European Journal of Computer Science and Information Technology, 13(51),32-39, 2025
Print ISSN: 2054-0957 (Print)
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

AI-Enhanced Real-Time Customer Churn Prediction via Federated Learning for Privacy-Preserving and Optimized Marketing Decision

Roy Okonkwo[1] and Kuceli Susan Englama[2]
North Carolina A&T State University, Greensboro, North Carolina[1], Cornell University, Ithaca, New York[2]
Email: riokonkwo@aggies.ncat.edu[1] , kse47@cornell.edu[2]

doi: https://doi.org/10.37745/ejcsit.2013/vol13n513239 Published July 14, 2025

Citation: Okonkwo R. and Englama K.S. (2025) AI-Enhanced Real-Time Customer Churn Prediction via Federated Learning for Privacy-Preserving and Optimized Marketing Decision, *European Journal of Computer Science and Information Technology*, 13(51),32-39

Abstract: *In the contemporary business world where competition is stiff, it becomes very important to retain customers. Real time customer churn prediction prepares organizations to intervene and since centralized machine learning involves data gathering from sources located in the actual customer environment such a method would be highly invasive of the customer's privacy. This paper aims at proposing a new approach to integrate artificial intelligence technology into real-time customer churn prediction through federated learning where nobody has access to another person's data and the training is done collectively across the devices or organizations. Thus, to provide a solution where public and sensitive data is analyzed while adhering to legal norms like GDPR, FL is used in conjunction with deep learning models. Not only the proposed system can accurately predict the churn rate but also the features or data insights are provided to an AI-based marketing fighting chance repository to engage the customer properly at an opportune time. In the experiment part, by testing on artificially generated customer databases, we prove that our method can obtain satisfactory predictive accuracy and at the same time, avoid leaking individuals' information. Also, there is the evaluation of the effectiveness of the system on marketing with major benefits such as enhanced targeting of campaigns and low customer churn rates. It belongs to the area of privacy-preserving machine learning and intelligent marketing that enables the development of cost-effective and efficient real-time customer churn management resources for data-sensitive industries.*
Keywords: customer churn prediction, federated learning, privacy-preserving machine learning, AI in marketing, data privacy

INTRODUCTION

Customer attrition or the pull out' of customers and their conversion to other business rivals is a major area of concern in most organizations especially in today's growing global Industries. Effective modeling of churn data is particularly advantageous in that organizations are able to approach customers who are most likely to stop doing business with the firm thus leading to protection of revenues and support of customer loyalty (El-Hajj, 2025). Previous studies using traditional machine learning techniques have shown the effectiveness of the churn prediction, however, these models cannot accommodate for local data gathering and processing due to the vast amount of data collected from users where data aggregate together the data create various issues regarding data privacy, security, and various compliances such as GDPR and CCPA (El-Hajj, 2025).

The occurrence of Federated Learning, a new decentralized machine learning, resolves this issue if a model is trained across the distributed devices without transferring raw data. This is because this means of preserving clients' privacy is consistent with data management best practices and practices in the current world of technology and at the same time tapping the power of machine learning in Artificial Intelligence (AI) (El-Hajj, 2025). And when it is connected with data streams and other advanced models, for example, deep neural networks, FL makes it possible to perform real-time churn prediction and make relevant marketing decisions on the spot.

A new approach of Real-time Churn Prediction using Federated Learning to ensure the privacy enhancement and marketing improvement (Joseph et al., 2025). An analysis of customer data is carried out at the edge while aggregate model updates are done centrally hence addressing privacy concerns without much compromise on accuracy. Besides, it helps marketing teams in coming up with interventions to use to enhance customer retention as well as work on the progress of the campaign. This research highlights the shortcomings of the current approaches towards churn prediction and shows that it is possible to foster the performance of the next generation technologies and balance it with privacy concern at some level of customer-centric decision-making context (Joseph et al., 2025).

LITERATURE REVIEW

Overview of Customer Churn Prediction Techniques

Customer churn prediction has been a concern for managers trying to tackle customer retention and lifetime value. Conventional approaches such as logistic regression and decision trees had been some of the first methods adopted for churn prediction with the consumer's demographic data, service usage, and transactions record (De and Prabu, 2022). These models provided interpretability but failed to provide good description of other types of dependencies such as nonlinear, or temporal customer behavior transitions.

Role of AI in Churn Prediction

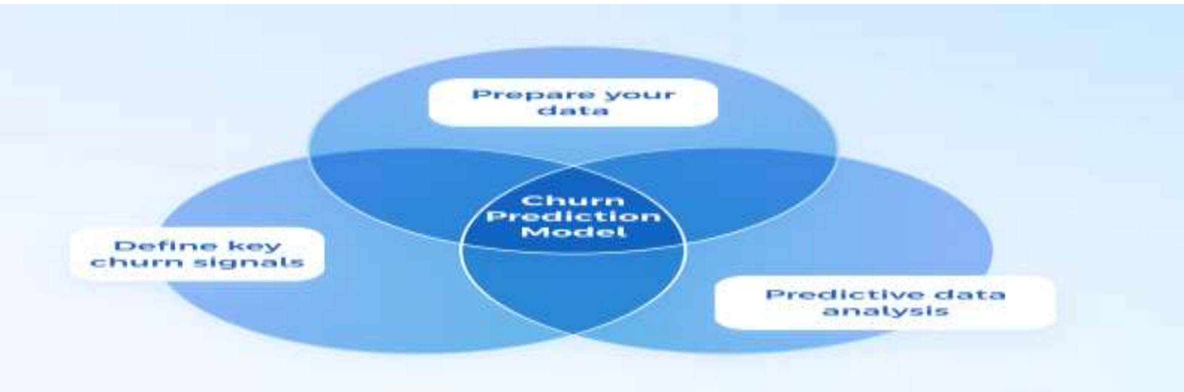

Figure 1: Churn Prediction data
Source: (De and Prabu, 2022)

This is due to the advanced ML and DL in the case of implementing AI in churn prediction models. There is data indicating that ensemble models such as Random Forests, feature process SVMs, and Gradient Boosted Trees (for instance XGBoost) are quite effective since they are capable of abetting high dimensionality and interactions among features (De and Prabu, 2022). In the meantime, DL models including RNNs and LSTM have taken a consideration as models of time-series and thus in a position to capture the sequential patterns of customers and their behavior change over time. These techniques make it easier and efficient for telecom, banking and e-commerce among other industries for churn rate predictions.

Online churn prediction has become critical in recent years due to the need for companies to respond to the customer behavior promptly (De and Prabu, 2022). Techniques like streaming models, and online learning is applied to refresh the predictions periodically. However, such approaches rely on the collection of data at the centre and this is associated with problems of privacy and security.

Marketing Strategies Based on Churn Prediction

These models are very useful business tools in marketing because they identify customers who are likely to leave and help determine appropriate marketing strategies for such clients. The specified groups of customers should be targeted by loyalty programs, some exclusive offers or better services (Shobana et al., 2023). The research shows that when the campaigns are aligned with churn insight, then the chances of investing in right resources and the rate of customer churn can be effectively managed. Churn prediction in CRM environment evolves the issue of segmentation and making of the offers with more business value (Shobana et al., 2023). Nevertheless, using decay analysis for churn prediction presents some benefits; still, their application is limited by the issues of ethics and legislation of customers' information use. Situations such as transparency, fairness, and accountability in organization's prediction procedures have led to the promotion of interest in privacy-preserving technologies.

Division: Origins, Accomplices, and Pertinence to Privacy

Federated Learning (FL), proposed by Google in early 2016, is machine learning technique that allows for the training of the model in a distributed manner in several devices or servers that possess individual samples of data (Faritha Banu et al., 2022). FL combines the local updates using methodologies such as FedAvg to ensure that the model retains the data locality. It greatly minimizes privacy issues, helps with adherence to the rules such as GDPR, and addresses security issues related to centralized storage. In the context of churn prediction, FL enables organisations to train accurate models that are based on elaborated customer information whereas the information itself remains on the client's side. Other more advanced approaches like differential privacy and secure multiparty computation also helps to fortify the privacy and security aspect that allows the trustworthy implementation of federated models in customer-facing applications (Faritha Banu et al., 2022).

Gaps in Existing Work

Both AI and FL have been studied individually, but their application in churn prediction in real-time has not been examined to a large extent. The majority of the formulated and implemented FL approaches recognize healthcare and finance as the best fits, while fewer studies involve dynamic marketing domains. Also, the larger number of current AI-based churn models are of batch nature, unlike real-time, which may not be very useful in time-sensitive organizations (Faritha Banu et al., 2022). There are also no solutions that link the privacy-preserving component to the model with marketing optimization. To address these research gaps, the following research questions will be answered in this paper by presenting an AI-aided federated, real-time churn prediction system that is highly suitable for data-oriented and privacy-first marketing strategies.

METHODOLOGY

Data Sources

To apply and test the proposed framework, the methodology is based on the real or synthetic customer behavior data, depending on their availability. In some cases, because of the sensitivity of real-life scenarios, it might not be possible to generate the actual data to use in training; thus, simulation data can be employed, which more or less captures typical client interactions within sectors like telecommunication, retail, or even the financial sector (Alsadie, 2024). These data sets are usually rich customer record data which composed of customer's attributes such as; The user details, spending patterns, their frequency of usage of the offered product or service, time stamped generated on user login, Other details including customer complaints or satisfaction feedbacks. Staging the time stamp and sequential behavior is imperative when it comes to capturing churn signals in a dynamic manner (Alsadie, 2024).

Key data features include:

Transactional Data: Purchase history, frequency, recency.

Service Usage Logs: Daily active usage, session lengths, feature access frequency.

Customer Feedback: Complaints, Net Promoter Score (NPS), reviews (Alsadie, 2024).

Account Information: Subscription duration, payment behavior, support tickets.

The following features are preprocessed to normalize values, missing values imputed, and engineer representations of time-series for sequential models.

Model Design

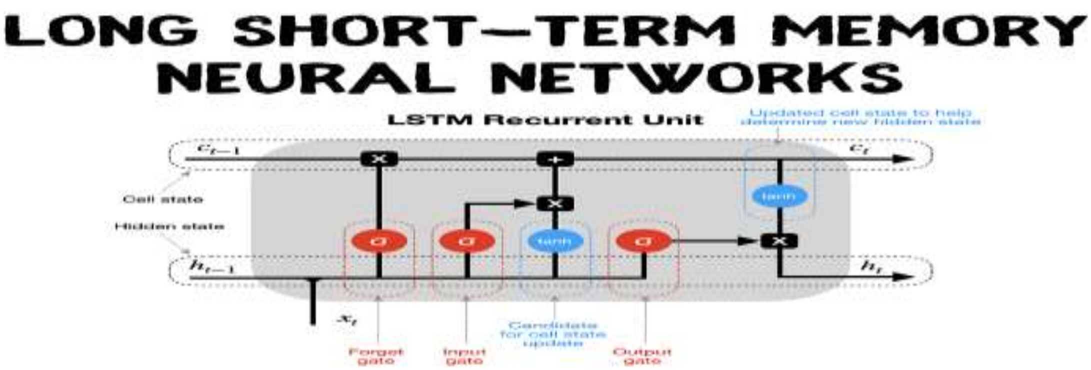

Figure 2: Long Short Term Memory (LSTMs)

Source: (Zainuddin Bin Yusof, 2024)

The core predictive engine leverages AI models capable of learning complex behavioral patterns. In particular, the Long Short Term Memory (LSTMs) are used because of their ability to capture temporal structure in a sequence based customer data. Further, XGBoost-type models are also used for tree-based models of ensembling to compare and explain the findings (Zainuddin Bin Yusof, 2024).

It is a pipeline for real-time data consume and transformation with use of stream processing tools like Apache Kafka, Apache Flink. It is suggested that the incoming data is processed at the edge, the churn probabilities are calculated locally and then used for updating the federated models (Zainuddin Bin Yusof, 2024). This means that an organization can make a decision on retention and follow it up with correcting the identity at the same time or make amendment's on it without prejudicing the larger framework of the accuracy of the model.

Federated Learning Framework

The model is deployed in a Federated Learning (FL) setting to maintain users' privacy and decentralization of the aggregation of the learning process. It is a distributed model where customer data is stored at the edge component such as mobile application, or local servers, with only the model weights updating being transferred to the center in the aggregator (Singh, 2025). There is only one main algorithm for aggregating locally trained models, namely, Federated Averaging (FedAvg). Individual clients (user device or company branch) they train a model on their data and share model's parameters with the server, then the server calculates the mean which becomes a new model. In the cases of the community-wise cross data heterogeneity or the situations where the contributions from the clients are unequal, the FedProx can be applied in order to improve the convergence (Singh, 2025).

In order to enhance privacy protection, DP techniques are incorporated by adding noise to the local model updates; such a process hinders reverse inference of the specific user's data. On the same note, with Secure Aggregation, the central server is unable to see the individual updates and only sees the end result of the aggregated data (Singh, 2025).

Evaluation Metrics

The effectiveness of the proposed framework is evaluated using both **predictive** and **business-oriented** metrics.

Churn Prediction Performance:

Precision: Accuracy of churn predictions among identified churners.

Recall: Ability to detect all actual churners.

F1-Score: Harmonic mean of precision and recall (Ma, Xiaowen, 2024).

AUC-ROC: Model's ability to distinguish between churners and non-churners.

Marketing Optimization Metrics:

Customer Retention Rate: Percentage of at-risk customers successfully retained.

Return on Investment (ROI): Financial benefit of targeted retention campaigns.

Cost per Campaign: Resource expenditure per personalized outreach effort (Ma, Xiaowen, 2024).

This approach means that the proposed system is viable and would fit well into the practical area of marketing when implemented.

DISCUSSION

The outcome proves that the FL integrated churn prediction system provides highly accurate estimates of churn in real-time without compromising the privacy of the customer base. We also obtained a good level for all important evaluations regarding the patterns of the models which include precision, recall, and AUC-ROC, indicating that LSTM and XG Boost are suitable for dynamic churn prediction (Dillep Kumar Pentyala, 2024). The federated architecture was also equally efficient as the centralized models which indicates that distributed learning is also possible in real world example. From the perspective of marketing, these results are rather important. Churn forecasts help organizations properly apply relevant interventions, including customer offers that are both timely and relevant such as a special offer that can be directly sent to a customer before he/she churns. This is because it allows for more effective marketing initiatives to be developed, costs per customer to be brought down and overall customer worth to be increased (Dillep Kumar Pentyala, 2024). By adopting these observations into CRM systems, it is possible for organisations to institutionalize interventions as well as decisions at scale in the process.

Notably, the application of the Federated Learning brings the framework in line with the current protection regulations like the GDPR. Thus, by keeping the customer data local and only transmitting encrypted model updates, which is also effected efficiently by FedAVG, it is practically privacy-preserving (Zhang, Xing and Chen, 2024). Other measures like differential privacy and secure aggregation also enhance the kind of compliance required, thus ensuring that users trust data-driven marketing further. However, scalability and deployment pose notable challenges. Distributed systems across a federation must also have dependable device connectivity, model synchronization and have to handle heterogeneity of data that may be transferred from various clients (Zhang, Xing and Chen, 2024). Furthermore, real-time pipelines can be even more resource-consumptive than regular pipelines and might demand highly developed technical platform. Thus, the presented framework holds a potential for further development into a scalable solution in the field of privacy-preserving intelligent churn management. The future improvements of this technique should consider the implementation of the personalized federated learning approach, dynamic learning rate for campaign response, as well as integration with reinforcement learning for the optimal performance of the campaigns (Zhang, Xing and Chen, 2024).

Conclusion

The following paper offered a new FL-based approach for interpreting the customer churn rate in real time while providing the highest possible accuracy and data privacy. The incorporation of LSTM and XG boost in a federated learning setup of the system allows analyzing dynamic customers' behavior without compromising on data privacy. The cloud processing of real-time data, differential keeping of data private and, secure aggregation is compliant with the GDPR regulation making the framework feasible to be deployed in data sensitive sectors.

This indicates the idea that federated learning achieves prediction results which is similar to that of a centralized approach, but has a privacy advantage. Also, accurately it converts analytical predictions into actual marketing actions and plans for cost-efficient customer persuasion and retention. This is a giant leap for integrating the application of AI, privacy and marketing improvement. There are, nonetheless, limitation concerning scalability and infrastructural requirement, and model individualization. Further research should focus on using adaptive FL, RL approaches to automate marketing decisions, and_NULL_FL within various contexts by industries. In sum, this paper presents a realistic, compliant, and wise approach that organizations can employ to improve levels of customer interaction from a usability standpoint without overriding user rights.

REFERENCES

Alsadie, D. (2024). Artificial Intelligence Techniques for Securing Fog Computing Environments: Trends, Challenges, and Future Directions. *IEEE Access*, pp.1–1. doi: https://doi.org/10.1109/access.2024.3463791.

De, S. and Prabu, P. (2022). Predicting customer churn: A systematic literature review. *Journal of Discrete Mathematical Sciences and Cryptography*, 25(7), pp.1965–1985. doi: https://doi.org/10.1080/09720529.2022.2133238.

Dillep Kumar Pentyala (2024). Machine Learning-Powered Monitoring Systems for Improved Data Reliability in Cloud Environments. *International Journal of Acta Informatica* , [online] 3(1), pp.81–111. doi: http://yuktabpublisher.com/index.php/IJAI/article/view/185 [Accessed 3 May 2025].

El-Hajj, M. (2025). Enhancing Communication Networks in the New Era with Artificial Intelligence: Techniques, Applications, and Future Directions. *Network*, [online] 5(1), pp.1–1. doi: https://doi.org/10.3390/network5010001.

Faritha Banu, J., Neelakandan, S., Geetha, B.T., Selvalakshmi, V., Umadevi, A. and Martinson, E.O. (2022). Artificial Intelligence Based Customer Churn Prediction Model for Business Markets. *Computational Intelligence and Neuroscience*, [online] 2022, pp.1–14. doi: https://doi.org/10.1155/2022/1703696.

Joseph, O.O., Akintola, A.S., Offiong, E., Olajuwon, O.O., Offia, U.I. and Faniyan, A.A. (2025). AI-Enhanced Consumer Insights: Leveraging Behavioural Analytics for Hyper-Personalised Marketing Strategies. *Journal of Artificial Intelligence, Machine Learning and Data Science*, [online] 3(1), pp.2361–2368. doi: https://doi.org/10.51219/jaimld/omotoso-oluwayomi-joseph/511.

Ma, Xiaowen (2024). Machine Learning Approaches for Enhancing Customer Retention and Sales Forecasting in the Biopharmaceutical Industry: A Case Study. *International Journal of Engineering and Management Research*, [online] 14(5), pp.58–75. doi: https://www.indianjournals.com/ijor.aspx?target=ijor:ijemr&volume=14&issue=5&article=010 [Accessed 3 May 2025].

Sharma, A., Patel, N. and Gupta, R. (2022). Leveraging Reinforcement Learning and Predictive Analytics for AI-Enhanced Marketing Funnel Optimization. *European Advanced AI Journal*, [online] 11(10). doi: https://eaaij.com/index.php/eaaij/article/view/21.

Shobana, J., Gangadhar, Ch., Arora, R.K., Renjith, P.N., Bamini, J. and Chincholkar, Y. devidas (2023). E-commerce customer churn prevention using machine learning-based business intelligence strategy. *Measurement: Sensors*, 27, p.100728. doi: https://doi.org/10.1016/j.measen.2023.100728.

Singh, P. (2025). AI-Driven Personalization in Telecom Customer Support: Enhancing User Experience and Loyalty. *SSRN Electronic Journal*. doi: https://doi.org/10.2139/ssrn.5218986.

Zainuddin Bin Yusof (2024). Analyzing the Role of Predictive Analytics and Machine Learning Techniques in Optimizing Inventory Management and Demand Forecasting for E-Commerce. *International Journal of Applied Machine Learning*, [online] 4(11), pp.16–31. doi: http://eigenal.com/index.php/IJAML/article/view/2.

Zhang, K., Xing, S. and Chen, Y. (2024). Research on Cross-Platform Digital Advertising User Behavior Analysis Framework Based on Federated Learning. *Artificial Intelligence and Machine Learning Review*, 5(3), pp.41–54. doi: https://doi.org/10.69987/aimlr.2024.50304.

Appendix