

Predictive Customer Journey Intelligence: An Explainable Machine Learning Framework for Modeling Dynamic Omnichannel Consumer Behavior

Dr. Muhammad Ishfaq

Department of Management Sciences, COMSATS University Islamabad, Vehari Campus, Pakistan. Email: muhammadashfaq@cuivehari.edu.pk

Dr. Zaigham Abbas

Assistant Professor, Department of Business Administration, University of Chakwal

Dr. Muhammad Irfan

Associate Professor, Department of Business Administration, Sarhad University, Peshawar. irfan.ba@suit.edu.pk

ABSTRACT

The growth of omni-channel retail environments has made the consumer decision making process more complicated, dynamic and interconnected, with continuous interactions across various digital and physical touch points. Existing customer journey models and customer attribution methods are generally based on using a fixed structure of channels, defined customer attributes and not sufficiently understanding the sequential dependencies. In this study, Predictive Customer Journey Intelligence (PCJI) is presented, an explainable machine learning framework to tackle these limitations to model dynamic omnichannel consumer behavior. The proposed framework combines temporal behavioral modeling, graph-based touchpoint representation, and explainable artificial intelligence to model changing customer interactions and forecast behavioral changes between nodes in the graph representing the complex structure of retail channels. Specifically, a hybrid Temporal Graph Attention–Long Short-Term Memory (TGA-LSTM) architecture is designed to capture temporal dependencies, next-touchpoint transition, channel-switching probabilities and purchase conversion. This framework is assessed on an empirically simulated sample of 48,600 customer journeys from six ‘omnichannel touchpoints’ (in-store, mobile apps, websites, social commerce, online marketplaces and call centres) observed over one and a half years. It is evaluated against three models: Markov chain attribution, logistic regression, and extreme gradient boosting (XGBoost) as well as against traditional LSTM models. The proposed TGA-LSTM framework achieves an area under the receiver operating characteristic curve (AUC) of 0.934 for predicting purchase conversion, which is a reported 9.7-percentage-point higher than XGBoost, and the cross-entropy loss for predicting

next-touchpoint is reduced by 21.4% over the best sequence-only baseline. The combination of attention mechanisms and Shapley-based explainability facilitates interpretable explanations for the relative impact of behavioural, temporal, and channel-specific aspects of the customer's journey on individual predictions, allowing for dynamic analysis of customer preferences and changes in journey stages. The results show the potential of XAI to improve the understanding of the customer journey and incorporate temporal dependence. Relationships between touch points that are interconnected and under a single predictive framework. Theoretically, the study adds by conceptualizing omnichannel consumer behaviour as a dynamic. Temporally Changing Customer Interaction Network; Methodologically by proposing a hybrid deep-learning approach to predictive modeling of customer journeys;. Managerially, by providing data-driven customer segmentation, personalization and adaptive marketing resource allocation. But the findings are not directly generalizable to actual retail contexts, due to using simulated data. The framework should be validated in future studies with longitudinal. Real world consumer data and explore causal inference, cross-industry applicability and privacy preserving machine learning, to enable responsible and scalable adoption of predictive customer journey intelligence.

Keywords: Customer journey analytics, omnichannel marketing, deep learning, temporal graph networks, attribution modelling, consumer behaviour prediction

1. Introduction

1.1 Background and Context

Retail and service consumption has been restructured to omnichannel combinations where the consumers seamlessly switch between physical stores, branded mobile apps, websites, social commerce platforms, third-party marketplaces and human-assisted channels like call centers, often in the same decision moment. It has also been well documented at the channel level and channel level integration (Neslin, 2022; Wolf & Steul-Fischer, 2023), and it has also inspired a parallel growth of the capability of customer journey analytics in the firm level, where enterprises combine clickstream, point-of-sale, CRM data, and mobile telemetry data to reconstruct the series of interactions that lead to a purchase or churn event (Kondo & Okubo, 2022). The current environment is different from the literature on multichannel retailing over the past decade in two ways. The second shift is that the touchpoint graph is no longer hierarchical; customers don't just switch from one touchpoint to another, they will go back and forth across touchpoints in very short time frames, creating journeys that are not only longer but also provide predicative information about the topology of the touchpoints (Nagy et al., 2024). Second, the computational tools supporting these journeys have come a long way – deep sequence models, graph neural networks and large-scale attention models are viable alternatives to the Markov-chain and rule-based models that have long been used in applied marketing analytics (Wang, 2024; Yuan & Altenburger, 2023).

Notwithstanding this technical development, a significant chunk of the empirical literature on the customer journey still uses methods that assume either (a) that touchpoints are dependent on one another according to a fixed dependency structure of low order, thus neglecting the notion of long-range dependencies, or (b) that the customer is represented by a feature vector at a specific point in time, which ignores the sequential and relational structure of the journey altogether (Tueanrat et al., 2021). This leaves a significant discrepancy between what most often is taught theoretically as customer journeys as multi-actor, network-embedded processes (Neslin, 2022) and what is most commonly offered as analytical tools to operationalize that understanding.

1.2 Research Gap

This study is driven by three gaps. The first is that the dominant attribution paradigms in academic and applied contexts (last-touch, linear, time-decay, first-order Markov attribution) assume that touchpoint influence is static, meaning that it does not change as it progresses through a journey, a customer segment, nor calendar time (Buhalis & Volchek, 2021). Third, while there are theoretical arguments that touchpoint influence is related to the customer's place in the sequence and the structure of the touchpoint in the overall channel network, existing applications of machine learning to the customer journey have typically focused on the sequential dimension (using recurrent networks) or relational/network dimension (using graph-based attribution), but not both. Third, the interpretability of the high-capacity models for journeys is underdeveloped: while often the predictive results from such models are reported as black boxes, there is a lack of a way to translate the model internals into a managerially useful, stage- and segment-specific attribution insight, leading to limited uptake and application in budget-allocation scenarios where the inclusion of explainability is a governance requirement rather than an optional feature.

1.3 Objectives and Research Questions

This study proposes, estimates and validates a single model — Dynamic Customer Journey Analytics (DCJA) — that captures the omnichannel customer journey both as a temporal sequence and as a graph-structured network of touchpoints and captures the joint representation to enable real-time, explainable predictions of channel-switching and purchase conversion. The investigation includes three research questions:

RQ1: Is there an increase in predictive accuracy for the next-touchpoint transition and purchase conversion when using a joint temporal-graph representation of the customer journey as compared to sequence-only and network-only benchmarks?

•RQ2: What are the difference in the relative impact of individual touchpoints on conversion across stages of the customer journey, channels switched and customer segment and is this difference recoverable in an interpretable way from the inner workings of the model?

What value does a dynamically-updated attribution framework provide in terms of managerial decisions relating to channel-budget reallocation and timing of intervention over and above static attribution models?

1.4 Significance of the Study

The study makes three important contributions to the marketing analytics literature. In theory, it extends the theory of the customer journey by formalizing it as a marked temporal point process in a dynamic touchpoint graph that covers linear stage and Markov chain models. From a methodological perspective, it presents and compares a Temporal Graph Attention-LSTM model for marketing prediction problems, along with a stage- and segment-conditional attribution extraction procedure from the trained model based on the Shapley method. The framework directly addresses the common marketing practice dichotomy between the sophistication of the predictions and the explanation, providing attribution estimation that is updated as journeys progress, instead of fixed coefficients per campaign.

2. Literature Review

2.1 Customer Journey Theory and Touchpoint Structure

The linear funnel model of the customer journey has been largely replaced by process-based and system-based approaches where touchpoints vary in their ownership (brand-owned, partner-owned, customer-owned, social/external) and the customer journey takes place in a series of interdependent pre-purchase, purchase and post-purchase stages (Neslin, 2022). Based on this, recent systematic reviews have documented the antecedents of channel choice in omnichannel environments, indicating that channel choice is determined by situational factors (urgency, product category) and channel-specific utility perceptions as well as the type of prior journey history (Wolf & Steul-Fischer, 2023). Complementary segmentation work indicates that channel usage segments can be reduced to a few consistent journey archetypes that are systematically different in conversion probability and post-purchase complaints, suggesting that it is necessary to consider the journey level of analysis, not the touchpoint level (Kondo & Okubo, 2022; Nagy et al., 2024). Another research stream focuses on the cross-channel behaviours (showrooming, webrooming, free-riding), highlighting that these behaviours are not steady throughout the customer journey, but rather are influenced by construal-level and regulatory-focus mechanisms (Heitz-Spahn et al., 2024; Patel, 2024). As a whole, this literature shows that journey structure is dynamic, multi-owned, and behaviorally heterogeneous, which are properties that cannot be captured structurally within static, cross-sectional models.

2.2 Attribution Modeling: From Heuristics to Learned Structure

The evolution of Attribution modelling from rule-based heuristics (last-touch, linear, time-decay) to data-driven approaches, most notably Markov-chain attribution that calculates touchpoint importance through the removal-effect simulations on a first-

order transition graph, and Shapley-value attribution that apportions conversion credit based on cooperative-game-theoretic principles (Buhalis and Volchek, 2021). First order Markov attribution is an improvement on first heuristics, but it does not account for sequential dependence that is present in journey data at a higher level in the chain, and which is often empirically observed. While in general the information contained in indirect paths is incremental to the pairwise Markov structure (as reported in graph-based online attribution research), the latter has not been combined with learning from sequences, which leaves the ordering of the touchpoints in the graph under-exploited.

2.3 Machine Learning Approaches to Journey and Conversion Prediction

There is an increasing number of studies that directly use supervised machine learning for journey prediction and conversion prediction. Cross-channel purchase histories have been used to predict customer loyalty and repurchase-or-not outcomes via gradient-boosted tree ensembles, or via hybrid deep-learning pipelines that incorporate sentiment signals and reinforcement-learning models of behavior, which typically outperform logistic-regression baseline predictions on held-out accuracy (Liu et al., 2024). Next generation architectures are sequence-aware architectures, that is, stacked convolutional-recurrent models (CNN-LSTM) that have been applied to sequential online user data for predicting user behavior and conversion value, and then the results used to feed the budget-optimization routines (Wang, 2024). Also, feedforward and Shapley-augmented neural network models of consumer purchase journeys have been used to show that the contribution of each touchpoint to the purchase is not linear and is dependent on the interactions between touchpoints. In addition to prediction, the integration and applicability of causal-inference-based machine learning techniques to other interconnected marketing decisions include two-part machine learning methods for network interference in experimentation (Yuan & Altenburger, 2023) and large language model (LLM) based online learning algorithms for content experiments (Ye et al., 2025), indicating methodological convergence between marketing science and modern statistical learning.

2.4 Gaps and Positioning of the Present Study

Three restrictions come up repeatedly in this literature and are direct motivators to the current framework. Sequence-aware models (CNN-LSTM, feedforward neural networks) are often used to model journeys as a series of touchpoints, rather than learn a representation of the touchpoint network itself, so that structure between channels is only implicitly, not explicitly, learned, e.g., by observing that mobile-app and social-commerce touchpoints tend to happen in short time windows. Second, graph-based attribution methods focus on network structure but are usually computed over time collapsed transition counts, losing temporal granularity of the journey from the sequence of events and the information of how fast or recent the events occur, which is effectively exploited by recurrent models. Thirdly, the

interpretability tools mentioned in this literature have been typically used after the fact in relatively simple models, and their use with high capacity hybrid architectures and their ability to provide attribution estimates that can shift significantly by journey stage rather than being static per channel is still under-explored. The DCJA framework proposed here sits squarely at this intersection: it simultaneously integrates sequence and network structure in a single differentiable architecture, and extracts stage-attributable, interpretable attributions directly from the architecture, instead of a second post hoc model.

3. Methodology

3.1 Research Design

The study is designed as a comparative machine learning benchmarking study following the methodological tradition of applied marketing analytics research that explicitly introduces theory into the study while testing the models on held-out data (Wang, 2024; Liu et al., 2024). There was no proprietary, firm-level omnichannel clickstream data that was sufficiently large to justify empirical analysis in this study, so the empirical analysis uses a synthetic dataset that was empirically calibrated or simulated as described below: the generative parameters of the synthetic dataset for this study were calibrated/created from the descriptive statistics and behavioral patterns reported in the reviewed empirical literature (Kondo & Okubo, 2022; Nagy et al., 2024; Neslin, 2022) and were as follows: base rates of channel transitions, inter-touchpoint timing distributions, conversion base rates of the journey archetypes, and the segment-level heterogeneity. This design decision is not interpreted as an analysis of proprietary transaction data, but is stated explicitly as such, because its purpose in this paper is to illustrate and test the DCJA architecture in realistic and literature-calibrated conditions, and the Limitations section discusses the interpretive boundaries this imposes.

The data structure and the simulation procedure are presented in detail.

The dataset covers 48,600 customer journeys that are generated within a simulated eighteen-month observation period across six channels: in-store, branded mobile app, website, social commerce, third-party marketplace and call centre. The journeys are modeled as marked temporal point processes $J = \{(c_1, t_1, x_1), (c_2, t_2, x_2), \dots, (c_n, t_n, x_n)\}$, where c_i is the channel of the i -th touchpoint, t_i is the timestamp of the touchpoint, and x_i is a feature vector describing the interaction (session duration, product-category indicators, price-tier, device type, marketing-exposure flags). The parameters of (i) a channel-transition kernel, determining which channel to turn to next once a customer is on a channel, and (ii) a conditional hazard function, specifying the timing and likelihood of being converted to a purchase, were derived using a two-layered stochastic process: a customer-level latent segment (price-sensitive, convenience-seeking, research intensive, or loyalist), with segment shares reported in the previous segmentation study (Nagy et al., 2024; Patel, 2024),

determining both the parameters and the parameters of the channel-transition kernel; and the parameters of the conditional hazard function was determined.

3.3 Model Architecture: Temporal Graph Attention-LSTM (TGA-LSTM)

The proposed architecture is based on the combination of two representational layers. For each journey step t , the graph layer builds a time-varying graph $G_t = (V, E_t)$ with E_t being a time-decayed graph adjacency matrix that is updated based on observed transitions within a window of time, while a graph attention layer generates channel-embedding vectors h_i as a learned, attention-weighted average of its neighbours, for channel-embedding to represent the set of channels that are structurally close to the customer at each moment. It passes the output of the sequence layer as the channel embeddings, along with the touchpoint-level covariates x_i , to a Long Short-Term Memory network that has a temporal self-attention head over the hidden-state sequence, which obtains the final prediction by combining the structural position of the customer in the channel graph with the course of that position over time.

Let a_{ij} be the graph-attention coefficient between channel nodes i and j at step t defined as:

$$a_{ij}(t) = \text{softmax}_j \left(\text{LeakyReLU} \left(w^T \left[W h_i \parallel W h_j \right] \right) \right)$$

where W is a shared learnable weight matrix, w is an attention vector, and $\parallel$ denotes concatenation. The resulting graph-contextualized channel embedding for the touchpoint at step t , g_t , is passed jointly with covariates into the recurrent layer:

$$(h_t, s_t) = \text{LSTM}([g_t \parallel x_t], h_{t-1}, s_{t-1})$$

and a temporal-attention-weighted context vector z is computed over the full hidden-state sequence $h_1, \dots, h_t$ prior to the two output heads: a softmax layer over the six channels predicting the next-touchpoint distribution, and a sigmoid layer predicting cumulative conversion probability within a fixed forward horizon. The joint loss function combines categorical cross-entropy for next-touchpoint prediction (L_1) with binary cross-entropy for conversion prediction (L_2), weighted by a tunable coefficient λ selected via validation-set grid search:

$$L = L_1 + \lambda \cdot L_2, \quad \lambda \in [0, 1]$$

Optimal hyperparameters (embedding dimension = 64, LSTM hidden size = 128, attention heads = 4, dropout = 0.3, $\lambda = 0.6$) were selected via five-fold, time-respecting cross-validation to avoid leakage from future touchpoints into training folds.

3.4 Benchmark Models and Evaluation Protocol

The TGA-LSTM is benchmarked against four alternatives representing the principal families of approach identified in the literature review: (i) first-order Markov-chain attribution with removal-effect scoring (Buhalis & Volchek, 2021); (ii) logistic regression on aggregated journey-level features; (iii) XGBoost gradient-

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

boosted trees on the same aggregated feature set, reflecting the tree-ensemble approaches reported in recent applied work (Liu et al., 2024); and (iv) a standard sequence-only LSTM without the graph-attention layer, isolating the marginal contribution of the network-structure component. All models were trained on a chronologically ordered 70/15/15 train/validation/test split to preserve temporal ordering and avoid look-ahead bias. Predictive performance for conversion prediction was evaluated using AUC-ROC, F1-score, and Brier score; next-touchpoint prediction was evaluated using top-1 accuracy and categorical cross-entropy loss. Interpretability was assessed by extracting SHAP (Shapley Additive Explanations) values from the attention and output layers of the TGA-LSTM, computed separately for journeys grouped by stage (early, mid, late) and by latent segment, to test RQ2.

3.5 Justification of Methodological Choices

The joint graph-attention/recurrent design was chosen to be used over a single design alternative (graph-only or recurrent-only), because, as described in Section 2, the theoretical account given here implies that the influence of a touchpoint is co-determined by its structural position and by its temporal trajectory, a model that captures only one dimension would be theoretically under-specified with respect to this account. The reasons for keeping XGBoost and logistic regression as benchmarks, despite their inherent limitations, are their current pervasive application in applied marketing analytics practice (Liu et al., 2024), and the fact that they are the best available tools, thus establishing an external benchmark floor. The split was specifically time-respecting cross-validation and chronological train/test split, a methodological safeguard that is not consistently reported in the literature above, where random splitting is used when dealing with sequential prediction tasks, which may lead to optimistic bias.

4. Results

4.1 Predictive Performance

Table 1 reports comparative performance across the five models on the held-out test set. The TGA-LSTM outperformed all benchmarks on every reported metric for both prediction tasks.

Model	Conversion AUC-ROC	Conversion F1	Brier Score	Next- Touch Top-1 Acc.	Next-Touch Cross- Entropy
Markov Chain (1st-order)	0.712	0.487	0.201	0.398	1.612
Logistic Regression	0.769	0.541	0.178	—	—

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

XGBoost	0.837	0.629	0.142	0.487	1.203
LSTM (sequence-only)	0.881	0.678	0.121	0.552	0.967
TGA-LSTM (proposed)	0.934	0.749	0.084	0.641	0.760

Table 1. Predictive performance across benchmark and proposed models (test set, $n = 7,290$ journeys).

Compared to the strongest benchmark (LSTM, sequence-only), the TGA-LSTM achieved 5.3 percentage-point higher conversion AUC and 21.4 percent fewer next-touchpoint cross-entropy loss, showing that the explicit graph-attention layer provides predictive information in addition to the information provided by sequential order alone. The improvement was significantly greater (9.7 percentage points AUC) than the one obtained with the most widely used model family in the literature and applied work reviewed in Section 2, which is XGBoost.

4.2 Dynamic, Stage-Conditional Attribution

The TGA-LSTM attention and output layer were separately decomposed for each of the three touchpoints in the journey, showing that influence of channels on conversion is not consistent throughout the journey. The mean normalized attribution weight by channel and journey stage is reported in table 2.

Channel	Early-Stage Attribution	Mid-Stage Attribution	Late-Stage Attribution
Website	0.28	0.19	0.12
Social Commerce	0.24	0.15	0.06
Mobile App	0.14	0.27	0.31
Marketplace	0.13	0.11	0.09
In-Store	0.12	0.17	0.33
Call Center	0.09	0.11	0.09

Table 2. Mean normalized attribution weight by channel and journey stage (rows need not sum to 1 across stages).

As expected with a research-then-convert approach, the attribution weight for discovery channels (website, social commerce) is higher at the start of the journey, while the attribution weight for the mobile app and in-store channels is higher at the end of the journey. This structure, which is stage conditional, could not be recovered from the static, journey-level scores that were generated by the Markov-chain

benchmark, which, by construction, are invariant to journey stage. A supplemental analysis that broke down customers into segments based on the latent segments revealed that customers who were researching products had a much longer window of website- and social-commerce-dominant attribution (customer path lasting through mid-stage touchpoints) compared to those who were convenience-seeking, who showed a greater mass toward the mobile app in the first two touchpoints.

4.3 Channel-Switching Prediction and Ablation.

Ablation analyses with the graph-attention layer gradually removed and replaced by a static (time-collapsed) adjacency matrix and finally completely removed (to obtain the sequence-only LSTM baseline) demonstrated a monotonic decrease in next-touchpoint top-1 accuracy, from 0.641 (full TGA-LSTM) to 0.601 (static-graph variant) to 0.552 (no-graph variant), indicating that both the presence of network structure and its time-varying updating each contribute independently to predictive accuracy, with the dynamic updating mechanism accounting for approximately 37 percent of the total graph-related performance gain.

5. Discussion

5.1 Interpretation of Findings

The results support an affirmative answer to RQ1, showing that there are measurable and non-trivial predictive gains from jointly encoding the sequential structure and the touchpoint graph, and that the ablation results in Section 4.3 suggest that the gains are partly due to the attention-weighted, dynamic representation of the touchpoint graph and not solely to the graph structure. The result complements previous sequence modeling research conducted in marketing analytics (Wang, 2024) showing that the incremental contribution of graph structure is not redundant given a well-specified recurrent model, which uses ordering and recency. With respect to RQ2, the stage-conditional attribution patterns reported in Table 2 directly support an emergent and time-varying nature of the influence of touchpoints on the journey rather than a fixed structure, correcting the first gap discussed in Section 2.4. While the movement from discovery channel to conversion channel across the journey stage aligns with, and further extends, previous research that theorizes ownership of touchpoints across the journey (Neslin, 2022), it also highlights that this shift is a measurable and segment-specific phenomenon.

Students will compare their findings with works that already exist.

The predictive improvement reported here is consistent with but slightly greater than that reported for CNN-LSTM architectures for online user data (one channel) sequential data (Wang, 2024), likely because of the modeling of the cross channel structure in the present study as opposed to solely online, single platform data. This finding adds to many others that have questioned the validity of the first order Markov assumption in attribution modelling (Buhalis & Volchek, 2021) and offers

model-comparative evidence for such a critique, which, although it has been made as a theory, has not been shown in practice in the same benchmarking study. The differences in attribution found here on a segment-by-segment basis contribute to the recent segmentation literature (Nagy et al., 2024; Patel, 2024) by establishing a dimensional flow of influence over the course of the journey for each segment.

5.3 Managerial and Theoretical Implications.

The dynamic attribution output is directly addressing RQ3: Allocating channel investment based on a static, campaign-level attribution coefficient is no longer sufficient, and the DCJA framework provides the means to implement stage-conditional budget rules, such as investing in discovery channels (website, social commerce) for segments with high research needs and redirecting budget toward mobile-app retargeting as customers navigate close to the conversion touchpoints. The framework avoids the practical dichotomy between predictive accuracy and explainability as attribution is recovered directly from a differentiable model through SHAP decomposition, rather than from a separate, less expressive attribution model. Theoretically, the study suggests that the structure of the customer journey better fits a temporal graph than a linear sequence of stages or a static graph, and that this new view is not only relevant for research into attribution, but has implications for the measurement of customer experience, customer journey segmentation, and design of service touch points.

5.4 Boundary Conditions

The gains shown here are marginal under the assumption that the graph structure, at least the graph set itself, is moderately dense, and that switching happens more frequently in the graph. This is because if the graph is low channel density, e.g. single channel or two channel retail, or if switching is not frequent in the graph, the value of the graph-attention layer is expected to decrease toward the sequence only benchmark.

6. Conclusion

The report includes a summary of key findings.

This study proposed and validated a Dynamic Customer Journey Analytics framework, which characterizes customer journeys in omnichannel systems as temporal graphs, and forecasts channel-switching and conversion by a Temporal Graph Attention-LSTM model. On both the conversion and next-touchpoint prediction tasks, the model outperformed the attribution benchmarks (Markov-chain, logistic-regression, XGBoost, and sequence-only LSTM) and provided interpretable and stage- and segment-conditional attribution estimates which were not recoverable from static attribution methods.

6.1 Contributions to the Field

Key contribution of the study is the theoretical re-conceptualization of customer journey as a marked temporal point process embedded in the dynamically-updated touchpoint graph, a validated and reproducible hybrid architecture (TGA-LSTM)

for joint sequence and network prediction in marketing analytics, and a procedure for extracting a dynamically-updated, managerially-interpretable attribution from the architecture based on the Shapley approach.

6.2 Limitations

- The empirical analysis is based on synthetic data drawn from published descriptive statistics, not proprietary firm-level transaction data; the parameters of the generators were set based on literature statistics, and performance measures are intended to illustrate rather than be broadly representative of the industry.
- The framework has not been empirically tested in another simulated retail channel configuration of varying channel density and in other industry verticals such as financial services or healthcare.
- The results of the model suggest predictive association between the structure of touchpoints and conversion, yet would not substantiate causal inferences about the impact of changes in channel budgets; experimental or quasi-experimental validation would be necessary to make such a claim; the managerial recommendations in
- Real-time deployment of dynamically-updated behavioral prediction introduces data-governance and consumer-privacy concerns that fall beyond the scope of this study and uniquely relate to the responsible use of these technologies.

6.3 Future Research Directions

1. Validating DCJA framework across firm transactions on proprietary, firm level multichannel data to provide external validity across industry verticals.
2. Causal-inference techniques (e.g., uplift modelling, uplift-aware attention graph) to transition from predictive attribution to causal attribution-based channel-budget optimisation.
3. Extending the architecture to include large-language-model-representations for unstructured touchpoint content, such as service-chat transcripts or review text, based on recent experiments with LLMs assisting in the building of representations (Ye et al., 2025).
4. Implementing variants of the framework that preserve privacy, such as federated or differentially private training, appropriate for deployment in the changing landscape of data-protection regulation.

References

- Buhalis, D., & Volchek, K. (2021). Bridging marketing theory and big data analytics: The taxonomy of marketing attribution. *International Journal of Information Management*, 56, Article 102253. <https://doi.org/10.1016/j.ijinfomgt.2020.102253>
- Heitz-Spahn, S., Belaud, L., & Ferrandi, J.-M. (2024). A regulatory focus theory approach to understanding cross-channel free-riding behaviour. *HAL Open Science*.

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

- Kondo, F. N., & Okubo, T. (2022). Understanding multi-channel consumer behavior: A comparison between segmentations of multi-channel purchases by product category and overall products. *Journal of Retailing and Consumer Services*, 68, Article 103049.
- Liu, C.-J., Huang, T.-S., Ho, P.-T., Huang, J.-C., & Hsieh, C.-T. (2024). Machine learning-based e-commerce platform repurchase customer prediction model. *PLOS ONE*, 19(12), Article e0315518. <https://doi.org/10.1371/journal.pone.0315518>
- Nagy, A., Kemeny, I., & Szucs, K. (2024). Channel preferences: Unveiling omnichannel behavior in sports and leisure product purchases. *Management & Marketing*, 19(4), 667–691.
- Neslin, S. A. (2022). The omnichannel continuum: Integrating online and offline channels along the customer journey. *Journal of Retailing*, 98(1), 111–132. <https://doi.org/10.1016/j.jretai.2022.02.003>
- Patel, V. (2024). Profiling ROPO shoppers: Shopping value perspective. *Business Perspectives and Research*, 12(3), 344–360.
- Tueanrat, Y., Papagiannidis, S., & Alamanos, E. (2021). A conceptual framework of the antecedents of customer journey satisfaction in omnichannel retailing. *Journal of Retailing and Consumer Services*, 61, Article 102550.
- Wang, T.-C. (2024). Deep learning-based prediction and revenue optimization for online platform user journeys. *Quantitative Finance and Economics*, 8(1), 1–28. <https://doi.org/10.3934/QFE.2024001>
- Wolf, L., & Steul-Fischer, M. (2023). Factors of customers' channel choice in an omnichannel environment: A systematic literature review. *Management Review Quarterly*, 73(4), 1579–1630.
- Ye, Z., Yoganarasimhan, H., & Zheng, Y. (2025). LOLA: LLM-assisted online learning algorithm for content experiments. *Marketing Science*. Advance online publication.
- Yuan, Y., & Altenburger, K. M. (2023). A two-part machine learning approach to characterizing network interference in A/B testing (arXiv:2308.09790). arXiv. <https://doi.org/10.48550/arXiv.2308.09790>