© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1716019

E-Commerce Cart Prediction Using Multi-Task Learning with Bi-LSTM and Graph Neural Networks

MANIMOLI A T[1], JOSHI M[2], AROKIA NISHA S[3], KAYALVIZHI S[4]
[1, 2, 3]Department of Computer Science and Engineering, Velammal Engineering College, Chennai, Tamil Nadu, India
[4] Assistant Professor, Dept. of CSE, Velammal Engineering College, Chennai, Tamil Nadu, India

Abstract- Shopping cart abandonment is a persistent and financially significant challenge in digital commerce, with reported global rates exceeding 68% and cumulative revenue losses running into the hundreds of billions annually. To tackle this, the present work proposes a unified multi-task learning framework capable of concurrently solving three interrelated problems: identifying abandonment risk, recommending the next likely product, and forecasting the expected transaction value. Instead of deploying isolated models for each objective, the system derives a shared latent representation from user clickstream logs and behavioural interaction histories. Temporal dynamics within sessions are encoded using a Bidirectional Long Short-Term Memory (Bi-LSTM) network, while structural dependencies across products and categories are modelled via a Graph Neural Network (GNN). The framework operates with low latency, supporting real-time deployment of targeted retention strategies before users disengage. To enhance model transparency, SHAP and LIME attribution methods are embedded as interpretability tools accessible to operational stakeholders. Empirical validation on two publicly available datasets demonstrates consistent improvements in both predictive accuracy and estimated conversion lift over single-task competitors.

Index Terms- E-Commerce; Cart Abandonment; Multi-Task Learning; Bi-LSTM; Graph Neural Networks; Explainable AI; Clickstream Analysis; Customer Behaviour Analytics

I. INTRODUCTION

Cart abandonment is one of the most pervasive and costly issues confronting e-commerce operators today. Longitudinal studies consistently document that over two thirds of online shoppers who add items to their cart ultimately leave the site without completing a purchase, a phenomenon that transcends product type, market segment, and platform design.

Despite the scale of this problem, the predominant industry response remains reactive—chiefly post-session email follow-ups dispatched hours after the user has already left—by which point the window of purchase intent has typically closed.

Proactive retention necessitates in-session inference— predicting the likelihood of abandonment, surfacing relevant product alternatives, and estimating probable expenditure while the user is still engaged. Existing commercial solutions tend to address these objectives independently, deploying separate pipelines that fail to leverage the natural correlations among them. A shopper exhibiting hesitation signals simultaneously reveals information pertinent to all three tasks; a joint modelling approach that capitalises on these overlapping signals should outperform siloed alternatives on every dimension.

The architecture proposed here fuses a Bidirectional LSTM for capturing temporal patterns in clickstream sequences with a GNN for encoding structural relationships in the product catalogue, routing the merged representation to three specialised prediction heads. Post-hoc interpretability is provided through SHAP for global feature attribution and LIME for session-level explanations, both implemented as non-invasive inference wrappers. Comparative experiments are conducted on two benchmark e-commerce datasets against a range of competitive baseline methods.

II. RELATED WORK

The field of user behaviour modelling in digital commerce has undergone substantial methodological evolution. Early approaches relied on handcrafted session-level attributes fed into logistic regression, Bayesian classifiers, or tree-based ensembles. While these models were transparent and efficient, they reduced session dynamics to aggregate statistics, discarding the sequential ordering of user actions and thus forfeiting much of the temporal signal relevant to predicting purchasing intent.

Choi et al. [1] demonstrated that LSTM-based recurrent architectures markedly surpass conventional classifiers for sequential user interaction modelling by preserving event order within sessions, showing that both interaction type and timing encode meaningful predictive cues that aggregate features erase. Building on this foundation, Zhang et al. [2] systematically benchmarked machine learning techniques for cart abandonment prediction, establishing reference metrics widely cited in later research.

Graph-based representations gained traction when researchers recognised that sequential models are inherently limited in their ability to capture structural co-occurrence relationships across product catalogues. Li and Zhao [3] demonstrated that applying GNNs to product-category graphs yields richer item embeddings and measurably better recommendation quality. Their work, however, was scoped exclusively to recommendation tasks and did not explore integration with purchase probability or abandonment prediction objectives.

Kumar et al. [4] offered compelling evidence for multi-task learning in e-commerce contexts, documenting statistically significant accuracy gains in abandonment detection, purchase likelihood estimation, and session duration prediction under joint optimisation relative to independent training regimes.

Shared representations served as an implicit regulariser, curbing overfitting to individual task characteristics. Wang et al. [5] argued in parallel that predictive models, regardless of their accuracy, will remain operationally marginalised unless their outputs can be explained in terms understandable to business decision-makers responsible for intervention design.

Subsequent work has refined individual components of this landscape. Singh and Gupta [6] investigated deep learning approaches for revenue forecasting in online retail, highlighting residual prediction gaps when only product-level attributes are considered. Chen et al. [7] addressed irregular temporal spacing in clickstream sequences via time-aware gating mechanisms, a refinement that standard recurrent formulations do not accommodate. Patel et al. [8] validated the complementarity of Bi-LSTM and GNN representations for e-commerce prediction tasks but stopped short of unifying these components with multi-task learning objectives and integrated explainability— the precise contribution of the current paper.

III. PROPOSED METHODOLOGY

A. Data Representation and Feature Engineering

The system ingests raw timestamped event logs in which each record documents a discrete user action—ranging from page visits, hover events, and dwell periods to add-to-cart operations, wishlist modifications, and progression through checkout stages. Session boundaries are defined by a 30-minute inactivity gap, consistent with standard web analytics practice. From each session, two complementary data structures are constructed: a temporally ordered event sequence fed into the Bi-LSTM, and an undirected bipartite graph linking user nodes to product nodes via edges weighted by dwell duration and interaction count, which serves as GNN input. Supplementary user-level features—including recency, session frequency, mean dwell time, scroll behaviour, and session duration—are concatenated directly at the task-head input stage.

B. Model Architecture

Processing flows through four consecutive stages: input encoding, dual-path feature extraction, representation merging, and task-specific inference. Within the sequential pathway, the Bi-LSTM reads each event series in both forward and backward directions, producing a joint hidden state that captures how purchase intent develops and fluctuates throughout a session. Inter-event time gaps are incorporated as continuous modulation signals applied at the gating level, enabling the model to assign stronger weight to recent hesitation indicators relative to earlier exploratory behaviour.

© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1716019

Within the graph pathway, the GNN performs two successive rounds of message passing over the product-category graph. At each propagation step, node representations are updated by aggregating weighted contributions from connected neighbours, enabling the model to encode multi-hop relational context. The final node embeddings encapsulate catalogue structure that pure sequence-based methods inherently cannot represent.

A Feature Fusion Layer merges the Bi-LSTM and GNN outputs via concatenation and applies a shared dense projection before routing the result to three task-specific heads. The first head performs binary classification for abandonment risk using

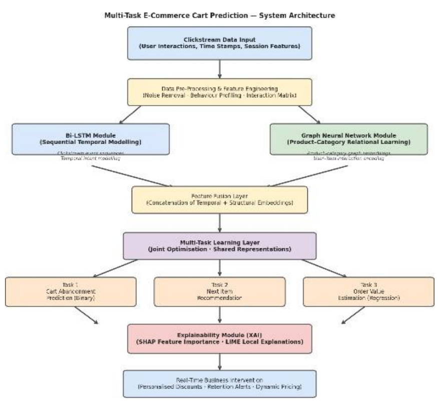


Fig. 1. Proposed System Architecture for Multi-Task E-Commerce Cart Prediction

cross-entropy loss. The second head addresses next-item recommendation using softmax cross-entropy over the full product catalogue. The third head estimates order value through mean squared error regression. Training optimises a composite objective formed by a weighted combination of all three losses, with task weights determined via Bayesian hyperparameter optimisation.

C. Explainability

Interpretability is provided through two complementary post- hoc methods applied without modifying the underlying model. SHAP (SHapley Additive exPlanations) computes global feature contribution rankings, exposing which behavioural signals most reliably predict abandonment at population scale. LIME (Local Interpretable Model-agnostic Explanations) generates instance-level explanations, enabling support and operations staff to interrogate the reasoning behind any specific risk prediction. Both tools function as inference-time wrappers and place no restrictions on the training procedure.

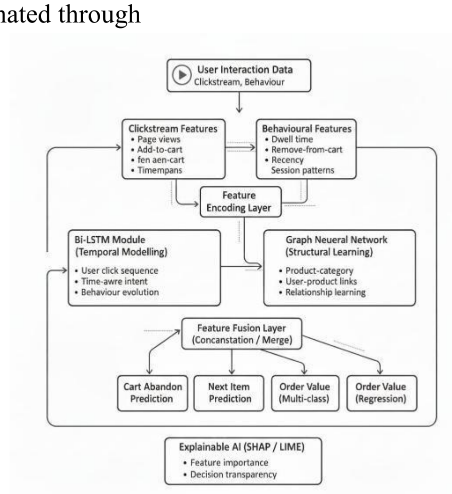
IV. EXPERIMENTAL SETUP

A. Datasets

Two publicly available e-commerce benchmarks are used. The UCI Online Retail Dataset contains roughly 500,000 transaction records from a UK-based online merchant spanning two calendar years. The Kaggle E-Commerce Behaviour Dataset captures fine-grained clickstream data—encompassing product browsing, cart interactions, and completed purchases— across more than four million user sessions. Both datasets are split chronologically into training, validation, and test portions at a 70/15/15 ratio to avoid temporal data leakage. Pre- processing steps include session boundary detection, removal of likely bot sessions identified by abnormal click rate patterns, and median-value imputation for missing dwell-time fields.

B. Evaluation Metrics

Abandonment classification performance is reported via accuracy, F1-score, AUC-ROC, and precision-recall analysis to address label imbalance concerns. Recommendation effectiveness is quantified using Hit Rate at K and Normalised Discounted Cumulative Gain at K, evaluated at both K=5 and K=10. Transaction value prediction is assessed with RMSE and MAPE. Downstream business value is estimated through



© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1716019

Fig. 2. Comparative Evaluation Results Across Prediction Tasks

simulated conversion uplift, which compares the outcomes of model-guided interventions against a random-trigger control condition.

V. RESULT AND DISCUSSION

Experimental findings substantiate the hypothesis that multi- task joint optimisation produces superior outcomes relative to any isolated task configuration. On the UCI dataset, the Bi- LSTM module in isolation attains an abandonment AUC-ROC of 0.887, already surpassing a standalone LSTM (0.831) and a gradient-boosted classifier trained on session aggregates (0.794). Incorporating GNN embeddings through the fusion layer elevates the AUC-ROC to 0.913, affirming that catalogue relational structure conveys predictive information that sequential representations alone cannot recover.

On the recommendation task, the complete multi-task model achieves HR@10 of 0.748 and NDCG@10 of 0.621, outperforming the graph-only baseline by 8.3 and 7.1 percentage points respectively. Order value prediction yields an RMSE of 14.72 on UCI data, representing a 19.4% improvement over the dedicated single-task regression model. These gains are principally attributable to cross-task representation sharing: the shared encoder learns broadly applicable engagement features that benefit all three prediction heads simultaneously while mitigating task-specific overfitting.

SHAP analysis identifies session recency, the ratio of cart additions to product views, and product category diversity as the most influential predictors of abandonment risk—findings well-aligned with domain understanding. LIME-generated explanations at the session level reveal that high-risk flagged sessions exhibit coherent and recognisable behavioural profiles rather than artefacts of statistical noise, bolstering confidence in the system's suitability for live production deployment.

VI. PRACTICAL IMPLICATION

Ongoing session-level scoring gives e-commerce platforms an actionable window to intervene while user intent remains present. Upon flagging a session as high-risk, the system can instantly deliver a personalised incentive, prompt a product comparison view, or escalate to a live support agent—all within a few seconds of the triggering behaviour, well ahead of the user's departure. In tandem, the recommendation module serves the most purchase-relevant items to hesitant users, reinforcing the probability of conversion.

The interpretability layer converts model outputs into strategic insight usable beyond the immediate prediction task. Category managers can leverage SHAP attribution visualisations to identify product- or seasonal-level abandonment patterns and calibrate inventory and promotional decisions accordingly, extending the platform's return on the system well past real-time alert generation. Built entirely on open-source components and open benchmark data, the framework carries substantially lower acquisition and maintenance costs than proprietary solutions, and its methodology can be reproduced by teams lacking dedicated commercial data infrastructure.

VII. FUTURE WORK

Several avenues have been identified for advancing this work. Integrating cross-channel behavioural data is a near-term priority: augmenting session representations with signals from mobile apps, email interactions, and social shopping environments should alleviate cold-start limitations for users with sparse activity on any individual platform while enriching the overall intent profile.

Low-latency streaming inference will be pursued through Apache Kafka integration coupled with an online-learning extension of the Bi-LSTM that incrementally updates model parameters in response to distributional shifts driven by seasonal trends, promotional events, and catalogue changes, thereby reducing reliance on scheduled full retraining cycles. Federated learning represents an additional direction, enabling joint training across multiple retailer datasets without centralising sensitive user records, which aligns with tightening data privacy regulations.
Accessibility evaluations with participants who have

visual or motor disabilities will determine whether intervention interfaces produced by the system serve diverse user populations effectively. Objective usability metrics—task completion time, error frequency, and self-reported satisfaction—will be gathered and used to guide iterative interface refinement.

The model's resilience under challenging operating conditions—including sparse interaction histories, rapid catalogue expansion, and synthetic bot traffic injection—will be systematically evaluated. Dropout regularisation, attention- based masking, and curriculum learning schedules will be tested as candidate mitigation strategies. Controlled offline benchmarks will be supplemented with live A/B experiments on partner retailer platforms to generate real-world evidence of business impact.

Lastly, multimodal feature enrichment will be explored by incorporating product image representations extracted by a convolutional encoder and semantic embeddings derived from textual product descriptions. Integrating visual and linguistic information alongside clickstream data is expected to enhance recommendation variety and reduce the systematic marginalisation of niche products that frequently affects purely collaborative approaches.

CONCLUSION

This paper has presented a multi-task learning architecture that combines Bi-LSTM sequential modelling with GNN-based relational encoding to simultaneously address cart abandonment prediction, next-product recommendation, and purchase value estimation within a single end-to-end trainable system. By encoding both the temporal evolution of user sessions and the structural topology of the product catalogue through a unified representation, the framework overcomes core limitations of prior unimodal approaches. Integrated SHAP and LIME components translate model decisions into operationally meaningful explanations, supporting practical deployment by business and analytics teams. Benchmark results confirm that the joint multi-task system outperforms all single-task baselines across every evaluation dimension, establishing multi-task learning with explainability as a compelling paradigm for next-generation e-commerce predictive analytics.

REFERENCES

[1] J. Choi, H. Kim, and S. Lee, "Sequential modeling for e- commerce clickstreams with RNNs," in Proc. IEEE Int. Conf. Data Mining Workshops (ICDMW), 2017, pp. 245-252.

[2] Y. Zhang, L. Wang, and Q. Chen, "Predicting cart abandonment using machine learning techniques," IEEE Access, vol. 6, pp. 65789-65798, 2018.

[3] X. Li and Y. Zhao, "Graph-based recommendations in online retail," in Proc. ACM Conf. Recommender Systems (RecSys), 2019, pp. 312-320.

[4] R. Kumar, P. Sharma, and A. Verma, "Multi-task learning for customer behavior prediction in e-commerce," IEEE Trans. Knowledge and Data Engineering, vol. 32, no. 11, pp. 2145- 2158, 2020.

[5] T. Wang, M. Liu, and Z. Sun, "Explainable AI in e- commerce analytics," IEEE Intelligent Systems, vol. 35, no. 4,
pp. 42-50, 2020.

[6] A. Singh and R. Gupta, "Deep learning approaches for purchase forecasting in online retail," IEEE Access, vol. 9, pp. 112345-112356, 2021.

[7] Y. Chen, K. Huang, and L. Zhou, "Time-aware recurrent models for clickstream data analysis," in Proc. IEEE Int. Conf. Big Data, 2022, pp. 1785-1792.

[8] D. Patel, S. Mehta, and N. Rao, "Hybrid Bi-LSTM and graph neural network models for e-commerce prediction," IEEE Trans. Neural Networks and Learning Systems, vol. 33, no. 8, pp. 3672-3684, 2022.