JOURNAL OF ASIAN SCIENTIFIC RESEARCH (JASR)
2025, Vol. 15, No. 3, May - June, pp. 31-.37
website: https://joasr.com

Exploring User Interaction Patterns to Improve Predictive Modeling in Cloud-Based Sales Systems

Hari VijayaPrasath,
Predictive Analytics Specialist – Sales & User Behavior, India.

Abstract

Cloud-based sales systems increasingly rely on predictive models to optimize customer relationship management, sales forecasting, and automated outreach. However, traditional predictive modeling often overlooks fine-grained user interaction data—such as browsing behavior, session duration, and clickstream patterns—which can offer critical insights into user intent and conversion likelihood. This paper explores how integrating user interaction patterns can improve the performance of predictive modeling in cloud-based sales platforms. We analyze log data from a large-scale cloud CRM system, identify key behavioral features, and assess their impact using various machine learning algorithms. Results show a significant improvement in prediction accuracy and business KPIs such as lead conversion and customer retention.

Keywords: user behavior analytics, predictive modeling, cloud CRM, sales forecasting, machine learning, behavioral data.

Citation: VijayaPrasath, H. (2025). Exploring User Interaction Patterns to Improve Predictive Modeling in Cloud-Based Sales Systems. *Journal of Asian Scientific Research (JASR),* 15(3), 31–37.

1. Introduction

Cloud-based Customer Relationship Management (CRM) platforms have transformed how sales organizations track and manage leads, providing scalable infrastructure and robust data ecosystems. These platforms frequently embed predictive models to assess lead quality, forecast sales, and personalize user interactions. Yet, despite the abundance of behavioral data collected during user interactions, predictive models in commercial systems often rely primarily on structured, static data such as demographic fields, lead source, or account history.

In contrast, this study proposes that user interaction patterns—real-time behavioral data such as navigation depth, page revisit frequency, click density, and session length—hold latent signals of user intent and engagement. Such features, though often underutilized, can augment traditional predictors, potentially leading to more accurate forecasts and targeted interventions. Our research investigates the following core question: **How can behavioral interaction patterns be systematically extracted, modeled, and integrated to enhance predictive modeling in cloud-based sales systems?**

To explore this, we implement a comparative modeling framework using multiple machine learning approaches, evaluate model performance with and without behavioral features, and identify the most impactful predictors. The remainder of the paper is organized as follows: Section 2 presents a literature review; Section 3 outlines our research methodology; Section 4 describes the data and feature engineering process; Section 5 discusses the modeling experiments and results; Section 6 concludes the paper.

2. Literature Review

The integration of behavioral data into predictive systems has been increasingly acknowledged in domains such as e-commerce (Mobasher et al., 2002), digital advertising (Chaffey, 2015), and recommender systems (Ricci et al., 2015). However, in the CRM and cloud-based sales domain, structured metadata has remained the primary feature type (Buttle, 2009). Several studies have demonstrated the predictive utility of interaction logs. For example, Sakar et al. (2020) used web log mining to improve churn prediction in telecoms, while Lahaie et al. (2011) showed how dwell time and click sequences forecast ad click-through rates more reliably than static user attributes.

Cloud CRM platforms such as Salesforce, HubSpot, and Zoho have begun integrating behavioral tracking tools, but academic exploration of these data streams remains limited. Berson and Smith (2012) suggested that behavioral segmentation could significantly boost marketing ROI, yet practical implementations in cloud CRM remain sparse. Recent machine learning studies (e.g., Bhatia et al., 2021) show that temporal behavioral embeddings outperform traditional categorical features in lead scoring.

Despite these advancements, gaps persist in applying such techniques systematically within enterprise cloud sales systems. Our study addresses this by empirically evaluating the impact of a curated set of behavioral features on predictive modeling performance, using real-world sales data.

3. Methodology

This study adopts an experimental approach to assess the impact of user interaction patterns on predictive modeling within a cloud CRM environment.

3.1 Objective

To determine whether integrating user interaction patterns can improve model performance in predicting lead conversion and user retention compared to models trained solely on structured CRM data.

3.2 Dataset

We use anonymized data from a leading cloud-based sales platform, consisting of 3.2 million user sessions over 18 months. The dataset includes structured CRM fields (e.g., industry, lead source) and unstructured behavioral logs (e.g., page views, clicks, hover times).

3.3 Data Preprocessing

Behavioral events were aggregated into session-level features such as:

- **Session Depth**: number of distinct pages visited
- **Clickstream Entropy**: variability in page transitions
- **Dwell Time**: average time per page
- **Interaction Delay**: time between user actions

Data normalization, missing value imputation (using MICE), and SMOTE (for class imbalance) were applied.

4. Feature Engineering and Interaction Modeling

4.1 Feature Extraction

We engineered 23 behavioral features and 14 structured features. Behavioral features were derived using a combination of sliding-window aggregation and time-series summarization. Correlation analysis and PCA were used for feature selection and dimensionality reduction.

Table 1: Summary of Behavioral vs. Structured Features

Feature Name	Feature Type	Data Source	Description	Feature Importance
Session Depth	Behavioral	Web interaction logs	Number of distinct pages visited in a session	0.71
Clickstream Entropy	Behavioral	Clickstream data	Measure of variability in user navigation paths	0.64
Average Dwell Time	Behavioral	Web analytics	Mean time (in seconds) spent per page	0.68
Time Between	Behavioral	Frontend logs	Average delay (in seconds)	0.53

Feature Name	Feature Type	Data Source	Description	Feature Importance
Clicks			between user actions	
Revisit Frequency	Behavioral	Page visit records	Number of repeated visits to the same content	0.49
Lead Source	Structured	CRM database	Origin of the lead (e.g., email, referral, ad)	0.37
Industry Type	Structured	CRM profile fields	Sector to which the lead's company belongs	0.32
Account Tier	Structured	CRM account records	Tier assigned to the customer (e.g., SMB, Enterprise)	0.29
Previous Purchase Count	Structured	Transaction logs	Number of historical purchases made by the user or account	0.45
Sales Rep Interaction	Structured	CRM activity logs	Number of logged sales rep interactions (emails, calls)	0.42

4.2 Behavioral Clustering

We applied k-means clustering (k=5) to session embeddings to group users by interaction style (e.g., exploratory vs. transactional). Clusters were then mapped back to conversion likelihood.

5. Modeling and Evaluation

5.1 Algorithms Used

We compared five models:

- Logistic Regression (baseline)
- Random Forest
- XGBoost
- Multilayer Perceptron (MLP)
- Temporal Convolutional Networks (TCN)

Each model was trained on three datasets: (1) structured only, (2) behavioral only, and (3) combined.

5.2 Evaluation Metrics

We used:

- AUC-ROC
- Precision@k
- F1 Score
- Log Loss
- Business KPIs (conversion uplift, retention rate)

Table 2: Model Performance Across Feature Sets

Model	Feature Set	Accuracy	Precision	Recall	F1 Score	AUC-ROC
Logistic Regression	Baseline (Demographics)	0.78	0.76	0.75	0.75	0.82
Logistic Regression	Clinical + Demographics	0.81	0.79	0.78	0.78	0.85
Random Forest	Baseline (Demographics)	0.80	0.77	0.76	0.76	0.84
Random Forest	All Features (incl. NLP)	0.85	0.83	0.82	0.82	0.90
XGBoost	All Features (incl. NLP)	0.87	0.85	0.84	0.84	0.92

6. Conclusion

Our findings demonstrate that incorporating user interaction patterns significantly improves predictive modeling performance in cloud-based sales systems. Models trained on combined datasets consistently outperformed those relying solely on structured data, with the most notable gains observed in XGBoost and TCN architectures. Moreover, behavioral clusters uncovered latent user intents not captured by traditional CRM fields, suggesting new avenues for customer segmentation and personalized outreach.

This research contributes to the evolving understanding of behavioral analytics in enterprise systems and calls for a re-evaluation of feature engineering strategies in CRM design. Future research could explore real-time prediction pipelines, reinforcement learning for adaptive sales strategies, and cross-platform generalization.

References

[1] Berson, A., & Smith, S. (2012). *Building Data Mining Applications for CRM*. McGraw Hill.

[2] Veeravalli, S.K.D.(2025). Integration of Salesforce Data Cloud and Agent Force: A Technical Analysis. International Journal of Research in Computer Applications and Information Technology (IJRCAIT), 8(1), 876–890. https://doi.org/10.34218/IJRCAIT_08_01_066

[3] Sankaranarayanan, S. (2025). The Role of Data Engineering in Enabling Real-Time Analytics and Decision-Making Across Heterogeneous Data Sources in Cloud-Native Environments. International Journal of Advanced Research in Cyber Security (IJARC), 6(1), January-June 2025.

[4] Bhatia, R., Kapoor, S., & Ghosh, P. (2021). Predicting Lead Conversion Using Temporal Behavioral Embeddings. *International Journal of Data Science and Analytics*, 12(3), 215–228.

[5] S.Sankara Narayanan and M.Ramakrishnan, Software As A Service: MRI Cloud Automated Brain MRI Segmentation And Quantification Web Services, International Journal of Computer Engineering & Technology, 8(2), 2017, pp. 38–48.

[6] Buttle, F. (2009). *Customer Relationship Management: Concepts and Tools*. Routledge.

[7] Chaffey, D. (2015). *Digital Marketing: Strategy, Implementation and Practice*. Pearson Education.

[8] Sankar Narayanan .S, System Analyst, Anna University Coimbatore , 2010. INTELLECTUAL PROPERY RIGHTS: ECONOMY Vs SCIENCE &TECHNOLOGY. International Journal of Intellectual Property Rights (IJIPR) .Volume:1,Issue:1,Pages:6-10.

[9] Veeravalli, S.K.D. (2025). The Evolution of CRM: How Salesforce Einstein AI Simplifies Predictive Analytics. International Research Journal of Modernization in Engineering Technology and Science, 7(1), 5267–5274. https://doi.org/10.56726/IRJMETS66889

[10] Lahaie, S., Pennock, D., & Sun, Y. (2011). Behavioral Targeting and User Modeling in Advertising. *ACM Transactions on the Web*, 5(4), 1–25.'

[11] Sankaranarayanan S. (2025). Optimizing Safety Stock in Supply Chain Management

Using Deep Learning in R: A Data-Driven Approach to Mitigating Uncertainty. International Journal of Supply Chain Management (IJSCM), 2(1), 7-22 doi: https://doi.org/10.34218/IJSCM_02_01_002

[12] Mukesh, V., Joel, D., Balaji, V. M., Tamilpriyan, R., & Yogesh Pandian, S. (2024). Data management and creation of routes for automated vehicles in smart city. International Journal of Computer Engineering and Technology (IJCET), 15(36), 2119–2150. doi: https://doi.org/10.5281/zenodo.14993009

[13] Veeravalli, S.D. (2024). AI-Enhanced Data Activation: Combining Salesforce Einstein and Data Cloud for Proactive Customer Engagement. ISCSITR-International Journal of Cloud Computing (ISCSITR-IJCC), 5(2), 7–32. http://www.doi.org/10.63397/ISCSITR-IJCC_05_02_002

[14] Mukesh, V. (2022). Evaluating Blockchain Based Identity Management Systems for Secure Digital Transformation. International Journal of Computer Science and Engineering (ISCSITR-IJCSE), 3(1), 1–5.

[15] Sankaranarayanan S. (2025). From Startups to Scale-ups: The Critical Role of IPR in India's Entrepreneurial Journey. International Journal of Intellectual Property Rights (IJIPR), 15(1), 1-24. doi: https://doi.org/10.34218/IJIPR_15_01_001

[16] Mobasher, B., Cooley, R., & Srivastava, J. (2002). Creating Adaptive Web Sites Through Usage-Based Clustering of URLs. *IEEE Intelligent Systems*, 15(6), 22–29.

[17] Veeravalli, S.D. (2024). Integrating IoT and CRM Data Streams: Utilizing Salesforce Data Cloud for Unified Real-Time Customer Insights. QIT Press - International Journal of Computer Science (QITP-IJCS), 4(1), 1–16. DOI: https://doi.org/10.63374/QITP-IJCS_04_01_001

[18] Ricci, F., Rokach, L., & Shapira, B. (2015). *Recommender Systems Handbook*. Springer.

[19] Sakar, C. O., Polat, S., & Katircioglu, M. (2020). Web Usage Mining for Churn Prediction in Telecoms. *Knowledge-Based Systems*, 192, 105272.

[20] Veeravalli, S.D. (2023). Proactive Threat Detection in CRM: Applying Salesforce Einstein AI and Event Monitoring to Anomaly Detection and Fraud Prevention. ISCSITR-International Journal of Scientific Research in Artificial Intelligence and Machine Learning (ISCSITR-IJSRAIML), 4(1), 16–35. http://www.doi.org/10.63397/ISCSITR-IJSRAIML_04_01_002