www.ijarp.com ISSN 2456-9992 Page: 01-07

International Journal Advanced Research Publications

Review Article

Volume: 02
Issue: 04

A DEEP LEARNING-BASED DATA MINING FRAMEWORK FOR CUSTOMER CHURN PREDICTION WITH INCREMENTAL LEARNING

M. Yamunadevi*[1], Dr. P. Ponmuthuramalingam[2]

[1]Research Scholar, Ph.D Computer Science, LRG Govt. Arts College for Women, Tirupur, Tamil Nadu, India.

[2]Associate Professor, Department of Computer Science, LRG Govt. Arts College for Women, Tirupur, Tamil Nadu, India.

Article Received: 28 February 2026, Article Revised: 18 March 2026, Published on: 08 April 2026

***Corresponding Author: M. Yamunadevi**

Research Scholar, Ph.D Computer Science, LRG Govt. Arts College for Women, Tirupur, Tamil Nadu, India.

DOI: https://doi-doi.org/101555/ijarp.7264

ABSTRACT

Customer churn prediction remains a critical challenge in data mining and business intelligence, as retaining existing customers is more cost-effective than acquiring new ones. Traditional methods often fail to model complex nonlinear relationships in high-dimensional behavioral data and necessitate complete retraining when new data is introduced. This paper introduces a hybrid deep learning framework that combines Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks with an incremental learning mechanism to enable adaptive churn prediction. The proposed model effectively captures both spatial and temporal patterns from structured demographic data and sequential behavioral logs, while dynamically updating model parameters without requiring full retraining. Experimental results show enhanced predictive accuracy, scalability, and reduced computational costs compared to traditional models. The framework offers an efficient solution for real-time churn prediction in dynamic environments.

KEYWORDS: Data Mining, Deep Learning, Customer Churn Prediction, Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), Incremental Learning, Artificial Intelligence.

I. INTRODUCTION

Companies like Amazon, Netflix, and Vodafone are now focusing a lot on keeping their customers. When a customer stops using a service or subscription, it's called churn. Being able to predict when this might happen helps businesses come up with better ways to bring customers back. In the past, people used methods like logistic regression and decision trees to predict churn. But customer data is really complicated—it has many different parts, changes over time, and isn't easy to understand. These challenges make older models less effective. Newer methods, like deep learning models such as CNN and LSTM, are better at handling these complex patterns. However, most systems today use batch training, which makes them less flexible when customer data changes. To fix this, this paper presents a new method that uses CNN and LSTM along with incremental learning. This allows the system to keep learning and adapting continuously without needing to start training from the beginning every time.

II. LITERATURE REVIEW

Customer churn prediction has garnered significant attention with the advancement of data mining and deep learning techniques. Traditional machine learning models such as logistic regression and decision trees have been widely used due to their simplicity; however, they are limited in capturing complex and dynamic customer behavior patterns [1]. To address these limitations, deep learning approaches have been increasingly adopted, particularly Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, which are capable of extracting high-level features and modeling temporal dependencies in sequential data [1], [2]. Studies show that LSTM and recurrent models are highly effective in analyzing time-dependent customer activity, while CNN models efficiently capture hidden patterns in large-scale behavioral datasets [2], [7]. Furthermore, hybrid and ensemble models combining multiple techniques have demonstrated superior performance compared to standalone models, with feature engineering and data quality playing a critical role in improving prediction accuracy [3]. The integration of big data technologies such as Hadoop and Spark has enhanced the scalability and applicability of churn prediction systems in real-world environments [4]. In addition, foundational works on deep learning provide strong theoretical support for designing advanced architectures, including CNN and LSTM-based hybrid models [10]. Practical implementations and real-world applications further highlight the effectiveness of deep learning models in handling large datasets and complex prediction tasks [8], [9]. Moreover, incremental learning techniques have emerged as a key advancement, enabling models to adapt continuously to streaming and evolving data without requiring complete retraining, thereby reducing computational cost and improving adaptability [5], [11]. Despite these advancements, challenges such as model interpretability, scalability, and real-time deployment remain open research issues. Overall, the literature indicates that hybrid deep learning frameworks integrated with incremental learning provide a promising and effective solution for developing accurate, scalable, and adaptive customer churn prediction systems.

III. RELATED WORK

Early churn prediction models used statistical and machine learning methods like logistic regression, decision trees, and support vector machines. These methods gave a basic level of performance but had trouble with time-based patterns and complex relationships between features.

More recent studies have used deep learning methods, such as LSTM networks for handling sequences and CNNs for extracting features from customer behavior. Combining CNNs and LSTMs in hybrid models has helped improve prediction accuracy. Still, there are not many approaches that allow models to learn incrementally. Most models need to be completely retrained when new customer data comes in, which makes the process expensive and hard to scale.

IV. SYSTEM MODEL

A. Overview

Customer churn prediction is crucial for industries such as telecom, banking, and subscription-based services, where retaining existing customers is more cost-effective than acquiring new ones. Conventional approaches depend on static data and are unable to reflect changing customer behavior. To tackle this issue, the proposed system presents a hybrid deep learning and data mining framework with incremental learning for adaptive churn prediction. The model combines structured data (including demographics, billing information, and subscription details) with behavioral logs (such as usage patterns, transactions, and interactions) to create a holistic understanding of customer behavior. A CNN is used to extract behavioral patterns, while an LSTM captures temporal dependencies. These features are integrated through a fusion layer to enhance prediction accuracy. Moreover, an incremental learning mechanism allows for continuous model updates without requiring complete retraining, ensuring the system remains adaptable in changing environments.

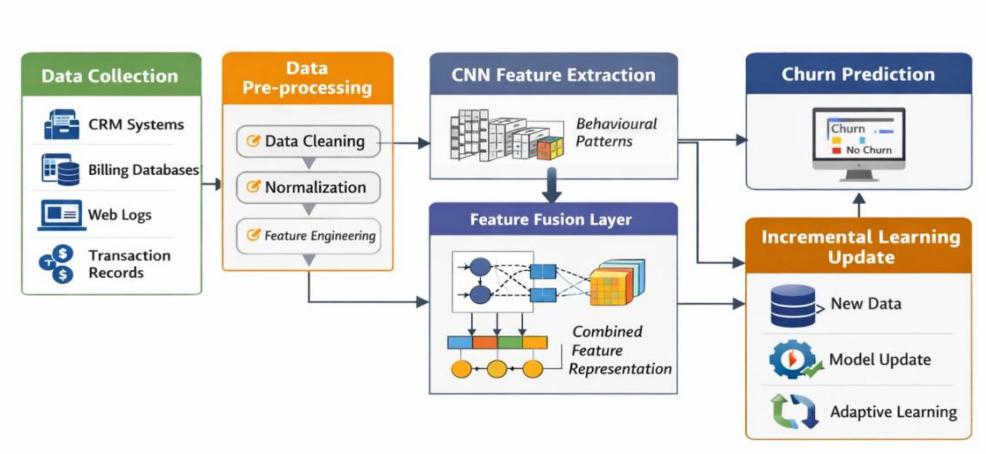

Fig.1. Hybrid Deep Learning – Based Churn Prediction System Architecture.

B. Architecture Description

The proposed framework consists of six modules:

1. Data Collection Module

This part gathers information from different sources like CRM systems, billing platforms, and web logs. It collects details such as customer information, purchase history, and how they interact with services. All this data is kept in one main storage place.

2. Data Pre-processing Module

This module deals with incomplete or messy data.It fixes missing parts, removes unnecessary information, and changes the data into a usable format. Methods used include scaling numbers, converting text into numbers, and creating new features like how often a customer complaints or uses a service. The data is also arranged in order based on time.

3. CNN-Based Behaviour Feature Extraction

A type of neural network called CNN is used to find hidden patterns in customer behaviour, such as a drop in usage or an increase in complaints.It creates simpler, more useful versions of this data.

4. LSTM-Based Sequential Modelling

Another type of neural network, LSTM, is used to understand how customer behaviour changes over time.It can detect trends like decreasing interest or spending.

5. Feature Fusion Layer

This part combines the features from CNN, LSTM, and other structured data. It uses methods like joining data together and applying dense layers to create one complete set of features that can be used to predict if a customer will leave.

6. Incremental Learning Update Module

This allows the model to keep learning from new data without starting over each time. It helps the system stay up-to-date and work better in changing situations.

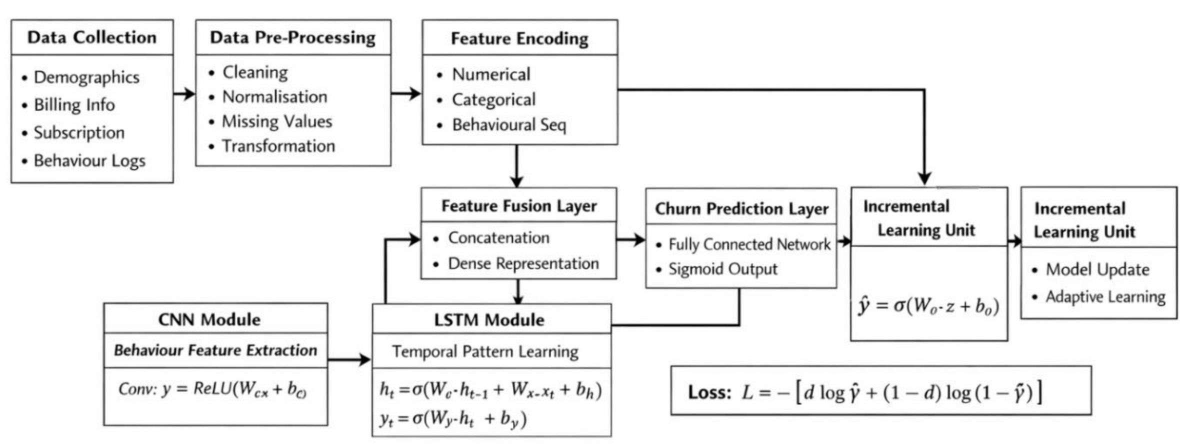

Fig. 2. A Deep Learning – Based Data Mining Framework for Customer Churn Predict Using Incremental Learning.

V. IMPLEMENTATION DETAILS

The churn prediction framework was built using Python because it has many tools that help with data handling and deep learning.

A. Development Environment

The model was made using TensorFlow and PyTorch, which both support training on graphics processing units (GPUs) for faster performance. Important tools used include NumPy for doing math with numbers, Pandas for organizing data, and Scikit-learn for preparing data and checking how well the model works. The system was set up on a workstation that has a GPU to make calculations faster.

B. Model Architecture Implementation

A mix of CNN and LSTM was used in the model. The CNN finds patterns in customer behavior from their activity data, and the LSTM looks at how things change over time. The result goes through layers of dense connections with a sigmoid function to calculate the chance that a customer will leave.

C. Training Configuration

The model uses Binary Cross-Entropy Loss, which is calculated like this: L = - (1/N) * sum over all examples of [y_i * log(p_i) + (1 - y_i) * log(1 - p_i)] The model was trained using the Adam method with a learning rate of 0.001 and a batch size of 64. It was trained for many rounds (epochs), and early stopping was used to stop training when the model wasn't improving anymore, which helps avoid problems like overfitting and makes the model work better on new data.

VI. EXPERIMENTAL RESULTS

The proposed CNN–LSTM model with incremental learning was tested using historical customer data and compared to other baseline models.

A. Evaluation Metrics:

The model's performance was measured using Accuracy, Precision, Recall, F1-Score, and ROC-AUC.

Accuracy is calculated as:

Accuracy=TP+TNTP+TN+FP+FN Accuracy = \ frac {TP + TN} {TP + TN + FP + FN}Accuracy=TP+TN+FP+FNTP+TN

B. Comparative Analysis:

Model	Accuracy	F1-Score
Logistic Regression	82%	0.78
Random Forest	86%	0.83
LSTM	88%	0.85
Proposed CNN–LSTM + Incremental	**92%**	**0.90**

The new model works better than the old methods because it extracts features more efficiently, handles time-based patterns well, and learns in a flexible way, which leads to higher accuracy and better performance in different situations.

VII. DISCUSSION

The results highlight the effectiveness of the proposed CNN–LSTM model in capturing both spatial and temporal patterns in customer behavior, facilitating accurate churn prediction. The incorporation of incremental learning enables continuous model updates with new data, ensuring adaptability to evolving customer trends without requiring full retraining. The framework also lowers computational costs by updating only essential components, making it suitable for large-scale implementation.

Key findings: Enhanced prediction accuracy, reduced computational overhead, improved generalization, and better early churn detection. These advantages support effective proactive customer retention strategies.

VIII. CONCLUSION

This paper introduced a new churn prediction system that mixes CNN and LSTM technologies and uses incremental learning. The model is good at understanding customer behavior and how it changes over time, and it can learn from new data without needing to retrain completely. The test results show that this method works better than older and simpler models in terms of accuracy and F1-score. Overall, this framework offers a scalable, efficient, and smart way to detect customers who might leave early and helps in making better plans to keep customers happy.

IX. REFERENCES

1. S. Wang, H. Liu, and Y. Zhang, "Customer churn prediction using machine learning: A systematic review," IEEE Access, vol. 11, pp. 45231–45250, 2023.
2. A. Kumar and R. Singh, "A review on deep learning approaches for customer retention and churn analysis," IEEE Transactions on Artificial Intelligence, vol. 4, no. 2, pp. 210–225, 2023.
3. M. Chen, X. Zhao, and L. Wu, "Data mining techniques for customer churn prediction: A survey and future directions," IEEE Access, vol. 12, pp. 102345–102367, 2024.
4. P. Sharma and K. Verma,"Recent advances in customer churn prediction using deep learning and big data analytics," IEEE Access, vol. 12, pp. 88765–88789, 2024.
5. R. Gupta, S. Iyer, and P. Nair, "A survey on incremental learning techniques in data mining and predictive analytics," IEEE Transactions on Knowledge and Data Engineering, vol. 36, no. 1, pp. 45–60, 2025.
6. S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
7. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.
8. A. Géron, "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow," 2nd ed., O'Reilly Media, 2019.
9. J. Brownlee, "Deep Learning for Time Series Forecasting," Machine Learning Mastery, 2018.
10. I. Good fellow, Y. Bengio, and A. Courville, "Deep Learning," MIT Press, 2016.
11. H. Bifet and R. Gavaldà, "Learning from time-changing data with adaptive windowing," in Proc. SIAM Int. Conf. Data Mining, 2007, pp. 443–448.