IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

CRM-Driven AI for Customer Retention: Model Deployment and Chatbot Integration

Chetna Dalvi and Pratiksha Bhapkar
Department of Computer Applications
S. M. Joshi College, Hadapsar, Pune, India

Abstract: *In the realm of customer relationship management (CRM), the integration of artificial intelligence (AI) has become a transformative force, offering innovative solutions to enhance customer retention strategies. This study explores the deployment of supervised machine learning algorithms, specifically Random Forest and XGBoost, for the task of churn prediction. Simultaneously, it examines the role of rule-based chatbot frameworks in facilitating real-time customer engagement. Through empirical experimentation using synthetic CRM datasets, the research demonstrates that the proposed hybrid system exhibits robust predictive performance and presents significant implications for automated, data-driven CRM systems*

Keywords: Customer Relationship Management, Churn Prediction, Artificial Intelligence, Random Forest, XGBoost, Chatbot Systems, Predictive Modeling

I. INTRODUCTION

Customer retention continues to pose a strategic challenge in the contemporary digital economy. Organizations increasingly rely on CRM platforms not only to store customer data but also to derive actionable insights through predictive analytics. AI, particularly machine learning (ML), has emerged as a key enabler in transforming CRM from a reactive system into a proactive decision-support mechanism. This study contributes to the theoretical foundation of CRM enhancement by integrating classification-based ML models with interactive AI agents, aiming to augment both predictive capacity and customer communication processes.

II. LITERATURE REVIEW

The application of statistical and machine learning techniques to CRM has received considerable academic attention. Traditional models, such as logistic regression and decision trees, have been widely deployed to forecast customer churn. Ensemble techniques, notably Random Forest and gradient boosting (XGBoost), have gained prominence for their higher accuracy and resilience against overfitting. Concurrently, chatbot systems—rooted in natural language processing (NLP) and decision-tree scripting—have evolved as efficient intermediaries in digital service delivery. However, limited studies have synthesized these two domains into a unified framework. This research addresses that gap by proposing a model that simultaneously predicts churn and engages users via chatbot responses informed by predictive outputs.

III. METHODOLOGY

The methodological approach adopted in this study follows a structured CRISP-DM-inspired framework. First, synthetic CRM data—emulating customer behavior, interaction logs, and purchase patterns—underwent preprocessing including feature encoding and normalization. Supervised learning models (Random Forest and XGBoost) were selected due to their ensemble learning characteristics, which enhance generalization. The models were trained and validated using an 80-20 data split. In parallel, a chatbot interface was developed using Python, designed to query and interpret model outputs. The chatbot logic adhered to a rule-based structure, enabling predefined conversational flows tied to CRM-relevant intents such as sales inquiries, churn risk alerts, and interaction history.

DOI: 10.48175/568

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

IV. IMPLEMENTATION

The implementation environment comprised Python 3.8 and a suite of scientific libraries, including pandas, numpy, scikit-learn, and xgboost. The CRM dataset was ingested, processed, and encoded using one-hot techniques to handle categorical variables. Feature-target separation was performed to train the models. Hyperparameter tuning was conducted via grid search to optimize model performance. The chatbot was constructed with conditional logic to simulate conversation flow, returning real-time CRM summaries, sales metrics, and churn predictions based on the underlying model outputs.

V. EVALUATION AND RESULTS

Model performance was evaluated using accuracy, precision, recall, and the ROC-AUC score. Random Forest yielded a classification accuracy of approximately 85% and an ROC-AUC of 0.89, while XGBoost closely followed with an ROC-AUC of 0.87. The confusion matrix indicated a balanced trade-off between false positives and false negatives, suggesting both models are suitable for CRM deployment. The chatbot effectively integrated with the prediction pipeline, delivering accurate churn risk summaries and relevant customer engagement metrics during test simulations.

VI. DISCUSSION

The results affirm the practical value of combining predictive modeling with AI-based interaction systems in CRM. Ensemble methods such as Random Forest and XGBoost not only exhibit superior classification capabilities but also integrate seamlessly with rule-based chatbots for interactive insights. While the models demonstrated high predictive capacity, further optimization could enhance sensitivity towards minority class detection. The chatbot, although functional, remains rule-driven; future iterations may explore machine learning-based dialogue systems for greater responsiveness and adaptability.

VII. CONCLUSION

This study demonstrates that AI-driven CRM systems combining ensemble-based churn prediction with chatbot-assisted engagement can substantially elevate business intelligence and operational efficiency. The synergy between predictive analytics and conversational interfaces facilitates real-time decision-making and customer communication. Future research should explore more dynamic chatbot frameworks, such as those incorporating NLP and reinforcement learning, to further enhance CRM automation.

REFERENCES

[1]. Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.
[2]. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD.
[3]. Scikit-learn documentation. https://scikit-learn.org
[4]. XGBoost documentation. https://xgboost.readthedocs.io
[5]. Python Software Foundation. https://python.org