

## Predictive Data Analytics for Understanding Customer Behavior in Digital Commerce Platforms

**Dr. Amelia Richardson**

Assistant Professor, Birmingham Institute of Computational Research, United Kingdom

### Abstract

Digital commerce platforms generate substantial volumes of behavioral information through customer browsing, product searches, shopping cart activities, purchase transactions, and repeated interactions. These digital records provide opportunities to examine consumer preferences, purchasing intentions, and behavioral patterns through predictive analytical techniques. However, customer behavior is influenced by numerous interconnected factors, including product characteristics, navigation experiences, perceived trust, pricing conditions, and previous purchasing experiences. Understanding these relationships requires analytical approaches capable of processing complex behavioral information and identifying meaningful patterns within customer interactions. The present paper develops a hypothetical methodological framework for examining the application of predictive data analytics to customer behavior in digital commerce environments. The proposed framework incorporates clickstream analysis, behavioral feature construction, customer segmentation, and supervised predictive modeling. Logistic regression, random forests, gradient boosting, and sequence-based feature models are considered as alternative approaches for examining purchase conversion patterns. A fictional dataset representing customer interactions across a digital retail platform is introduced to illustrate the proposed analytical procedure.

Two illustrative tables demonstrate how customer navigation stages and model classification performance could be reported. The first table describes a hypothetical purchasing funnel, while the second presents fictional values for area under the receiver operating characteristic curve, average precision, and Brier score. These values are manually constructed examples and do not represent actual experiments or verified customer behavior. The methodological discussion identifies data quality, behavioral variability, model interpretation, privacy protection, and evaluation reliability as important considerations in customer analytics. The paper concludes that predictive analytics offers a useful framework for investigating online purchasing behavior, provided that analytical procedures are supported by representative data, appropriate validation, and responsible information management. The proposed approach is intended to guide future empirical investigations rather than establish experimentally demonstrated commercial outcomes.

**Keywords:** Predictive Data Analytics, Customer Behavior, Digital Commerce, Purchase Prediction, Clickstream Analysis, Customer Segmentation, Machine Learning, Consumer Analytics, Conversion Modeling, Behavioral Prediction.

## 1. Introduction

Digital commerce platforms have changed the ways in which consumers discover products, compare alternatives, evaluate commercial information, and complete purchasing transactions. Online retail environments provide customers with continuous access to product catalogs, promotional information, purchasing facilities, and interactive services. These activities generate detailed digital records that can be examined to understand consumer interactions with commercial platforms. The availability of behavioral information has consequently increased the relevance of data analytics in electronic commerce research. Rather than relying exclusively on conventional sales reports, researchers can investigate the sequence and characteristics of customer activities occurring before, during, and after purchasing decisions.

Customer behavior within digital commerce environments is influenced by interactions between individual preferences, product information, platform characteristics, and situational conditions. Consumers may visit an online platform with clearly defined purchasing objectives, conduct extensive product comparisons, or browse available offerings without an immediate intention to purchase. These different behavioral patterns complicate the interpretation of digital activity records. A large number of page views, for example, does not necessarily indicate strong purchasing intention. Similarly, repeated visits may reflect customer interest, uncertainty, information gathering, or dissatisfaction with previous experiences. Understanding these distinctions requires analytical procedures capable of considering behavioral context.

Predictive data analytics provides methodological opportunities for examining relationships between previous customer activities and subsequent purchasing outcomes. Statistical models and machine learning algorithms can be used to estimate purchase probabilities, identify customers exhibiting similar behavioral patterns, and investigate factors associated with repeated transactions. These approaches may incorporate variables such as browsing frequency, product-category interest, cart activity, transaction history, and visit duration. Nevertheless, predictive models require careful development because customer behavior can change over time and observed correlations do not necessarily establish causal relationships. The quality of analytical conclusions therefore depends on suitable data preparation, feature selection, and evaluation procedures.

The development of predictive systems has also encouraged greater interest in personalization and customer relationship management. Digital commerce platforms may employ analytical information to organize product recommendations, identify potential customer disengagement, or improve the relevance of website interactions. However, predictive classification and behavioral intervention represent different activities. A model that identifies customers with higher purchase probabilities does not automatically demonstrate that a particular recommendation or promotional strategy will increase sales. Establishing such effects requires appropriate experimental evaluation. Consequently, research into digital customer behavior should distinguish between predictive performance, behavioral interpretation, and commercially meaningful outcomes.

The present paper examines predictive data analytics as a methodological approach for understanding customer behavior in digital commerce platforms. It introduces an illustrative framework combining customer activity records, behavioral indicators, predictive classification, and structured model assessment. A fictional online retail environment is considered to demonstrate how purchasing patterns and classification performance might be evaluated. The study does not analyze actual customers or commercial transactions. Instead, it provides a conceptually organized research design through which

future investigators could examine online behavioral relationships, predictive reliability, and the methodological limitations of data-driven customer analysis.

## 2. Objectives of the Study

The primary objective of this study is to examine the methodological application of predictive data analytics for understanding customer behavior within digital commerce platforms. Particular attention is given to browsing patterns, purchasing intentions, customer segmentation, and prediction reliability. Three specific objectives establish the analytical direction of the proposed research framework.

### 2.1 Examination of Customer Navigation and Purchasing Patterns

The first objective is to investigate how customer navigation activities may be represented through measurable behavioral indicators. Digital commerce platforms record interactions involving product searches, category exploration, page transitions, cart additions, and completed transactions. These activities provide a potential basis for distinguishing different purchasing behaviors. The proposed study considers how visit frequency, browsing depth, product interest, and interaction sequences could be organized into analytical variables. Particular attention is given to identifying patterns associated with customer progression through successive stages of an online purchasing process.

The investigation also aims to distinguish exploratory browsing from behavior that may indicate stronger purchasing interest. Customers frequently interact with commercial platforms for different purposes, and similar digital actions may reflect different intentions. Consequently, individual activities should be interpreted in relation to broader navigation sequences rather than treated as isolated behavioral signals. This objective emphasizes the value of examining customer pathways while recognizing that observed interactions cannot independently reveal every psychological or situational factor influencing consumer decisions.

### 2.2 Evaluation of Predictive Customer Classification Models

The second objective focuses on establishing a comparative framework for predictive models intended to estimate purchasing outcomes. Logistic regression, random forests, gradient boosting, and sequence-informed classification approaches are considered as alternative analytical techniques. Their methodological comparison would examine the ability to distinguish purchasing sessions from sessions without completed transactions. The proposed evaluation includes discrimination, precision-oriented assessment, and probability calibration. These measures would provide complementary information about predictive performance under the class distributions typically encountered in digital commercial activity.

A further objective involves examining how differences in behavioral feature construction may affect model usefulness. Models trained using aggregated visit statistics may represent customer activity differently from approaches incorporating the sequence of earlier interactions. Predictive comparison should therefore distinguish the contribution of additional behavioral information from the effects of algorithmic complexity. Consistent validation procedures, independent testing, and appropriate temporal ordering are necessary for meaningful interpretation. This objective emphasizes reproducibility and the avoidance of misleading performance estimates caused by inappropriate model development.

## 2.3 Identification of Factors Affecting Customer Prediction Reliability

The third objective is to identify methodological factors that may influence the reliability of customer behavior prediction. Incomplete interaction records, inconsistent customer identifiers, changing product availability, promotional conditions, and variation in customer preferences can complicate behavioral modeling. These characteristics may affect the relationship between observed activities and subsequent purchases. The proposed framework considers how data preparation, feature selection, and temporal validation could address selected difficulties while maintaining an appropriate distinction between observable activity and unmeasured purchasing motivations.

The study also seeks to identify limitations associated with privacy, fairness, and the interpretation of predictive results. Customer activity records may contain sensitive behavioral information, making responsible collection and analytical processing essential. Predictive models may also reflect existing differences in customer access, purchasing opportunities, or platform experiences. Therefore, analytical conclusions should be assessed in relation to representativeness, data governance, and potential decision-making consequences. This objective provides a foundation for examining predictive usefulness without assuming that numerical accuracy alone establishes commercial or ethical suitability.

## 3. Review of Literature

### 3.1 Digital Navigation and Online Purchase Behavior

Bucklin and Sismeiro (2003) examined website browsing behavior through clickstream information and demonstrated the importance of analyzing individual navigation activities. Moe (2003) investigated differences among online shoppers according to their underlying visit objectives. Montgomery et al. (2004) subsequently examined online browsing paths and demonstrated the analytical relevance of navigation sequences. These studies established important methodological foundations for interpreting digital interactions. They also indicated that customer activity should be examined as a sequence of related decisions rather than a collection of disconnected website measurements.

Moe and Fader (2004) developed a model for understanding purchasing conversion based on customer visit and transaction histories. Bucklin and Sismeiro (2009) reviewed developments in clickstream analysis and discussed the methodological opportunities and limitations of behavioral tracking. Their contributions demonstrate the relevance of customer navigation information to purchasing analysis. However, they also emphasize that observed browsing patterns require careful interpretation because digital records do not necessarily explain all motivations or contextual conditions influencing a consumer's decision.

### 3.2 Customer Retention and Predictive Commercial Analytics

Verhoef (2003) investigated the influence of customer relationship management efforts on retention and customer-share development. Gupta et al. (2004) examined the valuation of customers through expected future earnings. Fader et al. (2005) developed a purchasing model intended to support customer-base analysis and future transaction estimation. These contributions established important connections between customer behavior, purchasing frequency, and commercial value. Their findings suggest that predictive customer analytics should consider continuing relationships rather than focusing exclusively on individual transactions or immediate purchasing outcomes.

Neslin et al. (2006) examined methodological factors affecting the prediction of customer defection, while Lemmens and Croux (2006) investigated ensemble classification methods for predicting churn. Chen et al. (2012) discussed broader developments in business intelligence and analytical decision-making. Provost and Fawcett (2013) subsequently emphasized the relationship between predictive modeling and business-oriented analytical reasoning. Together, these contributions indicate that the practical value of customer prediction depends on appropriate modeling procedures, suitable evaluation criteria, and clear connections between analytical objectives and organizational decisions.

### 3.3 Consumer Trust, Personalization, and Recommendation Systems

McKnight et al. (2002) examined the measurement of trust within electronic commerce environments. Pavlou (2003) investigated the relationship among perceived risk, trust, technology acceptance, and online transaction intentions. Gefen et al. (2003) further examined how trust and technology acceptance influence online purchasing behavior. These studies indicate that digital consumer decisions cannot be explained solely through observable navigation activity. Perceived security, vendor reliability, and interaction quality may influence purchasing intentions even when such characteristics are not directly represented in behavioral datasets.

Adomavicius and Tuzhilin (2005) reviewed recommender-system approaches and their limitations, while Tam and Ho (2006) investigated how website personalization affects information processing and decision outcomes. Hu et al. (2008) examined recommendation methods using implicit behavioral feedback, and Koren et al. (2009) discussed matrix factorization techniques. Rendle et al. (2009) investigated ranking methods based on implicit interactions. These studies demonstrate that behavioral information can support personalization, although effective recommendations require appropriate evaluation and careful distinction between observed customer interest and actual purchasing preferences.

## 4. Research Methodology

### 4.1 Research Design and Analytical Orientation

The proposed study adopts a hypothetical quantitative research design intended to examine predictive relationships between digital customer activities and purchasing outcomes. The methodological framework combines descriptive behavioral analysis with supervised classification. Its primary analytical target is whether a customer session results in a completed purchase. The design considers four predictive model configurations to illustrate potential differences in classification performance. No actual commercial platform data have been collected or analyzed, and the proposed procedures are presented exclusively as a methodological demonstration.

The research framework would involve data preparation, behavioral variable construction, exploratory examination, model training, and independent evaluation. Navigation information would be organized into customer sessions, allowing the proposed investigation to represent sequential interactions before purchasing outcomes occur. All predictive features would be restricted to information available at the specified prediction point. This restriction is important because including subsequent checkout or transaction information could produce misleading performance estimates. The intended design therefore emphasizes temporal consistency and protection against target leakage.

## 4.2 Hypothetical Dataset and Customer Behavioral Variables

For illustrative purposes, the study assumes a fictional digital retail dataset containing 20,000 customer sessions associated with approximately 12,500 pseudonymous visitor profiles. The proposed observation period covers 90 consecutive days. Each session is assumed to contain recorded activities involving website entry, product exploration, shopping cart interaction, and possible purchase completion. These figures are hypothetical and have not been generated from an operational commerce system. They are introduced to demonstrate the organization of a potential customer behavior dataset.

The proposed analytical variables would include session duration, product views, search frequency, category diversity, previous visit frequency, and early cart interactions. Historical purchasing activity could also be considered when legitimately available before the prediction point. Missing records would be assessed, while inconsistent identifiers and duplicate events would require appropriate correction. Customer information would be pseudonymized, and unnecessary personal attributes would be excluded. These procedures are intended to support behavioral analysis without assuming access to identifiable customer information or actual commercial records.

## 4.3 Predictive Modeling and Evaluation Procedures

The proposed modeling stage would compare logistic regression, random forests, gradient boosting, and a classifier incorporating sequential behavioral features. Logistic regression would provide an interpretable baseline, while tree-based approaches would allow nonlinear relationships and interactions to be represented. The sequence-informed configuration would incorporate features describing the order of early session activities. Hyperparameter selection would be conducted using validation data, and model inputs would be limited to information available before the specified purchasing outcome.

A future experiment would employ appropriately separated training, validation, and independent testing partitions. Chronological and customer-level separation would be considered to reduce inappropriate information sharing between development and evaluation samples. Model performance would be assessed using area under the receiver operating characteristic curve, average precision, and Brier score. These metrics would examine discrimination, performance under class imbalance, and probability accuracy.

## 5. Analysis

### 5.1 Illustrative Assessment of Customer Purchase Funnel Behavior

Customer purchasing behavior can be examined by identifying successive stages within a digital commerce interaction. These stages may include initial website visits, product exploration, cart additions, checkout initiation, and completed purchases. Table 1 presents a hypothetical purchasing funnel containing manually constructed session counts. The numerical pattern illustrates how customer activity might decline across successive stages. These values have not been obtained from website records and should not be interpreted as actual customer conversion patterns or commercial performance.

The illustrative funnel demonstrates how stage-specific progression rates could be calculated from behavioral event records. A decrease between product exploration and cart addition may indicate a stage requiring further investigation, although it cannot independently identify the reason for customer departure. Pricing, product suitability, navigation difficulties, and external circumstances may contribute to such behavior. Consequently, descriptive funnel analysis should be interpreted alongside additional

contextual information. The proposed framework uses these indicators to organize potential research questions rather than establish causal explanations.

**Table 1. Hypothetical Customer Activity Across Digital Purchasing Stages**

| Customer Interaction Stage | Sessions | Progression From Previous Stage (%) |
|----------------------------|----------|-------------------------------------|
| Website Visits             | 20,000   | 100.0                               |
| Product Exploration        | 15,000   | 75.0                                |
| Shopping Cart Addition     | 9,000    | 60.0                                |
| Checkout Initiation        | 6,000    | 66.7                                |
| Completed Purchases        | 4,000    | 66.7                                |

The hypothetical overall purchase conversion rate is 20%, calculated from 4,000 completed-purchase sessions divided by 20,000 total sessions. This figure serves only to demonstrate the calculation of a descriptive commercial indicator and does not represent an observed industry benchmark.

## 5.2 Illustrative Evaluation of Predictive Purchase Classification

Predictive classification involves estimating the likelihood that customer behavior will be associated with a defined purchasing outcome. The hypothetical values in Table 2 illustrate possible differences among four analytical models. Area under the receiver operating characteristic curve represents discrimination across decision thresholds, while average precision is particularly informative when purchasing sessions are less frequent than nonpurchasing sessions. Brier score summarizes the squared differences between predicted probabilities and actual outcomes. All reported values are invented examples rather than measured classification results.

The hypothetical comparison assigns progressively stronger illustrative discrimination and probability performance to more flexible modeling approaches. However, such numerical differences do not demonstrate actual predictive superiority because the algorithms have not been implemented. A completed investigation would require appropriate model training, independent customer data, uncertainty estimates, and careful examination of classification errors. Predictive usefulness should also be evaluated in relation to operational objectives.

**Table 2: Hypothetical Predictive Model Performance for Session Purchase Classification**

| Predictive Model        | AUROC | Average Precision | Brier Score |
|-------------------------|-------|-------------------|-------------|
| Logistic Regression     | 0.68  | 0.31              | 0.150       |
| Random Forest           | 0.75  | 0.40              | 0.137       |
| Gradient Boosting       | 0.80  | 0.47              | 0.122       |
| Sequence-Informed Model | 0.83  | 0.51              | 0.116       |

## 6. Discussion

### 6.1 Behavioral Information and Purchase Intention

Customer interaction records provide useful information about how individuals navigate digital commerce platforms. Product searches, repeated visits, and category exploration may reflect different levels of familiarity or interest. However, observable behavior does not fully explain purchasing intention because customers may interact with commercial platforms for several reasons. The proposed analytical framework therefore emphasizes the interpretation of behavioral sequences rather than assuming that any single activity establishes a definite intention to purchase.

The hypothetical purchasing funnel illustrates how customer progression could be examined through successive interaction stages. Differences in progression rates may identify areas requiring further analytical attention, but they cannot independently explain customer decisions. A reduction in cart progression, for instance, may arise from product comparisons, changing preferences, or technical difficulties. Meaningful interpretation would require additional evidence, potentially including usability research, customer feedback, or appropriately designed controlled experiments conducted within an actual digital commerce environment.

### 6.2 Predictive Modeling and Customer Segmentation

Predictive modeling can support customer segmentation by identifying differences in observed behavioral patterns and estimated purchasing probabilities. Classification methods may distinguish sessions with relatively high predicted purchase likelihood from those associated with lower estimates. However, these distinctions represent statistical predictions rather than fixed psychological categories. The proposed framework therefore treats segmentation as an analytical process that requires continuous evaluation. Customer behavior can change, and models should be assessed for their ability to remain reliable across different conditions.

The hypothetical comparison of predictive algorithms demonstrates a possible approach to assessing model flexibility and performance. More complex methods may capture nonlinear behavioral relationships, although they may also create interpretability and maintenance challenges. A comparatively simple model can remain useful when its predictions are sufficiently accurate and understandable. Consequently, selecting a classification approach should involve consideration of discrimination, calibration, computational requirements, and the specific commercial decision that the predictive information is intended to support.

### 6.3 Analytical Interpretation and Commercial Decision-Making

Predictive customer analytics can provide an informational basis for examining purchasing patterns, potential disengagement, and behavioral differences among customer groups. Nevertheless, predictive associations should not be interpreted automatically as evidence that an intervention will change customer behavior. A model might identify customers who are already likely to purchase without determining whether additional promotional activity would influence their decisions. This distinction is important when analytical outputs are used to inform personalization, customer communication, or resource allocation.

Kohavi et al. (2009) discussed controlled experiments as a method for evaluating changes within online environments. Their methodological perspective is relevant because commercial effectiveness requires evidence beyond retrospective predictive accuracy. A future study could examine whether

decisions informed by predictive models produce measurable improvements when compared with suitable alternatives. Such an evaluation should consider customer outcomes, operational costs, and possible unintended consequences. The present paper does not establish these effects because its numerical illustrations are hypothetical.

## 7. Challenges

### 7.1 Incomplete Behavioral Records and Data Integration

Digital commerce platforms may generate customer information through websites, mobile applications, payment systems, and marketing interfaces. Integrating these records can be difficult when event definitions, identifiers, or reporting procedures are inconsistent. Missing navigation events and duplicate transactions may distort behavioral features and influence predictive estimates. In addition, customers can interact through multiple devices or anonymous sessions, making it difficult to establish complete activity histories. These limitations require carefully documented data integration and quality assessment procedures.

Incomplete records can also create uncertainty about whether observed inactivity represents genuine customer disengagement. Customers may continue shopping through another device, purchase through an alternative channel, or return after the observation period. Consequently, behavioral classifications should account for the boundaries of the available data. The proposed methodology emphasizes consistent session definitions, appropriate observation windows, and transparent treatment of missing information. Without these safeguards, predictive models may capture recording limitations rather than meaningful differences in customer behavior.

### 7.2 Privacy Protection and Responsible Personalization

Customer analytics may involve detailed information about browsing activities, purchasing preferences, and repeated interactions. These records can reveal commercially and personally sensitive behavioral patterns even when direct identifying attributes are excluded. Responsible analytical development therefore requires legitimate data collection practices, suitable access controls, and restrictions on unnecessary information processing. Pseudonymization may reduce certain privacy risks, but it does not automatically eliminate the possibility of identifying individuals through combinations of recorded activities.

Personalization also raises questions concerning the appropriate use of inferred customer preferences. Predictive models may classify customers according to behavior they did not explicitly intend to disclose or use for profiling. Such practices require careful attention to transparency, data minimization, and the potential consequences of automated targeting. The proposed framework therefore emphasizes responsible information handling and the evaluation of predictive outputs. Commercial usefulness should not be treated as sufficient justification for unnecessary or intrusive behavioral monitoring.

### 7.3 Behavioral Change and Model Generalization

Customer preferences and purchasing activities may vary according to product availability, seasonal demand, promotional conditions, and changes in platform design. These developments can alter the relationships between behavioral indicators and purchasing outcomes. A predictive model developed using one observation period may therefore provide unreliable estimates when applied to a different

period. Regular performance assessment is necessary to identify meaningful deterioration. However, changes should be interpreted carefully because short-term fluctuations do not always justify immediate model retraining.

Model generalization presents additional difficulties when a digital commerce platform serves different customer groups or product categories. Behavioral relationships observed among frequent customers may not apply to new visitors, while purchasing patterns may differ across product types. Consequently, aggregate performance measures can conceal important variations in predictive reliability.

## 8. Conclusion

Predictive data analytics provides a structured methodological approach for investigating customer behavior within digital commerce platforms. The present paper introduced a hypothetical research framework combining customer navigation analysis, behavioral feature construction, predictive classification, and comparative evaluation. The proposed methodology emphasizes the relationship between observable digital interactions and purchasing outcomes while recognizing the complexity of consumer decision-making. Its primary contribution is a clearly organized analytical design suitable for future empirical investigation.

The illustrative purchasing funnel and model comparison tables demonstrate how customer activity patterns and predictive performance could be reported. However, all numerical values are hypothetical and do not establish actual consumer behavior, model superiority, or commercial improvement. Reliable findings would require authentic behavioral records, reproducible computational procedures, appropriate data partitioning, and independently evaluated predictions. Moreover, potential business benefits would need to be assessed separately from the statistical accuracy of the predictive models.

Overall, the study identifies behavioral interpretation, data quality, privacy protection, and model generalization as essential considerations in digital customer analytics. Future research should examine these issues through representative datasets and carefully designed evaluations. Predictive systems may support informed commercial decisions when their outputs are reliable, interpretable, and responsibly applied. Nevertheless, genuine conclusions about customer behavior and operational effectiveness must be established through verified empirical evidence rather than illustrative numerical comparisons.

## References

1. Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), 734–749. <https://doi.org/10.1109/TKDE.2005.99>
2. Bucklin, R. E., & Sismeiro, C. (2003). A model of Web site browsing behavior estimated on clickstream data. *Journal of Marketing Research*, 40(3), 249–267. <https://doi.org/10.1509/jmkr.40.3.249.19241>
3. Fader, P. S., Hardie, B. G. S., & Lee, K. L. (2005). RFM and CLV: Using iso-value curves for customer base analysis. *Journal of Marketing Research*, 42(4), 415–430. <https://doi.org/10.1509/jmkr.2005.42.4.415>
4. Feng, Y., Yin, Y., Wang, D., Ignatius, J., Cheng, T. C. E., Marra, M., & Guo, Y. (2024). Enhancing e-commerce customer churn management with a profit- and AUC-focused prescriptive analytics approach. *Journal of Business Research*, 184, Article 114872. <https://doi.org/10.1016/j.jbusres.2024.114872>

5. Gadgil, K., Gill, S. S., & Abdelmoniem, A. M. (2023). A meta-learning based stacked regression approach for customer lifetime value prediction. *Journal of Economy and Technology*, 1, 197–207. <https://doi.org/10.1016/j.ject.2023.09.001>
6. Kannan, P. K., & Li, H. A. (2017). Digital marketing: A framework, review and research agenda. *International Journal of Research in Marketing*, 34(1), 22–45. <https://doi.org/10.1016/j.ijresmar.2016.11.006>
7. Kim, S., Shin, W., & Kim, H. W. (2024). Predicting online customer purchase: The integration of customer characteristics and browsing patterns. *Decision Support Systems*, 177, Article 114105. <https://doi.org/10.1016/j.dss.2023.114105>
8. Koehn, D., Lessmann, S., & Schaal, M. (2020). Predicting online shopping behaviour from clickstream data using deep learning. *Expert Systems with Applications*, 150, Article 113342. <https://doi.org/10.1016/j.eswa.2020.113342>
9. Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing*, 80(6), 69–96. <https://doi.org/10.1509/jm.15.0420>
10. Moe, W. W., & Fader, P. S. (2004). Dynamic conversion behavior at e-commerce sites. *Management Science*, 50(3), 326–335. <https://doi.org/10.1287/mnsc.1040.0153>
11. Necula, S.-C., & Păvăloaia, V.-D. (2023). AI-driven recommendations: A systematic review of the state of the art in e-commerce. *Applied Sciences*, 13(9), Article 5531. <https://doi.org/10.3390/app13095531>
12. Norouzi, V. (2024). Predicting e-commerce CLV with neural networks: The role of NPS, ATV, and CES. *Journal of Economy and Technology*, 2, 174–189. <https://doi.org/10.1016/j.ject.2024.04.004>
13. Sismeiro, C., & Bucklin, R. E. (2004). Modeling purchase behavior at an e-commerce Web site: A task-completion approach. *Journal of Marketing Research*, 41(3), 306–323. <https://doi.org/10.1509/jmkr.41.3.306.35985>
14. Sun, Y., Liu, H., & Gao, Y. (2023). Research on customer lifetime value based on machine learning algorithms and customer relationship management analysis model. *Heliyon*, 9(2), Article e13384. <https://doi.org/10.1016/j.heliyon.2023.e13384>
15. Valencia-Arias, A., Uribe-Bedoya, H., González-Ruiz, J. D., Sánchez Santos, G., Chapoñan Ramírez, E., & Martínez Rojas, E. (2024). Artificial intelligence and recommender systems in e-commerce. *Trends and research agenda. Intelligent Systems with Applications*, 24, Article 200435. <https://doi.org/10.1016/j.iswa.2024.200435>
16. Verhoef, P. C., Lemon, K. N., Parasuraman, A., Roggeveen, A., Tsiros, M., & Schlesinger, L. A. (2009). Customer experience creation: Determinants, dynamics and management strategies. *Journal of Retailing*, 85(1), 31–41. <https://doi.org/10.1016/j.jretai.2008.11.001>
17. Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. *Journal of Marketing*, 80(6), 97–121. <https://doi.org/10.1509/jm.15.0413>
18. Wen, Z., Lin, W., & Liu, H. (2023). Machine-learning-based approach for anonymous online customer purchase intentions using clickstream data. *Systems*, 11(5), Article 255. <https://doi.org/10.3390/systems11050255>
19. Yan, Y., & Resnick, N. (2024). A high-performance turnkey system for customer lifetime value prediction in retail brands. *Quantitative Marketing and Economics*, 22, 169–192. <https://doi.org/10.1007/s11129-023-09272-x>
20. Zhang, S., Yao, L., Sun, A., & Tay, Y. (2019). Deep learning based recommender system: A survey and new perspectives. *ACM Computing Surveys*, 52(1), Article 5. <https://doi.org/10.1145/3285029>