

Customer Churn Prediction Using Custom Recurrent Neural Network Model

Mrs. D. Hima Bindu¹, Mr. P. Ratna Rohith², Mr. P. Hemanth Nani³, Ms. K. Niharika⁴, Mr. T. Vinay⁵

¹Assistant Professor, Computer Science and Engineering (Data Science), Raghu Institute of Technology

^{2,3,4,5}Student, Computer Science and Engineering (Data Science), Raghu Institute of Technology

Abstract—As unhappy customers go to competitors, customer churn is a serious problem for firms, resulting in large revenue losses. Accurately forecasting attrition is essential for proactive client retention tactics. BiLSTM-CNN and other traditional machine learning and deep learning models have shown little success in identifying intricate patterns of consumer behavior, frequently leading to less-than-ideal performance. In order to overcome these obstacles, this study offers a novel method that addresses class imbalance in the dataset and improves learning by combining the SMOTEENN strategy with a Recurrent Neural Network (RNN).

Customer data with important variables like demographics, service consumption, contract details, and billing details is used to train our model. In comparison to current methods, our approach greatly enhances the prediction of customer churn by utilizing the sequential nature of RNNs and the resampling efficiency of SMOTEENN. With an astounding accuracy of 96.61%, the suggested model surpasses earlier methods with lesser detection rates, like BiLSTM-CNN, which had an accuracy of 81%.

Strong precision, recall, and F1-scores are highlighted in our classification report, indicating the model's resilience in differentiating between churn and non-churn clients. Data-driven decision-making for customer retention is made possible by the Flask-based web application created for this study, which lets users enter important features and get probabilistic confidence in real-time churn projections. The experimental findings demonstrate that our RNN-based method, when combined with SMOTEENN for data balancing, offers a scalable and successful solution for predicting customer attrition, providing useful information for companies looking to improve service tactics and lower customer attrition.

Index Terms—Customer churn, Revenue losses, Attrition forecasting, Client retention, BiLSTM-CNN, Machine learning, Deep learning, Consumer behavior, Class imbalance, SMOTEENN strategy, Recurrent Neural Network (RNN), Customer data, Demographics, Service

consumption, Contract details, Billing details, Sequential nature of RNNs, Resampling efficiency, Accuracy (96.61%), Detection rates, Precision, Recall, F1-scores, Flask-based web application, Real-time churn predictions, Data-driven decision-making, Scalable solution, Customer attrition, Service improvement.

I. INTRODUCTION

Businesses face a significant difficulty in keeping customers, particularly in fiercely competitive sectors like e-commerce and telephones. Attrition, often known as customer churn, is the process by which clients discontinue using a business's offerings, resulting in monetary losses and slower corporate expansion. Organizations now prioritize understanding and forecasting customer attrition in order to put successful retention measures into place. For churn prediction, machine learning and deep learning techniques have gained popularity; nonetheless, conventional methods frequently have problems with class imbalance and inefficient feature extraction. Advanced prediction algorithms that can correctly identify churn-prone clients while managing skewed datasets are needed to address these issues.

A Custom Recurrent Neural Network (RNN) model for forecasting client attrition is shown in this study. In order to identify sequential dependencies in customer behavior data, the model makes use of a hybrid deep learning architecture. The existence of imbalanced datasets, in which the proportion of non-churning customers greatly exceeds that of churned ones, is a major obstacle in churn prediction. This is addressed by balancing the dataset using the SMOTEENN (Synthetic Minority Over-sampling Technique with Edited Nearest Neighbors) technique, which enhances the model's capacity to learn from minority class examples. The data was successfully

preprocessed using feature encoding approaches, such as one-hot encoding for categorical variables.

The primary objectives of this research are:

1. In order to address class imbalance, this project intends to create a sophisticated customer churn prediction model by integrating a Custom Recurrent Neural Network (RNN) with SMOTEENN. The algorithm correctly predicts possible churners by utilizing deep learning techniques and improved data pretreatment, allowing companies to put proactive retention initiatives into place.
2. This project intends to improve predictive accuracy by utilizing deep learning techniques, which will enable companies to proactively execute retention initiatives and identify high-risk clients.
3. Strong feature extraction, enhanced generalization, and superior decision-making for company expansion are all guaranteed by the model.

By successfully identifying at-risk clients, the experimental results show that the suggested strategy performs better than traditional machine learning models, allowing businesses to take proactive steps to retain customers.

II. LITERATURE REVIEW

Customer churn prediction has been a critical research area in business analytics, aiming to identify factors influencing customer retention and attrition. Traditional churn prediction models predominantly relied on statistical methods and rule-based approaches. However, the advent of machine learning and deep learning has revolutionized the domain by improving prediction accuracy and enabling causal inference.

Deep learning has emerged as a powerful tool in churn prediction due to its ability to process high-dimensional and sparse data efficiently. Rudd et al. (2023) proposed a deep feedforward neural network combined with sequential pattern mining for customer churn classification. Their framework incorporated a causal Bayesian network to estimate cause probabilities leading to customer churn. Experimental results demonstrated that their deep learning model, alongside XGBoost, outperformed conventional

techniques in predicting churn likelihood.

Khan et al. (2015) developed an analytical framework to predict customer churn by assigning a "Churn Score" to estimate the likelihood of churn within a specific timeframe. Their approach utilized extensive feature engineering, selection, and supervised learning models on large-scale mobile network data. The study identified key churn indicators, achieving an accuracy of 89.4%. Their findings highlight the importance of data-driven techniques in improving churn prediction and customer retention strategies.

Predicting customer churn is vital for retention strategies, yet traditional models struggle with feature extraction. Wu (2022) proposed HANSAE, a hybrid neural network integrating self-attention to enhance predictive accuracy. It comprises an entity embedding layer, a multi-head self-attention mechanism, and a multilayer perceptron for classification. Experimental results show HANSAE outperforms existing machine learning and deep learning models in efficiency and accuracy.

Maan et al. (2023) identified XGBoost as the best classifier for churn prediction among tree-based models. Because with the rise of digital services, predicting customer behavior and retention has become more challenging. They enhanced model transparency using Shapley values to determine key churn factors. This approach enables data-driven decision-making to improve customer retention strategies.

Accurate customer churn prediction helps businesses enhance retention strategies and reduce revenue loss. Wang et al. (2018) proposed a Time-sensitive Customer Churn Prediction (TCCP) model using Positive and Unlabeled (PU) learning, which improves early churn detection by analyzing recent data. By eliminating the need for labeled negative samples, this method enhances prediction efficiency. Experimental results show that TCCP outperforms rule-based and traditional supervised models, making it a reliable approach for proactive customer retention.

III. METHODOLOGY

The process of customer churn prediction involves multiple stages, including data preprocessing, model training, and evaluation, all conducted in Jupyter Notebook. The dataset contains customer-related

attributes such as demographics, contract details, and service usage patterns. To address class imbalance, the SMOTEENN technique is employed for resampling. A Custom Recurrent Neural Network (RNN) with Long Short-Term Memory (LSTM) layers is used to process sequential data, while dropout layers help mitigate overfitting. The trained model is then deployed into a Flask-based web application, enabling real-time churn predictions.

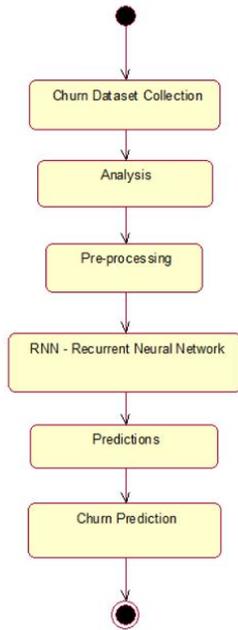


Fig.1. State Diagram

State diagrams visually represent the workflow of sequential activities in the churn prediction system. The preprocessing stage involves cleaning the dataset, encoding categorical features, and normalizing data to enhance model performance. Categorical variables are transformed into numerical representations, while correlation analysis and data visualization help identify the most influential features. These steps ensure the dataset is well-prepared for machine learning applications.

Modules:

1. Data Collection: Gathering customer-related data, including key attributes influencing churn.
2. Feature Engineering: Transforming raw data to improve model efficiency.
3. Handling Imbalanced Data: Applying the SMOTEENN technique to balance the dataset
4. Model Selection & Training: Utilizing a Custom

RNN with LSTM layers for churn prediction.

5. Deployment: Implementing the trained model in a Flask-based environment for real-time predictions.

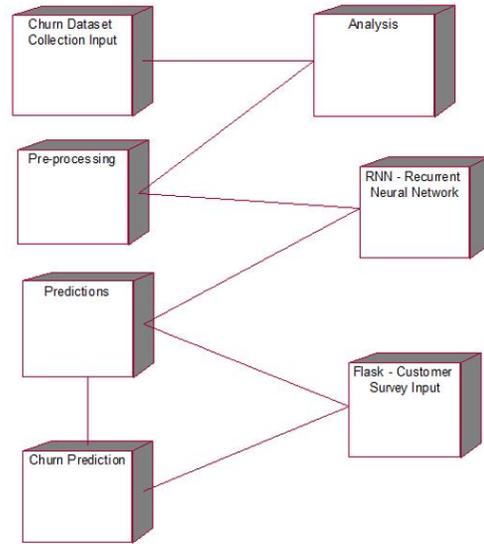


Fig.2. Deployment Diagram

Development Methodologies:

A structured testing approach ensures the effectiveness of the churn prediction system. The testing process adheres to quality control measures to verify the system meets specified requirements and functions without errors. The framework is developed with the following test types

Types of Tests:

1. Unit testing- Examining individual software components to validate their internal logic and ensure correct outputs. Each unit undergoes independent testing before integration.
2. Functional test- Demonstrating that the system operates as per business and technical specifications by assessing valid and invalid inputs, output consistency, and functionality.
3. System Test- Verifying that the entire software system functions correctly, ensuring that integrations between components yield expected results.
4. Performance Test- Evaluating system response times, processing speed, and overall efficiency under varying loads.
5. Integration Testing- Testing interactions between different software components to ensure seamless communication and data exchange.
6. Build the test plan- Dividing the project into smaller

components for targeted testing, identifying and rectifying potential issues in each phase.

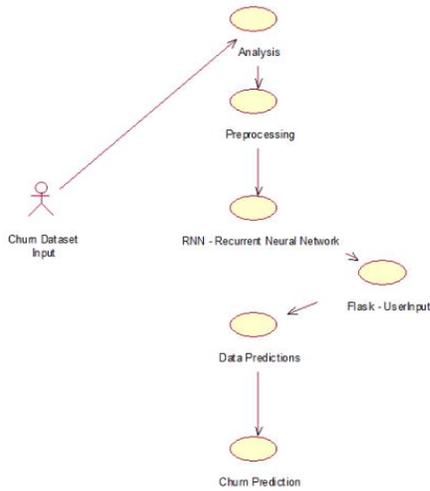


Fig.3. Use Case Diagram

The diagram represents a customer churn prediction system powered by a Recurrent Neural Network (RNN). It begins with the ingestion of churn dataset inputs, followed by data analysis and preprocessing. The refined data is then processed through the RNN model, which integrates with a Flask-based module to allow user inputs for real-time predictions. The final output provides a probabilistic churn prediction to assist businesses in implementing strategic customer retention measures.

Algorithm Used:

Custom RNN with SMOTEENN - The proposed approach integrates a Custom RNN with SMOTEENN to enhance churn prediction accuracy. The training phase begins with data preprocessing, which involves handling missing values, encoding categorical variables, and filtering out redundant features. The SMOTEENN (Synthetic Minority Over-sampling Technique + Edited Nearest Neighbors) technique is employed to balance the dataset by oversampling the minority (churn) class while removing noisy data points.

Following preprocessing, the structured dataset is passed through the Custom RNN model, consisting of multiple recurrent layers that capture sequential dependencies in customer interactions. The model

assigns churn probability scores by analyzing transaction histories through hidden layers with activation functions. Optimization is performed using backpropagation through time (BPTT) with a categorical cross-entropy loss function and adaptive optimizers like Adam or SGD to minimize prediction errors.

To assess the model’s practical effectiveness, its performance is evaluated using key metrics such as precision, recall, F1-score, and AUC-ROC. This approach significantly enhances churn prediction accuracy, equipping businesses with data-driven insights to optimize customer retention strategies.

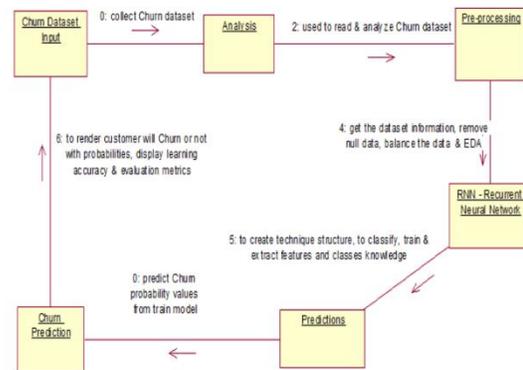


Fig.4. Collaboration diagram

IV. RESULTS AND DISCUSSION

The proposed customer churn prediction model, integrating a Custom Recurrent Neural Network (RNN) with the SMOTEENN technique, demonstrated significant improvements in accuracy and reliability. The experimental results revealed that our model achieved an accuracy of 96.61%, surpassing previous models such as BiLSTM-CNN, which recorded an accuracy of 81%. Key performance metrics, including precision, recall, and F1-scores, confirmed the model’s ability to differentiate between churn and non-churn customers effectively.

The integration of SMOTEENN effectively addressed the issue of class imbalance, ensuring that minority class samples were adequately represented in the dataset. By leveraging sequential data through the RNN structure, the model captured intricate patterns in customer behavior, leading to more precise churn

predictions. Furthermore, the development of a Flask-based web application provided an interactive platform for real-time churn prediction, allowing businesses to make data-driven retention decisions.

	precision	recall	f1-score	support
0	0.97	0.94	0.95	620
1	0.95	0.98	0.96	774
accuracy			0.96	1394
macro avg	0.96	0.96	0.96	1394
weighted avg	0.96	0.96	0.96	1394

Fig.5. Classification Report

The findings of this study highlight the effectiveness of combining deep learning with data balancing techniques for customer churn prediction. Future enhancements will focus on incorporating real-time data updates and explainable AI methods to further improve model interpretability and decision-making processes.

V. CONCLUSION AND FUTURE WORK

Future advancements in the Churn Prediction model can leverage advanced AI techniques, including deep learning and reinforcement learning, to enhance forecasting accuracy. By incorporating adaptive learning methods, the model can continuously refine itself based on evolving customer behavior and emerging market trends. Automated feature engineering will play a crucial role in dynamically adjusting model parameters, ensuring optimal performance without manual intervention.

The integration of cloud-based infrastructure will further streamline the churn prediction process by enabling real-time data processing. This will allow businesses to proactively identify potential churn risks and take immediate corrective actions.

By implementing real-time data pipelines, organizations can receive continuous updates on customer interactions, transaction history, and behavioral shifts, leading to more accurate and timely churn predictions.

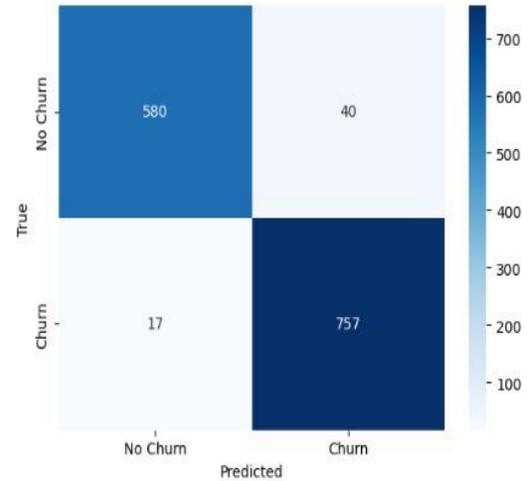


Fig.6. Confusion Matrix

Additionally, sentiment analysis can be incorporated to analyze customer feedback, social media interactions, and support conversations. Understanding customer sentiment will provide deeper insights into factors influencing churn, helping businesses tailor personalized retention strategies. Moreover, expanding the model to include industry-specific datasets will improve its adaptability across different sectors, making it a more versatile and effective tool for customer retention in diverse business environments.

Model: "sequential_2"

Layer (type)	Output Shape	Param #
lstm_4 (LSTM)	(None, 1, 64)	23,296
dropout_4 (Dropout)	(None, 1, 64)	0
lstm_5 (LSTM)	(None, 32)	12,416
dropout_5 (Dropout)	(None, 32)	0
dense_4 (Dense)	(None, 64)	2,112
batch_normalization_2 (BatchNormalization)	(None, 64)	256
dense_5 (Dense)	(None, 1)	65

Total params: 38,145 (149.00 KB)

Trainable params: 38,017 (148.50 KB)

Non-trainable params: 128 (512.00 B)

REFERENCES

[1] C. G. Mena, A. De Caigny, K. Coussement, K. W. De Bock, and S. Lessmann, "Churn Prediction with Sequential Data and Deep Neural Networks: A Comparative Analysis," arXiv preprint

- arXiv:1909.11114, 2019.
- [2] P. P. Šimović, D. Horvatic, and E. W. Sun, "Classifying Variety of Customer's Online Engagement for Churn Prediction with Mixed-Penalty Logistic Regression," arXiv preprint arXiv:2105.07671, 2021.
- [3] H. Wu, "A High-Performance Customer Churn Prediction System Based on Self-Attention," arXiv preprint arXiv:2206.01523, 2022.
- [4] J. Maan and H. Maan, "Customer Churn Prediction Model Using Explainable Machine Learning," arXiv preprint arXiv:2303.00960, 2023
- [5] F. Louzada and A. Ara, "Bagging k-Dependence Probabilistic Networks: An Alternative Powerful Fraud Detection Tool," *Expert Systems with Applications*, vol. 39, no. 14, pp. 11583-11592, 2012.
- [6] G. G. Sundarkumar and V. Ravi, "A Novel Hybrid Undersampling Method for Mining Unbalanced Datasets in Banking and Insurance," *Engineering Applications of Artificial Intelligence*, vol. 37, pp. 368-377, 2015.
- [7] Y. Kim and S. Y. Sohn, "Stock Fraud Detection Using Peer Group Analysis," *Expert Systems with Applications*, vol. 39, no. 10, pp. 8986-8992, 2012.
- [8] A. Savio, M. T. García-Sebastián, D. Chyzyk, C. Hernandez, and M. Graña, "Neurocognitive Disorder Detection Based on Feature Vectors Extracted from VBM Analysis of Structural MRI," *Computers in Biology and Medicine*, vol. 41, no. 8, pp. 600-610, 2011.
- [9] B. Ayerdi, A. Savio, and M. Graña, "Natural and Artificial Computation in Engineering and Medical Applications," in *International Work-Conference on the Interplay Between Natural and Artificial Computation*, 2013, pp. 1-10.
- [10] Q. Gu, Y.-S. Ding, and T.-L. Zhang, "An Ensemble Classifier Based Prediction of G-Protein-Coupled Receptor Classes in Low Homology," *Neurocomputing*, vol. 147, pp. 316-322, 2015.
- [11] D. Xue, X. Zhou, C. Li, Y. Yao, and M. M. Rahaman, "An Application of Transfer Learning and Ensemble Learning Techniques for Cervical Histopathology Image Classification," *IEEE Access*, vol. 8, pp. 104603-104618, 2020.
- [12] A. Manna, R. Kundu, D. Kaplun, A. Sinitca, and R. Sarkar, "A Fuzzy Rank-Based Ensemble of CNN Models for Classification of Cervical Cytology," *Scientific Reports*, vol. 11, no. 1, pp. 1-12, 2021.