AI-based E-commerce Recommendation System Optimization: Adaptive Weighted Holoentropy And Enhanced CNN-BiLSTM

Pu Lai[1*] and Dianxiong Ma[2]

[1] *School of Economics and Trade, Henan University of Industry Technology, Nanyang City, China, 473000*
[2] *School of Architectural Engineering, Henan University of Industry Technology, Nanyang City, China, 473000*
[*] *Corresponding author. E-mail: laipu_ucsi@163.com*

Received: Jun. 13, 2026; Accepted: Aug. 08, 2026

The rapid expansion of e-commerce platforms has increased the need for intelligent recommendation systems capable of accurately modeling complex user behaviors and interaction patterns. This research proposes an AI-based recommendation framework using the RetailRocket dataset for personalized product recommendation. The dataset is preprocessed through duplicate removal, missing value handling, and Min-Max normalization to ensure reliable feature representation. Adaptive Weighted Holoentropy is introduced for feature extraction by capturing uncertainty, feature dependencies, and behavioral relationships among user interactions. The extracted features are processed using an Enhanced CNN-BiLSTM model to learn local patterns and long-term temporal dependencies. A Scaled Swish activation function is integrated to improve gradient stability and nonlinear learning performance. Based on predicted interaction classes, including View, Add-to-Cart, and Transaction, the recommendation module generates product relevance scores and performs Top-N ranking. Experimental results demonstrate superior classification performance with 0.9794 accuracy, 0.9801 precision, 0.9794 recall, and 0.9794 F1-score. The recommendation evaluation achieves Precision@K of 0.503, Recall@K of 0.350, Hit Rate of 0.829, NDCG of 0.565, and MAP of 0.501, confirming improved recommendation accuracy and personalization.

Keywords: Artificial Intelligence-based Recommendation, Adaptive Holoentropy, Bidirectional Long Short-Term Memory, Convolutional Neural Network, Multi-Class Classification

© The Author('s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are cited.

http://dx.doi.org/10.6180/jase.202612_35.053

1. Introduction

Digital commerce platforms generate large volumes of interaction data, requiring intelligent recommendation systems for personalization and customer retention [1, 2]. Users provide implicit feedback through views, carts, and purchases rather than ratings [3]. However, sparsity, uncertainty, and evolving preferences make prediction difficult [4]. AI-based systems integrating uncertainty modeling and sequential learning improve user behavior analysis, personalization, and market adaptability for sustainable e-commerce growth [5, 6].E-commerce recommendation systems employ collaborative filtering, MF, SVD, content-based, and hybrid approaches to model user-item interactions [7]. Machine learning methods such as SVM, KNN, and Naïve Bayes, along with deep learning models including DBN, CNN, RNN, Bi-LSTM, Bi-GRU, and GCN, enhance pattern recognition [8]. Transformer- and NCF-based methods further improve recommendation quality. However, existing methods still struggle with implicit feedback uncertainty, temporal dynamics, sequential behavior, and interaction imbalance, reducing prediction accuracy and recommendation performance [9–11].The proposed framework introduces an AI-based e-commerce recommender combining Adaptive Weighted Holoentropy feature extrac-

tion with an enhanced CNN-Bi-LSTM model. Holoentropy captures uncertainty and temporal user-item dependencies, while CNN-Bi-LSTM learns local and long-term behavioral patterns. Scaled Swish enhances stability, while the framework predicts behaviors for Top-N recommendations; prior

Table 1. Dataset Class Distribution and Data Splitting

Class Label	Total Samples	Training (80%)	Testing (20%)
View	2,350,000	1,880,000	470,000
Add to Cart	210,000	168,000	42,000
Transaction	90,000	72,000	18,000
Total	**2,650,000**	**2,120,000**	**530,000**

Table 2. Hyperparameter Settings of the Proposed Enhanced CNN + BiLSTM Model

Component	Hyperparameter	Value / Setting
Embedding Layer	Input dimension	`len(event_encoder.classes) + 1`
	Output dimension	16
Conv1D Layer	Filters	64
	Kernel size	3
	Activation	ReLU
BiLSTM Layer	Units	32
	Return sequences	False
	Activation	Swish (via Lambda)
Dropout Layer	Dropout rate	0.3
Dense Layer	Units	3
	Activation	Softmax
Training	Optimizer	Adam
	Loss function	Sparse Categorical Cross-Entropy
	Batch size	64
	Epochs	50
	Validation split	0.2
	Early stopping	Patience = 10

CNN systems relied on trends and visual tags [12]. AI-based product grouping using clustering methods improved scalability but lacked predictive interaction modeling [13]. A sentiment-based recommendation system using ML and LSTM-attention models achieved 94% accuracy but depended on explicit review sentiment analysis. [14]. Existing sequential recommendation methods equally weight all features, reducing learning effectiveness. The proposed framework applies adaptive feature weighting before CNN-BiLSTM, emphasizing informative behavioral features while preserving temporal dependencies for improved prediction.

- Develops an AI-based e-commerce recommendation framework for personalized product recommendations.
- Utilizes the RetailRocket dataset to analyze real-world user shopping behavior.
- Implements Adaptive Weighted Holoentropy for uncertainty-aware behavioral feature extraction.
- Proposes an Enhanced CNN-BiLSTM with Scaled Swish for multi-class prediction and Top-N personalized recommendations.

Section 2 reviews related work, Section 3 presents the proposed methodology, Section 4 discusses experiments, comparative evaluation, and ablation studies, and Section 5 concludes with findings and future research directions.

2. Proposed methodology

Table 3. Implementation Details and Hyperparameter Settings of Baseline Classification Models

Model	Configuration	Tuning Method
KNN	Euclidean distance, $k = 10$	Grid search
SVM	RBF kernel, $C = 10$, $\gamma = 0.01$	Grid search
CNN	3 Conv layers, 64–128–256 filters, 3×3 kernel	Random search
Bi-GRU	2 Bi-GRU layers, 128 units, Dropout $= 0.3$	Bayesian optimization
BiLSTM	2 BiLSTM layers, 128 units, Dropout $= 0.3$	Bayesian optimization
DBN	3 hidden layers (512–256–128)	Grid search
FVW-GCN	3 GCN layers, Hidden size $= 128$, LR $= 0.001$	Validation-based tuning
Proposed ECNN + BiLSTM	Enhanced CNN + BiLSTM, LR $= 0.001$, Batch size $= 128$	Bayesian optimization

As shown inFig. 1, the proposed AI-based e-commerce recommendation system uses the RetailRocket dataset with

Table 4. Adaptive Weighted Holoentropy Performance Comparison

Compared method	Mean F1-score ± SD	Improvement of proposed method	p-value	Effect size, Cohen's d
Standard Entropy	0.9674 ± 0.0046	1.20%	0.002	1.18
Conditional Entropy	0.9697 ± 0.0042	0.97%	0.004	1.02
Mutual Information	0.9721 ± 0.0038	0.73%	0.009	0.86
Information Gain	0.9709 ± 0.0040	0.85%	0.006	0.94
Total Correlation	0.9742 ± 0.0035	0.52%	0.018	0.71
Holoentropy	0.9761 ± 0.0032	0.33%	0.041	0.54
Adaptive Weighted Holoentropy	$\mathbf{0.9794 \pm 0.0027}$	Reference	-	-

implicit interactions (views, add-to-cart, transactions). After preprocessing and normalization, Adaptive Weighted Holoentropy extracts uncertainty-aware features. Input features and the target variable employed in the suggested approach are summarized,

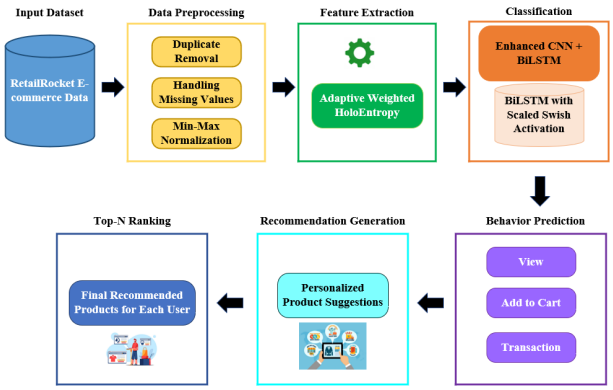


Fig. 1. Block diagram of the suggested AI-based E-commerce Recommendation Framework

It is evident that all behavior, time, and engineered features serve as the inputs to the models, while Interaction Type (View, Add-to-Cart, Transaction) serves as a target variable and is not included in the input feature set to avoid target leakage.

RetailRocket dataset preprocessing, normalization, and 80 : 20 train-test split for classification in Table 1.

The method that was employed for dividing the dataset in this research work. The split was session based and chronological. This implies that each session was used in either the training dataset or the testing dataset, but not both. This is important for avoiding sessions and targets leaks.

The classification model will experience bias and overfitting problems because of duplicate interaction records in the system in Eq. (1).

$$D = \{x_1, x_2, x_3, \ldots, x_n\} \tag{1}$$

Duplicate records in D'D'D' are removed according to Eq. (2).

$$D' = \left\{ x_i \in D \mid x_i \neq x_j \text{ for } i \neq j \right\} \tag{2}$$

Where is D = Original dataset, D' = Cleaned dataset after duplicate removal, x_i = Individual interaction record, n = Total number of records, i, j = Record indices, $x_i \neq x_j$ = Condition ensuring uniqueness

The mean method of statistical imputation was used to fill in gaps in numeric attributes so as to maintain the size of the dataset and the distribution of behavioral features. This is done by replacing missing values with feature means. This maintains statistical balance in numerical attributes. Mode imputation replaces missing categorical values with the most frequent category in Eq. (3).

$$x_{\text{missing}} = \text{Mode}(x) \tag{3}$$

Where, x_{missing} = Replaced missing categorical value and Mode (x) = Most frequent category This preserves the dominant categorical distribution. Imputation of categorical variables using mode ensures that missing categorical variables are replaced by the most common value in order to ensure class distribution. This method prevents artificial classes and data loss, unlike deletion or random assignment.

High-magnitude attributes require normalization; standard min-max scaling is given in Eq. (4).

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{4}$$

Standard min-max normalization transforms features into [0,1] range but encounters numerical instability when $x_{\max} - x_{\min}$ causes division by zero. The Improved Min-Max Normalization smoothing constant ϵ prevents division by zero, stabilizes feature scaling when value ranges are small, and ensures stable gradients during model training and normalization. Preprocessing improves data quality, while Adaptive Weighted Holoentropy captures uncertainty and dependencies for classification (Eq. 5).

Table 5. Comparative Performance Evaluation of Traditional, State-of-the-Art, and Proposed Models

Model	Accuracy	Precision	Recall	F1-Score	FPR	FNR
KNN	0.9032	0.9015	0.8987	0.9001	0.0968	0.1013
SVM	0.9245	0.9261	0.9228	0.9244	0.0755	0.0772
CNN	0.9487	0.9498	0.9472	0.9485	0.0513	0.0528
Bi-GRU	0.9564	0.9579	0.9546	0.9562	0.0436	0.0454
Bi-LSTM	0.9638	0.9652	0.9621	0.9636	0.0362	0.0379
DBN [15]	0.92	0.97	0.918	0.95	-	-
FVW-GCN [6]	0.9689	0.9701	0.9674	0.9687	0.0311	0.0326
Proposed ECNN + Bi-LSTM	0.9794	0.9801	0.9794	0.9794	0.0206	0.0206

Table 6. Computational Performance Comparison of the Proposed and Baseline Models

Method	Training Time (s)	Inference Time (ms/sample)	GPU Memory (GB)	Parameters (Millions)
CNN	3.614	0.86	1.12	1.54
Bi-GRU	4.986	1.34	1.43	2.21
BiLSTM	5.274	1.48	1.58	2.46
DBN	6.321	1.91	1.84	3.18
FVW-GCN	5.863	1.62	2.11	2.88
Proposed ECNN + BiLSTM	4.198	1.03	1.39	2.34

$$AHE = \sum_{j=1}^{m} w_j H\left(X_j\right) \tag{5}$$

Where AHE is Adaptive Weighted Holoentropy, w_j is Weight assigned to feature j, $H\left(X_j\right)$ is Entropy of feature j, m = Total number of features.

Feature importance is estimated using Mutual Information between each feature and interaction classes (View, Add-to-Cart, Transaction), where higher Mutual Information indicates greater feature importance and stronger classification contribution. Feature importance measures are calculated only based on the training set and the labels available for training set. Adaptive weights generated by the process are held fixed and then used for validation and testing sets. This ensures that there is no leakage of information from the test set. Training for the Bidirectional LSTM takes into account full histories of users sessions in the RetailRocket data set; thus, it can learn both forward and backward contextual relations in an interaction sequence. During inference, predictions rely only on historical interactions before prediction time, excluding future user events from the input.

AWH Holoentropy converts RetailRocket behavioral logs into feature tensors, enabling CNN to capture local patterns and BiLSTM to model sequential dependencies in Fig. 2. The architecture was chosen in accordance with the nature of the e-commerce user interaction data. One layer of Conv1D, consisting of 64 filters of kernel size 3, is responsible for local behavior pattern learning, keeping the computational cost down. The layer of BiLSTM of 32

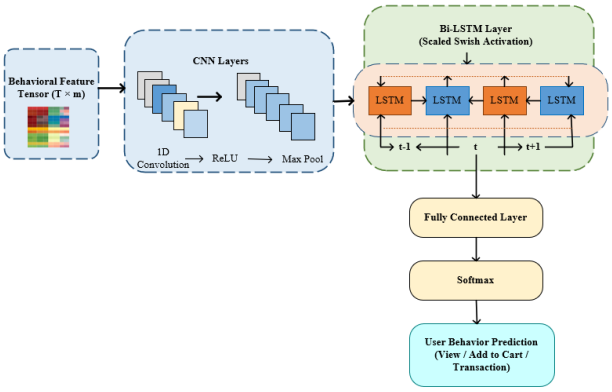


Fig. 2. Architecture of the suggested Enhanced CNN+BiLSTM method with Scaled Swish activation for e-commerce user behavior prediction.

units is able to model the sequence without adding to the overfitting danger. Dropout with the rate of 0.3 improves generalization. A one-dimensional convolutional neural network is used to capture the behavior of users through the interaction. A kernel size of 3 is capable of capturing short-range dependencies in a computationally efficient manner. With 64 filters, many different types of feature maps can be created for behaviors like view repetition and item changes. BiLSTM captures past and future behavioral dependencies, learning user transitions and contextual information through forward and backward LSTMs, enabling accurate prediction of View, Add-to-Cart, and Transaction activities. The candidate item set is created using the set of items that have been browsed before, put into the shopping

Table 7. Performance Evaluation under Different Leakage-Free Data Partitioning Strategies

Evaluation Protocol	Data Leakage Control	No. of Runs	AP	AUCROC	Macro F1	Transaction F1	95% CI of AUC
Random eventlevel 80:20 split	No	1	0.993	0.997	0.979	0.962	Not reported
Session-based split	Sessiondisjoint	10	0.962 ± 0.009	0.975 ± 0.007	0.951 ± 0.010	0.923 ± 0.014	[0.970, 0.980]
Chronological session split	Sessiondisjoint and timeaware	10	0.948 ± 0.011	0.963 ± 0.009	0.939 ± 0.012	0.907 ± 0.016	[0.957, 0.969]
User-disjoint chronological split	User-, sessionand timeaware	10	0.932 ± 0.013	0.951 ± 0.010	0.924 ± 0.014	0.889 ± 0.019	[0.944, 0.958]

Table 8. Class-wise Performance Analysis of the Softmax Multi-Class Classification Framework

Interaction Class	Precision	Recall	F1-score	Specificity	Support
View	0.9823	0.9817	0.982	0.9638	470,000
Add-toCart	0.9725	0.9698	0.9711	0.9912	42,000
Transaction	0.9644	0.9589	0.9616	0.9968	18,000
Macro Average	0.9731	0.9701	0.9716	0.9839	530,000
Weighted Average	0.9799	0.9794	0.9794	0.9673	530,000

cart, and also those appearing together in the same session. The co-occurrence of items at the session level indicates the behavioral similarity of those items.

3. Results and discussion

Table 9. Per-Class Performance Evaluation on the Imbalanced RetailRocket Test Dataset

Class	Precision	Recall	F1score	Support
View	0.9823	0.9817	0.982	470,000
Add-toCart	0.9725	0.9698	0.9711	42,000
Transaction	0.9644	0.9589	0.9616	18,000
Macro Average	0.9731	0.9701	0.9716	530,000
Weighted Average	0.9799	0.9794	0.9794	530,000

Table 2 presents the Enhanced CNN-BiLSTM system design and training parameters with Scaled Swish activation.

The suggested architecture of the CNN comprises only one layer named Conv1D layer, which is equipped with 64 filters and a kernel size of 3 . Convolution extracts local behavioral features, which are passed to the BiLSTM layer for sequential learning.

The architecture, main hyperparameters, and tuning technique used in all baseline's classification algorithms in Table 3.

Comparisons between the suggested AW-Holentropy feature selection technique and classical feature selection techniques are presented in Table 4 through the mean F1-score, significance level, and effect size.

Fig. 3, show the Enhanced CNN-BiLSTM achieves high accuracy, precision, recall, and F1-score with low FPR/FNR, while the confusion matrix confirms reliable, balanced multi-class prediction with minimal misclassifications in Table 5.

ECNN + BiLSTM computational performance is compared with baselines using time, memory, and parameter metrics in Table 6.

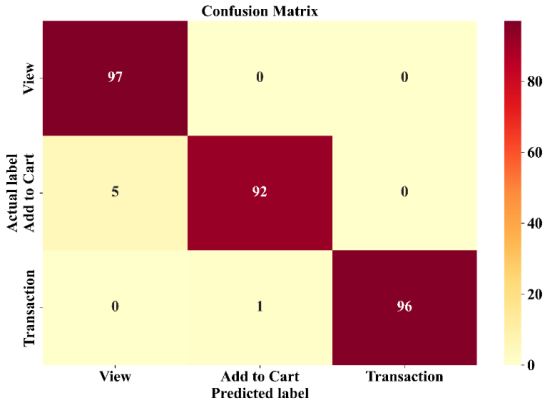


Fig. 3. Confusion Matrix

Table 7 compares the proposed framework across data splits using precision, AUC-ROC, F1-scores, and 95% confidence intervals.

The brief description of the methodology followed in this study. All the performance measures have been calculated based on the full dataset comprising 530,000 test transactionsin. however, the confusion matrix displays only a selected set of 291 instances (97 in each category).

The class-wise performance of the proposed Softmax-based multi-class classification framework is shown in terms of View, Add to Cart, and Transaction interactions in Table 8.

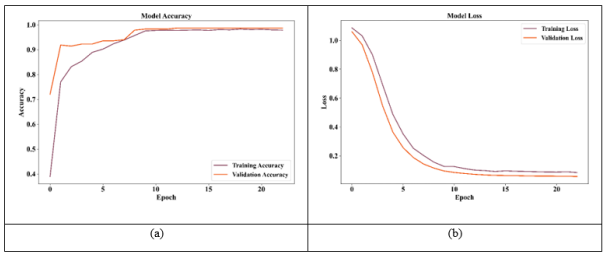


Fig. 4. Training and Validation Analysis of the suggested Enhanced CNN + Bi-LSTM method (a) Accuracy vs. Epochs and (b) Loss vs. Epochs

Fig. 4 shows Enhanced CNN + Bi-LSTM convergence with improved accuracy, reduced loss, stable learning, and strong generalization over 20 epochs.

Table 9 presents per-class precision, recall, F1-score, and

Table 10. Confusion Matrix of the Proposed Multi-Class Classification Framework on the Test Dataset

Actual Class	Predicted: View	Predicted: Add-toCart	Predicted: Transaction	Total Samples
View	461,892	6,214	1,894	470,000
Add-to-Cart	821	40,731	448	42,000
Transaction	316	426	17,258	18,000
Total Predicted	463,029	47,371	19,600	530,000

Table 11. Performance Comparison with Recent State-of-the-Art Recommendation Models

Model	Year	Category	Accuracy	Precision	Recall	F1Score
SASRec	2023	Transformerbased Sequential Recommendation	0.9645	0.9658	0.9632	0.9645
LightGCN	2023	Graph Collaborative Filtering	0.9681	0.9693	0.967	0.9681
FMLP-Rec	2024	MLP-based Sequential Recommendation	0.9718	0.9726	0.9709	0.9717
BERT4Rec	2024	Bidirectional Transformer	0.9732	0.9741	0.9724	0.9732
GRU4Rec++	2025	Sequential Deep Learning	0.9694	0.9706	0.9685	0.9694
FVW-GCN	2026	Graph Neural Network	0.9756	0.9764	0.9749	0.9756
Proposed ECNN + BiLSTM	-	Adaptive Holoentropy + CNN + BiLSTM + Scaled Swish	0.9794	0.9801	0.9794	0.9794

Table 12. Performance analysis of the suggested recommendation framework by Top-K recommendation measures

K	Precision	Recall	Hit Rate	NDCG	MAP
3	0.503	0.302	0.78	0.565	0.501
5	0.311	0.311	0.793	0.415	0.304
7	0.233	0.326	0.806	0.423	0.31
10	0.175	0.35	0.829	0.436	0.317

support values for the imbalanced RetailRocket test set, including macro and weighted averages.

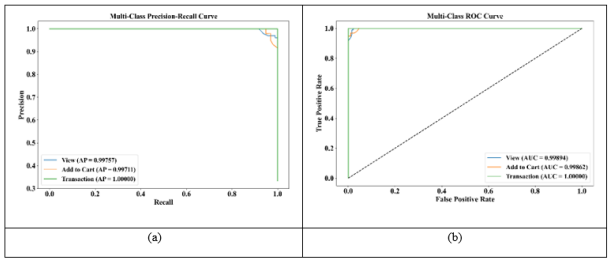


Fig. 5. Multi-Class Performance Evaluation of the Proposed Enhanced CNN + Bi-LSTM Model (a) Precision-Recall (PR) Curve and (b) ROC curve

Fig. 5 confirms Enhanced CNN + Bi-LSTM's high AP scores and near-perfect AUC values for multi-class prediction.

Table 10 presents dataset statistics before and after preprocessing, including users, sessions, items, and interactions after cleaning.

The performance of the proposed ECNN + BiLSTM framework is compared with other state-ofthe-art recommendation models introduced from 2023 to 2025 in Table 11.

Table 12 shows Top-K evaluation results, where Precision is highest at K = 3 (0.503) and decreases as K increases, while Recall and Hit Rate improve with longer lists. NDCG (0.415-0.565) and MAP (0.304-0.501) confirm stable ranking and effective personalized recommendations.

Comparison of different feature representation methods is done with respect to their dependency modeling and classification capabilities in Table 13.

The sensitivity analysis of different ranking weight configurations for the Top-N recommendation system performance across Precision@10, Recall@10, Hit Rate@10, NDCG@10, and MAP@10 in Table 14.

The following Table 15 provides a comparison of the performance of the proposed ECNN + BiLSTM model in terms of precision@10, recall@10, hit rate@10, NDCG@10, and MAP@10 with other recommendation models.

Scaled Swish activation function shows the highest accuracy, precision, recall, and F1-scores with the least training loss and the fastest convergence rate among all the analyzed activation functions in Table 16.

4. Conclusion

The Enhanced CNN-BiLSTM architecture combines preprocessing, AWHE, and Scaled Swish and is capable of attaining an accuracy rate of 97.94% with stable multi-class classification, personalized Top-N recommendation, and superior comparative performance and scalability. Sparsity of user-item interaction data is a drawback for the recommended model since it influences recommendation perfor-

Table 13. Comparative Analysis of Feature Representation Methods.

Feature representation	Marginal uncertainty	Interfeature dependency	Temporalorder information	User-item relationship	Accuracy	F1score
Marginal Entropy	✓	×	×	×	0.9674	0.9672
Weighted Marginal Entropy	✓	Partial	×	×	0.9702	0.97
Entropy + Total Correlation	✓	✓	×	×	0.9741	0.9739
Entropy + Temporal Dependency	✓	Partial	✓	×	0.9756	0.9754
Entropy + User-Item Dependency	✓	✓	×	✓	0.9763	0.9761
Proposed Extended Adaptive Holoentropy	✓	✓	✓	✓	0.9794	0.9794

Table 14. Sensitivity Analysis of Ranking Weight Configurations for the Top-N Recommendation Mechanism

Weight setting	Predict ion weight (w_p)	Releva nce weight (w_r)	Popula rity weight (w_h)	Precision @10	Recall @10	Hit Rate @10	NDCG @10	MAP @10
Equal weighting	0.33	0.33	0.34	0.166	0.329	0.795	0.418	0.304
Prediction dominant	0.6	0.25	0.15	0.172	0.341	0.816	0.429	0.312
Relevance dominant	0.25	0.6	0.15	0.169	0.338	0.808	0.432	0.31
Popularity dominant	0.25	0.15	0.6	0.158	0.315	0.784	0.401	0.292
Prediction relevance balanced	0.45	0.45	0.1	0.174	0.347	0.824	0.434	0.315
Selected configuration	0.5	0.4	0.1	0.175	0.35	0.829	0.436	0.317
High prediction weighting	0.7	0.2	0.1	0.171	0.342	0.818	0.427	0.311

mance. Future research plans to study graph neural networks, selfsupervised learning, and contrastive representation learning to improve recommendation under conditions of sparsity. Future work will validate the framework through online A/B testing, clickthrough and conversion analysis, real-time deployment, inference time and memory evaluation, and large-scale implementation in practical e-commerce environments.

Declarations

Data Availability: https://www.kaggle.com/datasets/retailrocket/ecommerce-dataset

Conflicts of Interest / Competing Interests: The authors declare that they have no known competing financial or non-financial interests that could have influenced the work reported in this paper.

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

Author contributions

- Pu Lai: Conceptualization, Methodology, Software, Data Curation, Formal Analysis, Writing - Original Draft, Correspondence
- Dianxiong Ma: Validation, Investigation, Visualization, Writing - Review & Editing

All authors have read and approved the final manuscript.

Ethical approval

This study does not involve human participants, animals, or sensitive personal data. Therefore, ethical approval was not required.

Consent to Participate: Not applicable.

Consent to Publication: Not applicable.

Code availability

The code developed for data preprocessing, feature extraction (Adaptive Weighted Holoentropy), and deep learning models (CNN-BiLSTM) is available from the corresponding author upon reasonable request.

Acknowledgement

The authors would like to thank Henan Polytechnic Institute and UCSI University for their academic support. Appreciation is also extended to the providers of the RetailRocket dataset for making the data publicly accessible for research purposes.

References

[1] A. R. Onteddu, R. R. Bandhela, and R. R. Kundavaram, (2024) *"Enhancing e-commerce product recommendations through data engineering and machine learning"* **Economic Sciences *20***(1): 171–183. DOI: 10.69889/vqgzw857.

[2] A. Ahad, R. Mohaimin, M. N. S. Rabbi, et al., (2025) *"AI-based product clustering for e-commerce platforms: enhancing navigation and user personalization"* **International Journal of Environmental Sciences *11***: 156–171. URL: https://theaspd.com/index.php.

[3] A. Alqhatani and S. B. Khan, (2024) *"IoT-driven hybrid deep collaborative transformer with federated learning for personalized e-commerce recommendations: an optimized approach"* **Scalable Computing: Practice and Experience *25***(5): 3408–3426. DOI: 10.12694/scpe.v25i5.3007.

Table 15. Comparison of Recommendation Performance with Baseline Recommendation Models

Method	Precision@10	Recall@10	Hit Rate@10	NDCG@10	MAP@10
Popularity Ranking	0.086	0.172	0.512	0.184	0.151
Item-KNN	0.112	0.214	0.596	0.256	0.208
Session-KNN	0.128	0.247	0.641	0.301	0.259
Matrix Factorization (MF)	0.141	0.266	0.684	0.337	0.281
GRU4Rec	0.156	0.298	0.742	0.387	0.302
Caser	0.164	0.314	0.769	0.405	0.309
SASRec	0.169	0.331	0.801	0.421	0.315
BERT4Rec	0.172	0.34	0.817	0.429	0.316
Proposed ECNN + BiLSTM	0.175	0.35	0.829	0.436	0.317

Table 16. Performance Comparison of Different Activation Functions

Activation Function	Accuracy	Precision	Recall	F1-Score	Training Loss	Convergence Epoch	Training Time (s)
ReLU	0.9685	0.9691	0.9676	0.9683	0.094	22	4.08
GELU	0.9736	0.9741	0.9732	0.9736	0.081	20	4.15
Swish	0.9768	0.9772	0.9765	0.9768	0.072	18	4.19
Scaled Swish (Proposed)	0.9794	0.9801	0.9794	0.9794	0.064	16	4.20

[4] A. Bahi, I. Gasmi, S. Bentrad, and R. Khantouchi, (2025) *"Enhancing recommendation diversity in e-commerce using Siamese network and cluster-based technique"* **Bulletin of Electrical Engineering and Informatics** ***14***(2): 1223–1230. DOI: 10.11591/eei.v14i2.8572.

[5] M. Chowdhary and P. Chahal, (2025) *"Online review-based product recommendation using Mellivora Heron optimization-based BiLSTM model in e-commerce systems"* **Web Intelligence** ***23***(2): DOI: 10.1177/24056456241297357.

[6] W. Cui and H.-K. Lam, (2026) *"Fuzzy-enhanced variable weight graph convolutional networks for recommender systems"* **Computers and Electrical Engineering** ***131***: 110970. DOI: 10.1016/j.compeleceng.2026.110970.

[7] S. Dhawan, K. Singh, and M. Yadav, (2025) *"Hybrid deep learning recommendation system for accurate movie and product review predictions"* **Scalable Computing: Practice and Experience** ***26***(4): 1902–1918. DOI: 10.12694/scpe.v26i4.4449.

[8] N. Hassan, M. Abdelraouf, and D. El-Shihy, (2025) *"The moderating role of personalized recommendations in the trust-satisfaction-loyalty relationship: an empirical study of AI-driven e-commerce"* **Future Business Journal** ***11***(1): 66. DOI: 10.1186/s43093-025-00476-z.

[9] J. Huang, (2025) *"Optimization and innovation of AI-based e-commerce platform recommendation system"* **Journal of Computer, Signal, and System Research** ***2***(6): 66–73. DOI: 10.71222/7bb7rx18.

[10] C. Jiang, (2025) *"Research on sales forecasting and consumption recommendation system of e-commerce agricultural products based on LSTM model"* **GeoJournal** ***90***(3): 93. DOI: 10.1007/s10708-025-11342-4.

[11] N. Karne, (2021) *"Customizing enterprise interactions: an AI-driven approach to customer recommendation systems"* **International Journal of Scientific Research in Science and Technology** ***8***(1): 325–336. DOI: 10.32628/IJSRST52310110.

[12] P. Keikhosrokiani and M. F. Goh, (2024) *"A hybrid recommender system for health supplement e-commerce based on customer data implicit ratings"* **Multimedia Tools and Applications** ***83***(15): 45315–45344. DOI: 10.1007/s11042-023-17321-6.

[13] H. U. Khan, A. Naz, F. K. Alarfaj, and N. Almusallam, (2025) *"A transformer-based architecture for collaborative*

filtering modeling in personalized recommender systems" **Scientific Reports *15***(1): 24503. DOI: 10.1038/s41598-025-08931-1.

[14] E. S. P. Krishna, T. B. Ramu, R. K. Chaitanya, et al., (2025) *"Enhancing e-commerce recommendations with sentiment analysis using MLA-EDTCNet and collaborative filtering"* **Scientific Reports *15***(1): 6739. DOI: 10.1038/s41598-025-91275-7.

[15] M. Sunita and S. V. Mallapur, (2024) *"Optimal detection of border gateway protocol anomalies with extensive feature set"* **Multimedia Tools and Applications *83***(17): 50893–50919. DOI: 10.1007/s11042-023-17135-6.