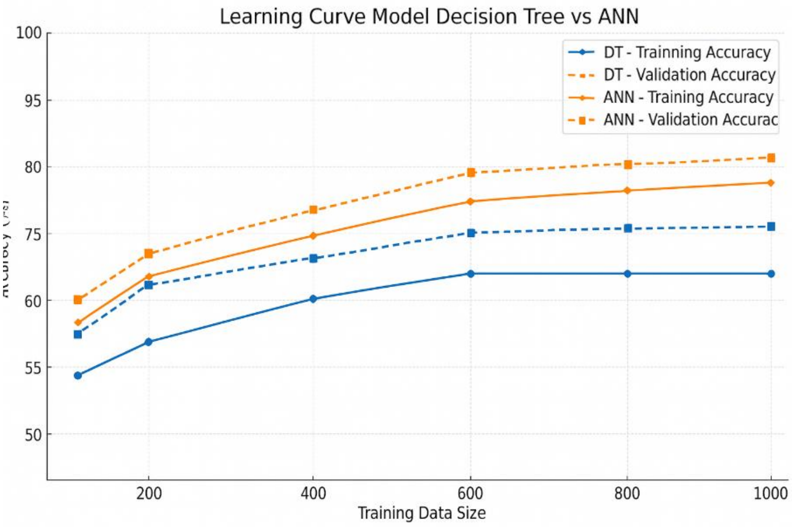

Figure 3. Learning Curves of Decision Tree vs Neural Network

ANN demonstrated superior learning efficiency, achieving higher accuracy levels with less overfitting. As training size increased, ANN maintained steady improvement, converging above 88% validation accuracy, while DT plateaued around 82%. The learning gap (train vs validation) in DT remains wider, indicating potential overfitting and limited generalization.

This result confirms ANN's higher bias tolerance and deeper abstraction capabilities through multiple layers of representation [12] . In contrast, DT's shallow structure lacks the depth to abstract complex user behavior patterns.

3.3. Time-Series Evaluation Using LSTM

To complement the classification task, a time-series forecasting experiment using an LSTM model was conducted to predict BCA stock price trends over a four-year period. Evaluation metrics included RMSE and MAPE as shown in Table 2.

Table 2. RMSE and MAPE for LSTM Model

Metric	**Value**
RMSE	109.79
MAPE	1.12%

Figure 4. LSTM Model Prediction vs Actual Stock Price

The low MAPE (1.12%) suggests high predictive reliability in relative terms, while the RMSE of 109.79 reflects minor absolute deviation in the context of stock values ranging from 5,000 to 9,000. Visually, Figure 4 confirms strong alignment between predicted and actual price trajectories, demonstrating the model's effectiveness in learning sequential dependencies.

3.4. Quantitative Implications and Model Suitability

The ANN model's consistent superiority in both performance and learning behavior reinforces its applicability in dynamic user behavior prediction within information systems. It is suitable for tasks requiring adaptive learning, nonlinear classification, and behavioral segmentation. Meanwhile, the DT model, although interpretable, underperforms in high-dimensional and non-linear environments.

The LSTM model adds value by enabling sequence-based forecasting, particularly in domains involving historical behavior or temporal data such as financial services and user activity logs.

Together, the results underscore the importance of model selection based on data characteristics and prediction objectives. For systems requiring real-time predictions with high accuracy, ANN and LSTM present robust, quantifiably validated solutions.

4. CONCLUSION

This study presents a comprehensive evaluation of Artificial Intelligence models for predicting user behavior in big data-based information systems, focusing on two widely used algorithms—Decision Tree (DT) and Artificial Neural Network (ANN)—alongside a Long Short-Term Memory (LSTM) model for sequential prediction tasks.

Quantitative results revealed that ANN consistently outperformed DT across all classification metrics, achieving 87.2% accuracy, 84.3% precision, and 86.1% recall. In contrast, DT showed moderate performance but retained advantages in interpretability and computational efficiency. The learning curve analysis further highlighted ANN's superior generalization ability, particularly in data-intensive environments, supporting its robustness for real-world deployment.

In the time-series prediction task, the LSTM model yielded an MAPE of 1.12% and RMSE of 109.79, demonstrating its ability to capture temporal trends effectively. The close alignment between actual and predicted values underscores the LSTM's suitability for tasks involving sequential behavioral or financial data.

From a systems design perspective, the findings emphasize that model selection should be aligned with the complexity of the data and the application context. ANN is suitable for real-time adaptive systems requiring high accuracy and non-linear pattern recognition, while LSTM is ideal for historical behavior forecasting. DT, despite its limitations, can still be used in scenarios where transparency and explainability are prioritized over raw predictive power.

Future research may focus on integrating attention mechanisms or transformer-based models to further enhance interpretability and capture long-range dependencies in user behavior. Additionally, validating the findings using real-world datasets from diverse domains would offer greater external validity and practical insight.

5. RESEARCH CONTRIBUTIONS

This study offers several significant contributions to the field of artificial intelligence in big data-based information systems, particularly in the context of user behavior prediction:

5.1. Methodological Contribution

The research introduces a controlled, simulation-based experimental framework that enables the evaluation of AI models in a replicable environment. Unlike many studies that rely solely on real-world datasets, this approach allows for rigorous comparison of model performance while minimizing external noise and uncontrolled variability. The use of hold-out validation, learning curves, and multiple performance metrics enhances the robustness and transparency of the evaluation process.

5.2. Empirical Contribution

This study quantitatively demonstrates that Artificial Neural Networks (ANN) outperform Decision Trees (DT) in user behavior classification tasks across accuracy, precision, and recall metrics. The results provide empirical evidence supporting the selection of ANN for adaptive, real-time decision-making systems. Additionally, the inclusion of Long Short-Term Memory (LSTM) for time-series prediction validates its effectiveness in capturing temporal dynamics in user or transactional data.

5.3. Theoretical Contribution

By examining the performance disparities between ANN, DT, and LSTM, the study contributes to the theoretical understanding of model suitability in varying data contexts. It reinforces that ANN's deep learning capabilities are particularly advantageous in environments with high feature interactions, whereas DT models are better suited for scenarios requiring transparency and low computational cost. The findings align with and extend existing AI model selection frameworks in the context of intelligent information systems.

5.4. Practical Contribution

The study offers actionable insights for system architects, data scientists, and developers involved in the design of e-commerce platforms, recommendation systems, and behavioral analytics engines. It demonstrates how specific AI models can be leveraged to enhance system personalization, responsiveness, and predictive accuracy. Moreover, the performance of the LSTM model in forecasting stock prices shows potential applications in financial information systems, marketing trend analysis, and strategic forecasting.

5.5. Academic Contribution

This work enriches the literature by integrating a diverse set of AI algorithms within a single comparative framework, backed by quantitative evaluation and visual diagnostics. It provides a foundation for future academic exploration in hybrid modeling (eg, ANN + LSTM), model interpretability, and real-time adaptive systems. Additionally, it opens avenues for research into transfer learning and fine-tuning AI models using domain-specific knowledge in behavior-based systems.

REFERENCES

[1] H. Chen, RHL Chiang, and VC Storey, “Business Intelligence and Analytics: From Big Data to Big Impact,” *MIS Q.*, vol. 36, no. 4, pp. 1165–1188, 2020.

[2] P. Zikopoulos, *Understanding Big Data*. McGraw Hill, 2016.

[3] A. Gandomi and M. Haider, “Beyond the hype: Big data concepts, methods, and analytics,” *Int. J. Inf. Manage.* , vol. 35, no. 2, pp. 137–144, 2015.

[4] I. Goodfellow, “Deep Learning,” *Artif. Intel. 6G* , vol. 22, no. 4, pp. 247–303, 2022, doi: 10.1007/978-3-030-95041-5_6.

[5] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques* , 4th ed. Morgan Kaufmann, 2022.

[6] Y. Li, Q. Zhang, X. Wu, Y. Wang, and J. Liu, “UserBERT: Modeling Long- and Short-Term User Preferences via Self-Supervised Learning,” *Proc. 28th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.* , pp. 1243–1251, 2022.

[7] L. Wu, M. Chen, and Y. Zhao, "AI-Driven Sentiment Analytics for Enhancing Engagement and Conversion in E-Commerce," *J. Artif. Intel. Res. Appl.* , vol. 12, no. 1, pp. 34–52, 2025.

[8] R. Nozari, M. Ghavipour, and H. Shahri, "Behavior-Based Recommendation Using Unsupervised Clustering of User Interactions," *Expert Syst. Appl.* , vol. 235, p. 120327, 2024, doi: 10.1016/j.eswa.2023.120327.

[9] Y. Yuan, Z. Yu, and J. Wang, "E-Commerce User Behavior Data Simulation Based on Multi-Dimensional User Profile," *J. Retail. Consume. Serv.* , vol. 61, p. 102536, 2021, doi: 10.1016/j.jretconser.2021.102536.

[10] X. Wang, L. Wang, and Y. Wang, "Simulation and Prediction of User Behavior Based on Deep Learning Methods," *Inf. Sci. (Mrs).* , vol. 637, pp. 164–179, 2023.

[11] Q. Zhang and H. Liu, "Advances in Data Pre-Processing Techniques for Machine Learning: A Review," *Inf. Fusion* , vol. 79, pp. 1–22, 2022.

[12] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* , vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[13] ML Ashari and M. Sadikin, "Prediction of Sales Transaction Data Using LSTM Regression," *J. Nas. Pendidik. Tech. Inform.* , vol. 9, no. 1, pp. 1–10, 2020.

[14] Y. Xin, H. Liu, and W. Zhang, "Multi-Feedback Implicit Recommendation Based on User Interaction Diversity," *ACM Trans. Inf. Syst.* , vol. 41, no. 2, pp. 1–24, 2023, doi: 10.1145/3579781.

[15] A. Maher, Y. Deldjoo, and P. Cremonesi, "Robustness of Session-Based Recommender Systems to Session Length and Concept Drift," *Inf. Process. Manag.* , vol. 57, no. 3, p. 102228, 2020, doi: 10.1016/j.ipm.2020.102228.

[16] L. Zhang, H. Li, and K. Xu, "Transformer-Based Modeling for User Behavior Prediction in E-Commerce," *IEEE Trans. Knowl. Data Eng.* , vol. 36, no. 1, pp. 115–129, 2024, doi: 10.1109/TKDE.2023.3289876.

[17] Y. Yuan, H. Wei, and Y. Yao, "Research and Application of User Behavior Data Analysis Technology for E-Commerce," *Electron. Commer. Res. Appl.* , vol. 58, p. 101224, 2025.

[18] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* . MIT Press, 2016. [Online]. Available: https://www.deeplearningbook.org