Optimizing Omni-Channel Marketing Campaigns: A Reinforcement Learning-Based Intelligent Allocation System

Zhuoxi chen[1] and Baitong Zhong[2*]

[1] *Hunan Vocational College of Electronic Technology, School of Humanities and Education , changsha Hunan province, China, 410000*

[2] *College of Information Engineering, Hunan Mechanical & Electrical Polytechnic, Hunan Province, China, 410000*

[*] *Corresponding author. E-mail: bvbv1234weih@163.com, baitongzhong78@outlook.com*

Received April 18, 2026; accepted June 16, 2026

The rapid expansion of digital platforms together with the intricate nature of customer engagement across multiple channels which include email and social media and search engines and display advertising have made it more difficult to optimize omni-channel marketing campaigns. Traditional methods operated with fixed models which could not track ongoing user behavior and their changing choices rendering those methods ineffective. The research presented a new hybrid framework which combined Long Short-Term Memory (LSTM) and Deep Q-Network (DQN) based Reinforcement Learning framework to increase campaign success while delivering better return on investment. The model applied the Marketing Attribution Path Dataset to employ an LSTM in learning customer interaction patterns. The DQN framework used the predictions to train its agent to develop marketing channel selection policies which used conversion rate and click-through rate and revenue as their reward system. Additionally, the proposed framework effectively integrates sequential prediction with adaptive decision optimization, improving performance in dynamic marketing environments. The results showed that the proposed model achieved better results because it achieved an accuracy of 0.9745 and a precision of 0.9697 and a recall of 0.9796 and an F1-score of 0.9746 which exceeded the performance of existing methods. The reinforcement learning agent demonstrated its learning capacity by successfully identifying the channels that produced optimal performance. The study showed that the LSTM-DQN framework which researchers developed as their solution for intelligent omni-channel marketing optimization achieved both scalable and adaptive capabilities while delivering major improvements in customer engagement and campaign efficiency.

Keywords: Omni-channel marketing, LSTM, Deep Q-Network, Reinforcement learning, Campaign optimization, Customer journey modeling.

http://dx.doi.org/10.6180/jase.202611_34.016

1. Introduction

Digital platforms have rapidly expanded, leading businesses to adopt omni-channel marketing strategies across email, social media, search engines, and display advertising [1, 2]. Deep learning models like LSTM capture temporal customer behavior, while reinforcement learning enables adaptive decision-making [3–5]. Traditional marketing methods such as logistic regression, Random Forest, SVM, and Gradient Boosting are widely used for prediction and segmentation [6–8]. However, these approaches are static and fail to model sequential customer behavior and real-time optimization [9, 10]. This study proposes a hybrid LSTM–DQN framework for intelligent omni-channel marketing optimization. The LSTM model predicts cus-

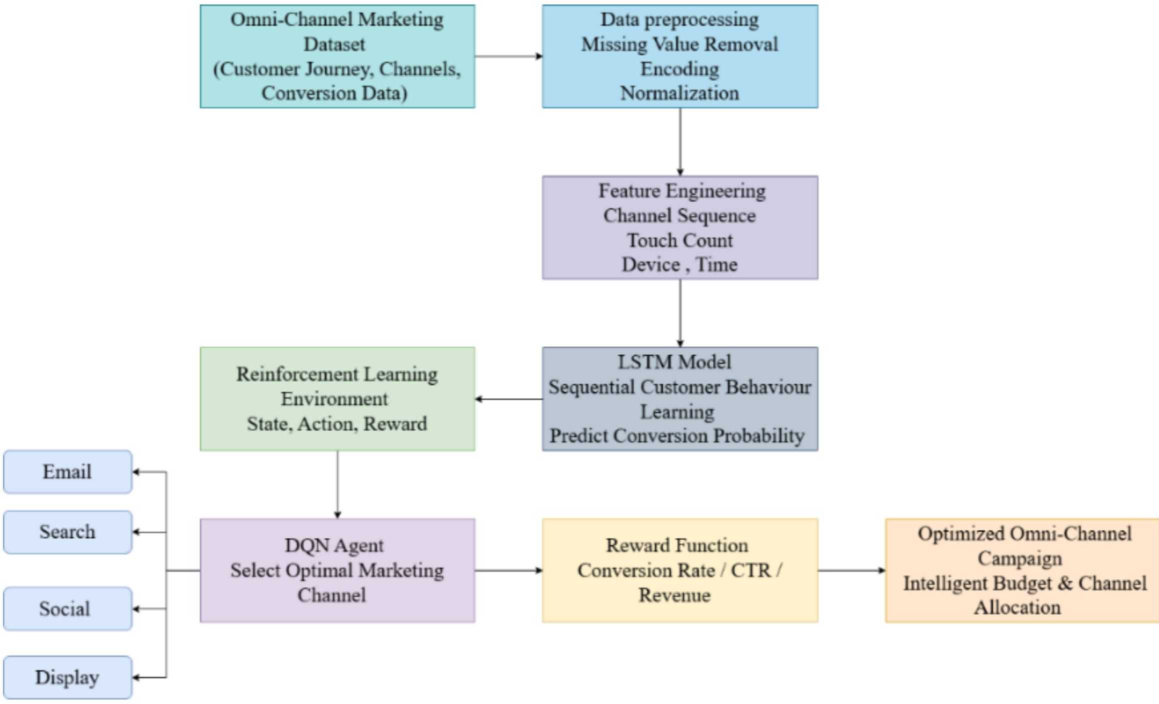

Fig. 1. LSTM–DRL Omni-Channel Marketing Optimization Workflow

tomer conversion probability from sequential interactions, while the DQN agent learns optimal marketing channel selection through reward-based learning. The proposed framework combines temporal behavior modelling with adaptive decision-making for real-time campaign optimization. Experimental results show that the model outperforms conventional machine learning and deep learning approaches across multiple evaluation metrics [11].

2. Materials and methods

Figure 1 shows the workflow of the proposed framework, starting with data collection and feature engineering (channel sequence, touch count, device type, interaction time). These features are then processed by the LSTM model to learn customer behavior patterns and predict conversion probabilities.

The LSTM-predicted conversion probability serves as the state for the reinforcement learning model, where the DQN agent selects the optimal marketing channel. The reward function evaluates performance using conversion rate, click-through rate, and revenue to enable adaptive optimization. Feature engineering converts raw interaction data into structured inputs, allowing effective modeling of sequential customer journeys over time. It is calculated as Eq. (1):

$$f_t = \sigma\left(W_f\left[h_{t-1}, x_t\right] + b_f\right) \tag{1}$$

The forget gate produces its output through equation f_t which uses the previous time step hidden state, h_{t-1} and current time step input, x_t along with weight matrix, W_f and b_f The input gate determines which new information should be stored in the memory cell as represented in the Eq. (2):

$$i_t = \sigma\left(W_i\left[h_{t-1}, x_t\right] + b_i\right) \tag{2}$$

The candidate cell state $\widetilde{C}_t$ in the LSTM network shows the new information which the current input x_t and the previous hidden state h_{t-1} bring. Then the memory cell state is updated by combining the previous memory and the new candidate information as expressed in the Eq. (3):

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \widetilde{C}_t \tag{3}$$

Figure 2 shows the LSTM-based customer behavior analysis, where sequential interaction data is used to learn temporal patterns and predict conversion probability. Lastly, the output gate is calculated to decide which information at the current time step will be sent on to the following step as given in the Eq. (4):

$$o_t = \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right) \tag{4}$$

The hidden state represents the output at each time step, while the output gate controls the flow of information using the cell state. These gated mechanisms enable the model to

Table 1. Hyperparameter Settings for Proposed LSTM-DQN Model

Component	Hyperparameter	Value
LSTM Model	Input Features	20 features
	Hidden Units	128, 64
	Activation	ReLU
	Regularization	L2 (1e-4)
	Batch Normalization	Yes
	Dropout	0.2, 0.1
	Output Activation	Sigmoid
	Optimizer	Adam
	Learning Rate	0.001
	Loss Function	Binary Cross entropy
RL Environment	State Dimension	20
	Action Space	7 channels
	Episode Length	50 steps
	Reset Strategy	Random
DQN Agent	Discount Factor (γ)	0.95
	Initial Epsilon	1
	Minimum Epsilon	0.05
	Epsilon Decay	0.995
	Batch Size	32
	Memory Size	2000
	Episodes	100
	Target Update	Every 10 episodes

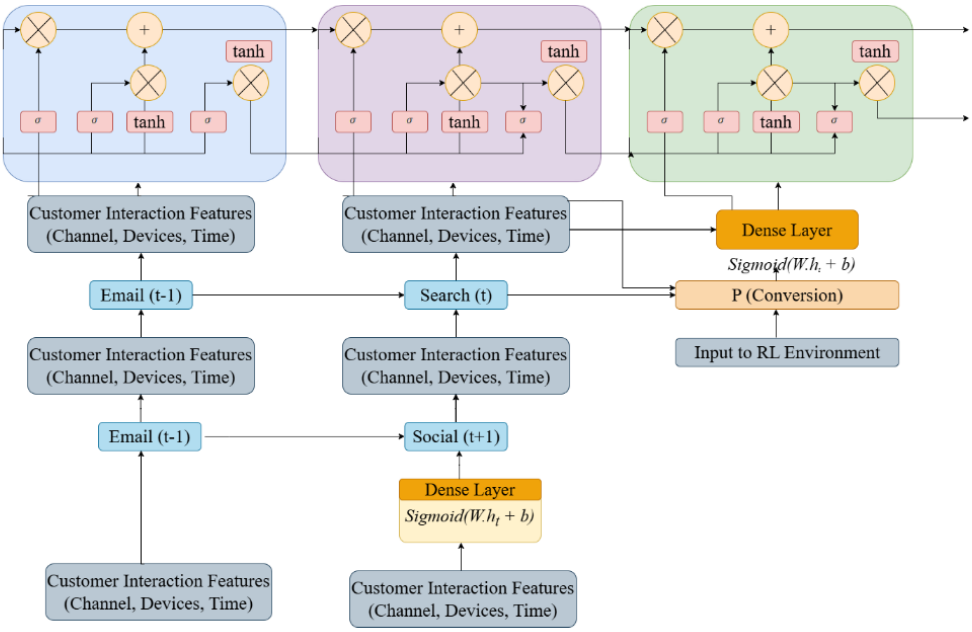

Fig. 2. Architecture of the LSTM Model for Customer Interaction Sequence Modeling

capture long-term dependencies and generate conversion predictions, which are used as state inputs for reinforcement learning-based channel optimization. The LSTM outputs, including conversion probability and hidden states, form the state representation for the DQN agent, enabling adaptive marketing channel selection. The reinforcement learning framework uses state, action, and reward components, where the state includes customer features such as channel sequence, interaction count, device type, and predicted conversion probability. The reward function evaluates the effectiveness of selected actions, assigning higher rewards to channels that result in customer engagement or conversion. The reward as formulated Eq. (5):

$$R_t = f\left(C_t, CTR_t, Rev_t\right) \tag{5}$$

Here, Rt denotes the reward, Ct the conversion outcome, CTR the click-through rate, and Revt the generated revenue. The DQN agent learns optimal marketing channel selection by maximizing rewards based on customer behavior and LSTM-predicted conversion probabilities. The Q-function evaluates the expected reward for each action under a given state Eq. (6).

$$Q(s,a) = r + \gamma \max_{a'} Q\left(s', a'\right) \tag{6}$$

The reinforcement learning agent interacts with the environment to learn optimal marketing channel selection policies that maximize campaign performance. Training is further improved using experience replay and an epsilon-greedy strategy, which balance learning stability and exploration for better channel optimization.

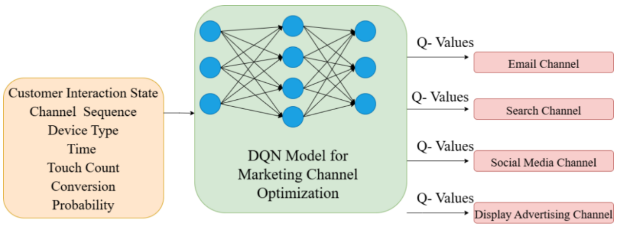

Fig. 3. DQN Architecture for Omni-Channel Marketing Decision Making

Figure 3 shows the DQN architecture for marketing channel optimization, where customer interaction features and LSTM-predicted conversion probability are used to compute Q-values for different actions The reward function as mathematically represented in the Eq. (7):

$$R_t = \alpha C_t + \beta CTR_t + \gamma Rev_t \tag{7}$$

The reward function evaluates performance based on conversion, click-through rate, and revenue, with adjustable weights for multi-objective optimization. The policy can be mathematically as indicated in the Eq. (8):

$$\pi^*(s) = arg \max_{a} Q(s,a) \tag{8}$$

The trained DQN agent selects the channel with the highest Q-value, enabling real-time adaptive campaign allocation and improved marketing performance. Table 1 summarizes the hyperparameters of the proposed LSTM–DQN framework. The LSTM model uses two hidden layers with ReLU activation and a sigmoid output, along with dropout, L2 regularization, and Adam optimization to ensure stable training and reduce overfitting.

The reinforcement learning setup defines state and action spaces based on customer features and marketing channels, with training conducted over 100 episodes using an epsilon-greedy strategy and experience replay for improved learning stability. The implementation is carried out in Python using TensorFlow and PyTorch, with standard libraries for data processing and visualization on the Marketing Attribution Path Dataset. Table 2 summarizes the dataset [12], which includes 200 customer journeys with 14 features across seven marketing channels. It contains sequential data such as timestamps, touchpoints, and device type, with targets for both binary conversion and conversion value.

The dataset is split into training (80%) and testing (20%), with preprocessing steps including handling missing values, class balancing, and feature encoding. Missing values are replaced with default values, and categorical data is converted using one-hot encoding. The framework uses binary cross-entropy loss for LSTM-based conversion prediction and temporal difference loss for the DQN model, enabling effective learning of customer behavior and optimal channel selection policies.

3. Results and discussion

The model is evaluated using Accuracy, Precision, Recall, and F1-score, achieving strong performance with balanced and closely aligned values. The results indicate reliable classification with no major bias, supported by statistical validation confirming the significance of the predictions.

Figure 4 demonstrates strong performance through high AUC values, stable accuracy improvement, and low error rates. Table **??** confirms that removing key components like LSTM, DQN, or feature engineering significantly reduces performance, highlighting their importance.

Figure 5 shows cumulative reward growth, where the model improves over time as it learns optimal strategies.

Figure 6 highlights learning progress, where rewards stabilize after initial fluctuations, indicating a reliable policy.

Figure 7 shows epsilon decay, reflecting a gradual shift from exploration to exploitation.

Table **??** confirms improved learning through higher rewards, strong decision-making, and stable policy behavior. Figure 8 demonstrates that channel performance varies, with some channels achieving higher rewards while others show lower and more fluctuating performance.

Table 5 compares marketing channels, showing that Social, Email, and Referral achieve high conversion rates and

Table 2. Summary of Dataset Characteristics

Component	Details
Total Samples	200 customer journeys
Features	14 columns
Channels	7 (affiliate, display, email, referral, search, social, video)
Key Features	`channels_sequence`, `touch_timestamps`, `touch_count`, `device_types`
Target Variables	conversion (binary), `conversion_value` (continuous)
Training Set	80%
Testing Set	20%
Missing Values	`conversion_value` → 0; `geo_country`, `geo_city` → "Unknown"

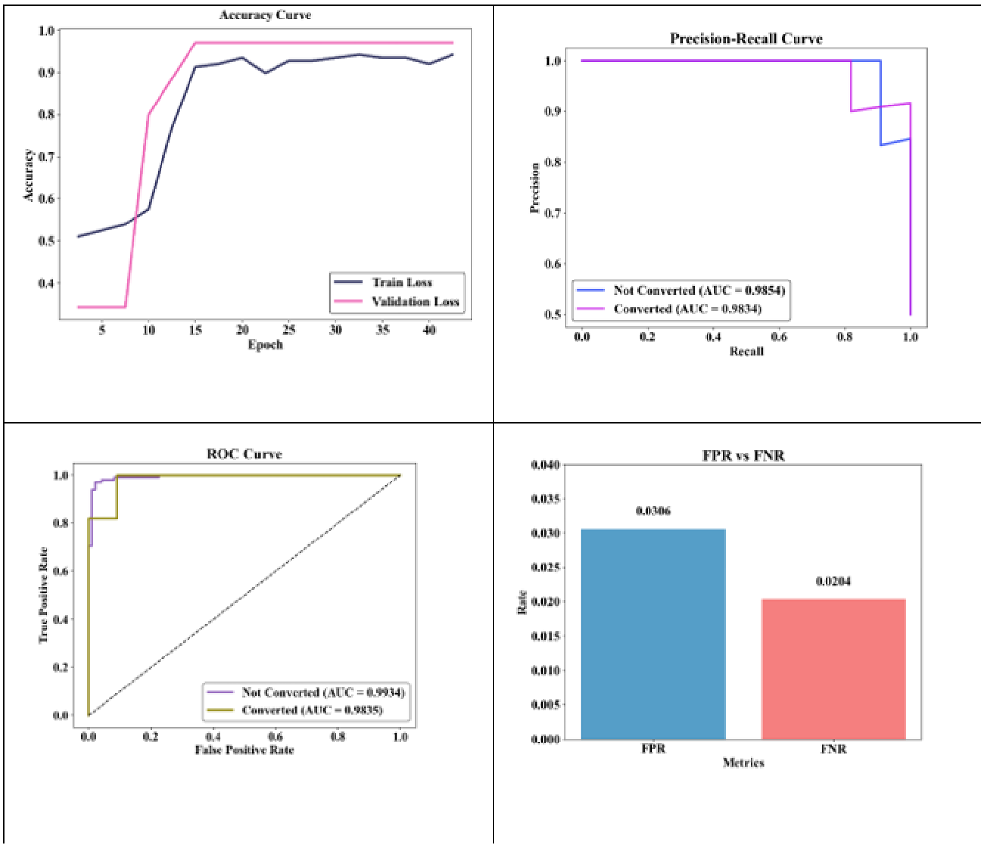

Fig. 4. Model Performance Evaluation Using Multiple Metrics

revenue, with Referral having the highest reward score. Video and Affiliate perform poorly, while Search and Display provide steady but average results. Table 5 compares marketing channels, showing that Social, Email, and Referral achieve high conversion rates and revenue, with Referral having the highest reward score. Video and Affiliate perform poorly, while Search and Display provide steady but average results.

Figure 9 shows how loss curves behave during training and validation tests which lasted for 45 epochs. The model starts to make incorrect predictions because both training and validation losses begin at high levels which reach approximately 0.7. The loss decreases significantly during the first 15 epochs because the model learns and adapts to the data at a fast rate.

Figure 10 shows a confusion matrix indicating a highly accurate classification, with most predictions correctly matching the actual labels.

Table 6 shows the LSTM–DQN model outperforms existing methods (QMC+RL, DDPG) with high accuracy, precision, recall, and F1-score.

Overall, the hybrid framework effectively combines tem-

Table 3. Ablation Study of Model Components in Reinforcement Learning

Model Variant	Cumulative Reward	Avg Reward	Success Rate	Policy Entropy	Q-Mean
Only DQN (No LSTM + Basic Features)	812.67	8.12	0.623	0.832	2.76
Without DQN (Random Policy)	642.18	6.42	0.512	1.245	1.12
Without Feature Engg	954.87	9.54	0.682	0.612	3.21
Without LSTM	1012.32	10.12	0.701	0.589	3.85
Full Model (Proposed)	1235.45	12.35	0.784	0.6729	4.79

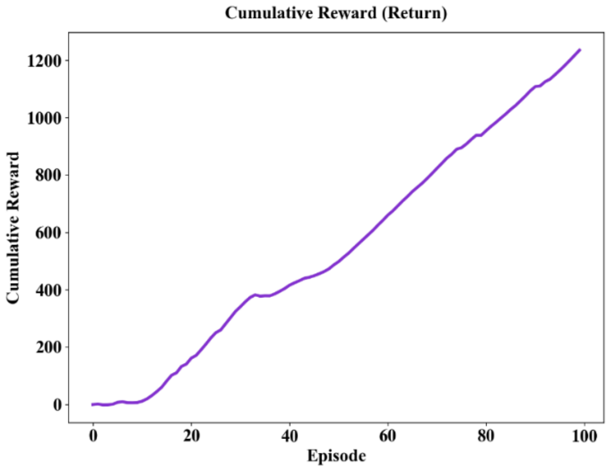

Fig. 5. Cumulative Reward Growth Over Training Episodes

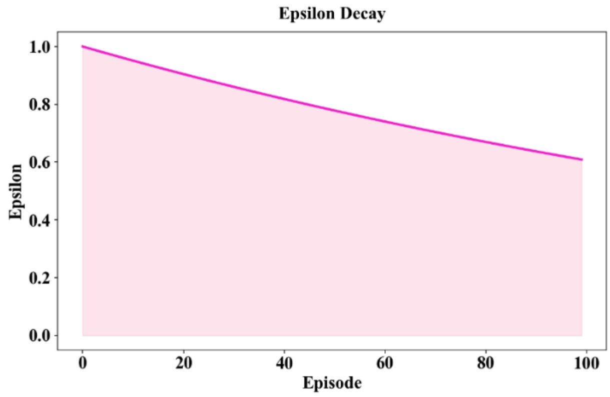

Fig. 7. Epsilon Decay Strategy Over Training Episodes

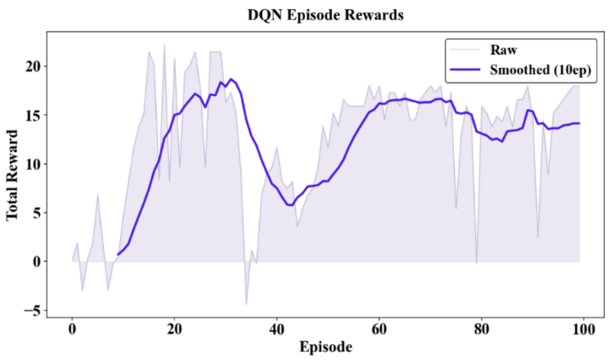

Fig. 6. DQN Training Performance. Episode Rewards with Smoothed Trend

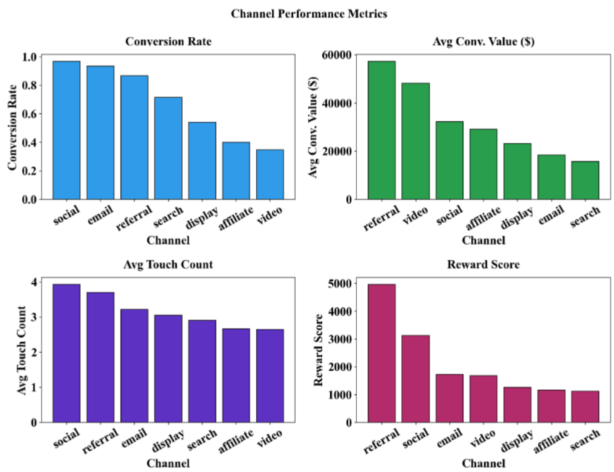

Fig. 8. Channel-wise Performance Comparison Across Key RL Metrics

poral learning and adaptive decision-making, enabling accurate conversion prediction and optimal marketing channel selection. Traditional machine learning and deep learning models fail to capture dynamic customer behavior and adaptive decision-making effectively. The proposed LSTM–DQN framework improves performance by combining temporal pattern learning with reinforcement learning-based optimization. The removal of LSTM reduces the model's ability to capture sequential patterns, leading to poor prediction and unstable learning. Channel analysis shows that Referral, Social, and Email perform better, while Affiliate and Video require improvement. Despite strong performance, the model has limitations such as a small dataset, fixed reward design, and lack of external factors like seasonality and market trends. Future improvements include using larger datasets, incorporating real-world factors, and enhancing reward functions with long-term metrics like customer lifetime value.

Table 4. Detailed Performance Metrics of the Reinforcement Learning Model

Metric	Value
Cumulative Reward	1235.446
Average Reward	12.3545
Initial Reward	7.0628
Final Reward	14.8276
Reward Improvement	7.7648
Policy Entropy	0.6729
Action Diversity	1.7452
Success Rate	0.784
Average Regret	0 as T increases
Q-Value Mean	4.7949
Q-Value Max	10.9203
Q-Value Min	-3.3419
Reward (email)	0.4735
Reward (search)	0.2492
Reward (social)	0.0334
Reward (display)	0.4932
Reward (affiliate)	0.0771
Reward (video)	0.3194

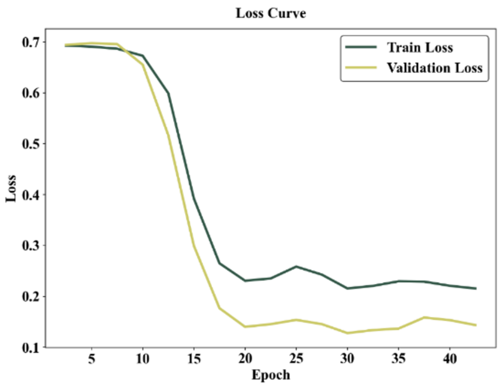

Fig. 9. Training and Validation Loss Across Epochs

4. Conclusion

The research proposes an advanced framework that integrates LSTM and DQN to optimize omni-channel marketing campaigns. The model effectively captures temporal customer behavior and selects optimal marketing channels, improving conversion prediction and campaign performance. The framework demonstrates strong stability and performance through high evaluation metrics and consistent learning patterns. Its key contribution lies in combining temporal pattern recognition with adaptive decision-making, overcoming the limitations of traditional static models. This hybrid approach enables intelligent, data-driven marketing optimization, resulting in improved customer engagement and better campaign outcomes. The

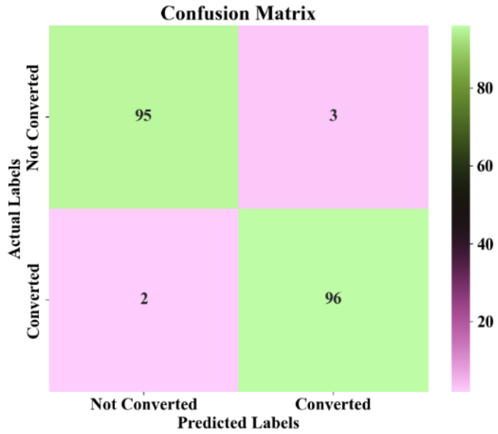

Fig. 10. Confusion Matrix for Binary Classification

framework provides actionable insights by identifying high-performing channels and enabling efficient resource allocation. It is well-suited for modern digital marketing environments that require continuous adaptation across multiple channels. The model also supports scalability, allowing integration of various marketing channels such as email, search, social media, affiliate, display, referral, and video platforms. Future work should focus on using larger and more diverse datasets, incorporating contextual factors, and improving reward mechanisms with long-term metrics like customer lifetime value. Additionally, including external factors such as seasonal trends, market dynamics, and changing customer preferences can further enhance adaptability. Overall, the framework offers a scalable and effective solution for real-world omni-channel marketing optimization.

Declaration

Data availability

https://gomask.ai/marketplace/datasets/marketing-attribution-path-dataset

Conflicts of interest

The authors declare no conflicts of interest regarding this study.

Funding statement

This research received no external funding.

Acknowledgement

Not Applicable

Table 5. Channel-wise Performance and Revenue Contribution Analysis

Channel	N_Journeys	Conv_Rate	Avg_Value	Avg_Touches	Total_Revenue	Reward_Score
Referral	30	0.867	57,270.00	3.70	1,718,099.86	4,963.40
Social	60	0.967	32,310.78	3.93	1,938,646.76	3,123.38
Email	105	0.933	18,494.26	3.22	1,941,897.78	1,726.13
Video	23	0.348	48,284.37	2.65	1,110,540.48	1,679.46
Display	83	0.542	23,218.39	3.05	1,927,126.29	1,258.83
Affiliate	45	0.400	29,183.79	2.67	1,313,270.40	1,167.35
Search	123	0.715	15,762.26	2.91	1,938,757.83	1,127.71

Table 6. Comparison of Proposed Model with Existing Methods

Model	Accuracy	Precision	Recall	F1-Score
QMC + RL [13]	0.9564	0.9482	0.9625	0.9553
DDPG + Markov decision making [14]	0.9417	0.9326	0.9504	0.9415
Proposed LSTM + DQN	**0.9745**	**0.9697**	**0.9796**	**0.9746**

Author contribution

Zhuoxi Chen contributed to conceptualization, methodology, data analysis, and manuscript writing. Baitong Zhong supervised the study, reviewed, and edited the manuscript.

Ethical approval

This study does not involve human participants or animals; therefore, ethical approval was not required.

Consent to participate

Not applicable.

Consent to publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

[1] M. Kramarz and M. Kmiecik, (2024) *"The role of the logistics operator in the network coordination of omni-channels"* **Applied Sciences** ***14***(12): 5206. DOI: 10.3390/app14125206.

[2] I. D. Nagy, D.-C. Dabija, R. E. Cramarenco, and M. I. Burcă-Voicu, (2024) *"The use of digital channels in omni-channel retail—An empirical study"* **Journal of Theoretical and Applied Electronic Commerce Research** ***19***(2): 797–817. DOI: 10.3390/jtaer19020042.

[3] Z. He, L. Chen, and B. Liu, (2024) *"Application of integrating reinforcement learning and intelligent scheduling in logistics distribution"* **Intelligent Decision Technologies** ***18***(1): 57–74. DOI: 10.3233/IDT-230528.

[4] Y. Wan, Z. Yan, and S. Wang, (2025) *"Pricing and return strategies in omni-channel apparel retail considering the impact of fashion level"* **Mathematics** ***13***(5): 890. DOI: 10.3390/math13050890.

[5] T. A. Rainy, (2025) *"AI-driven marketing analytics for retail strategy: A systematic review of data-backed campaign optimization"* **International Journal of Scientific Interdisciplinary Research** ***6***(1): 28–59. DOI: 10.63125/0k4k5585.

[6] J. Pan, C.-J. Lu, W.-J. Chen, K.-S. Wu, and C.-T. Yang, (2024) *"A study on the production-inventory problem with omni-channel and advance sales based on the brand owner's perspective"* **Mathematics** ***12***(19): 3122. DOI: 10.3390/math12193122.

[7] S. Mou, (2022) *"Integrated order picking and multi-skilled picker scheduling in omni-channel retail stores"* **Mathematics** ***10***(9): 1484. DOI: 10.3390/math10091484.

[8] Y. Xie, H.-Q. Ye, and W. Zhu, (2025) *"Prediction and optimization for multi-product marketing resource allocation in cross-border e-commerce"* **Journal of Theoretical and Applied Electronic Commerce Research** ***20***(2): 124. DOI: 10.3390/jtaer20020124.

[9] C. Ziakis and M. Vlachopoulou, (2023) *"Artificial intelligence in digital marketing: Insights from a comprehensive review"* **Information** ***14***(12): 664. DOI: 10.3390/info14120664.

[10] J. Jeong, D. Hong, and S. Youm, (2022) *"Optimization of the decision-making system for advertising strategies of small enterprises—Focusing on company A"* **Systems** ***10***(4): 116. DOI: 10.3390/systems10040116.

[11] Y. Zhang, (2025) *"A method to manage the energy consumption of cloud centers for predictability in neuro-fuzzy networks"* **Journal of Engineering and Applied Science** ***72***(1): 85. DOI: 10.1186/s44147-025-00646-4.

[12] GoMask.ai. *Marketing Attribution Path Dataset (Synthetic Dataset)*. Retrieved March 31, 2026. 2026.

[13] S. Roosta, S. J. Sadjadi, and A. Makui, (2025) *"Dynamic pricing modeling and inventory management in omnichannel retail using Quantum Decision Theory and reinforcement learning"* **PLOS ONE** ***20***(10): e0333068. DOI: 10.1371/journal.pone.0333068.

[14] S. Huang and Y. Yang, (2025) *"Reinforcement learning optimization strategies for dynamic pricing and inventory control in e-commerce retail"* **International Journal of Housing Science and Its Applications** ***46***(4): 29–44. DOI: 10.70517/ijhsa46403.