www.ijecs.in
International Journal of Engineering and Computer Science
Volume 14 Issue 04 April 2025, Page No. 27022- 27029
ISSN: 2319-7242 DOI: 10.18535/ijecs/v14i04.5055

Customer Churn Prediction using Deep Learning

A Sherill PG Student [1]*, [2] Dr. R. Porkodi Professor

[1] Department of Computer Science Bharathiar University Coimbatore-641046
[2] Department of Computer Science Bharathiar University Coimbatore-641046

Abstract
Customer churn is a common challenge in the banking sector, severely affecting the profitability of financial institutions. Although various strategies have been implemented to address this issue, churn remains a persistent problem. To effectively mitigate this, the use of diverse predictive models is crucial. These models, built using machine learning and deep learning techniques, include methods such as classification, clustering, and hybrid approaches. The models like Artificial Neural Network, Convolutional Neural Network, Recurrent Neural Network, Deep Neural Network, Long Short-Term Memory are compared in this study. Across diverse models, DNN achieved the highest recall of 91%. Among the evaluation metrics, recall ensures to capture the most potential churners.

Keywords: Customer churn, Machine learning, Deep learning, Banking

1. Introduction

In the banking sector, customer churn referring to the rate at which customers close accounts, switch banks, or discontinue specific financial services is a critical issue with broad implications for both individual banks and the industry as a whole. High churn rates are not only disruptive but also costly; they can erode a bank's revenue base, diminish its profitability, and impact long-term growth. Additionally, churn can signal dissatisfaction or competitive pressure, prompting banks to allocate considerable resources toward managing and reducing turnover through customer retention strategies. Due to high customer mobility, individuals can easily move from one bank to another, seeking improved service quality, competitive rates, or personalized products. To prevent churn, banks employ various strategies to enhance customer retention, such as offering attractive rates, superior customer service, and tailored financial products. Despite these measures, high churn persists and poses a major challenge to profitability. Consequently, many banks are now turning to customer churn prediction models as a strategic approach to mitigate these impacts. By leveraging advanced analytics and machine learning, these models enable banks to understand, anticipate, and pre-emptively address factors that may lead customers to leave. Churn prediction models analyze historical and real-time data to identify which customers are most likely to discontinue their relationship with the bank within a specified timeframe. Using inputs such as customer demographics, transaction history, account activity, and satisfaction levels, these models identify patterns associated with churn. This predictive power enables banks to proactively reach out to customers who might be at risk, allowing them to intervene with customized retention strategies.

The financial impact of customer churn extends beyond the immediate loss of revenue from fees and interest. Losing a customer triggers additional costs associated with acquiring new clients, who may not yield the same revenue initially. Onboarding new customers involves marketing expenses, account setup costs, and time invested in cultivating these relationships. Therefore, retaining existing customers is often more cost-effective than acquiring new ones, particularly for banks that aim to maximize the lifetime value of each client. Moreover, high churn rates can damage a bank's reputation and weaken trust, potentially deterring prospective clients from opening accounts. High customer attrition could also indicate underlying issues, such as customer dissatisfaction with services or increased competition, that necessitate strategic adjustments. This places additional pressure on banks to address the root causes of churn, whether by refining service offerings, enhancing digital platforms, or improving customer support. Churn prediction models generate valuable insights that inform a range of strategic decisions, from product development to marketing and customer service enhancements. For example, understanding the behaviors and characteristics of churn-prone customers allows banks to develop products and services that better meet customer needs. Similarly, insights from churn models can guide marketing campaigns, helping banks allocate resources effectively and tailor messages to at-risk customer segments. For a churn prediction model to be effective, it must be built on a comprehensive dataset that accurately captures the nuances of customer relationships with the bank. Relevant data typically includes transaction histories, product usage patterns, service interactions, and demographic details. By leveraging this data, models can gain a holistic view of each customer's engagement, enabling more precise predictions.

The performance of deep learning models depends not only on their overall accuracy but also on their ability to minimize false negatives instances where likely churners go undetected. This is where recall, a metric that measures the proportion of actual churners correctly identified by the model, becomes critical. While precision and accuracy are commonly used metrics in machine learning and deep learning, recall is often emphasized in churn prediction scenarios due to the significant cost of undetected churners. By focusing on recall, businesses can better identify customers at the highest risk of leaving, enhancing the impact of targeted retention strategies and improving customer retention outcomes. DL models, known for their ability to capture complex, non-linear relationships, are especially suited to handle the intricacies of customer behavior data.

2. Literature Review

The prediction models use various data mining techniques. Agrawal, Sanket, et al. [1] highlights recall's importance in customer churn prediction due to the imbalance in churn datasets, where churners are fewer than non-churners. High recall is crucial to identify at-risk customers effectively, enabling proactive retention actions and minimizing revenue loss. Deep learning models, like multi-layer neural networks, help maximize recall by accurately capturing churn patterns, which is more valuable than solely focusing on overall accuracy.

Saha et al. [2] emphasized that recall is crucial for customer churn prediction, especially in the telecom industry where datasets are highly imbalanced. High recall is prioritized because it ensures the model captures a large portion of actual churn cases, allowing companies to identify and address dissatisfied customers who are likely to leave. The proposed ChurnNet model, using attention mechanisms and data balancing techniques like SMOTE, focuses on improving recall, which helps telecom companies better retain customers and reduce revenue loss.

Rudd et al. [3] used deep learning and Bayesian networks to accurately identify potential churners, emphasizing that high recall is crucial for retaining customers who may otherwise leave. By identifying likely churners, companies can proactively address customer dissatisfaction. Recall is prioritized in this framework, as it ensures most potential churners are captured, which is critical for implementing effective retention strategies.

Mahalekshmi et al. [4] examined the effectiveness of deep learning models, specifically Convolutional Neural Networks (CNNs), for predicting churn in the telecom industry. The researchers emphasize the importance of maximizing recall to identify as many potential churners as possible, which is crucial for customer retention and minimizing losses due to churn. They found that CNNs achieved higher recall compared to other neural network models.

Cao, Shulin et al. [5] used a stacked autoencoder to enhance recall, focusing on accurately identifying a broad range of potential churn risks. By maximizing recall, the model ensures that more at-risk customers are detected, allowing for timely intervention strategies. This comprehensive detection of churn risks is designed to reduce customer retention losses, providing the company with actionable insights to better engage and retain valuable customers.

Spanoudes et al. [6] examine the impact of feature engineering on achieving high recall rates in deep learning models for churn prediction, with a particular focus on the telecom sector. The study underscores the importance of capturing even small percentages of potential churners in industries like telecommunications, where customer retention is essential and missed predictions can lead to substantial financial losses. Through refined feature engineering, the research seeks to enhance the model's ability to detect critical behavioural and transactional signals that are strongly indicative of churn. Liu et al. [7] integrated an LSTM-CNN model with an attention mechanism, specifically designed to enhance recall in detecting potential churners within banking data. The combination of Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs) allows the model to capture both sequential and spatial patterns in customer behavior, making it highly effective at identifying subtle signals that may indicate churn. The addition of an attention mechanism further amplifies this capability by directing the model's focus to the most relevant features in the data, enabling it to prioritize patterns strongly associated with churn.

Damien Dablain, Bartosz Krawczyk et al. [8] explores a novel integration of deep learning with SMOTE for handling imbalanced datasets. DeepSMOTE combines synthetic minority sample generation with deep learning architectures, aiming to improve model performance on minority classes in highly skewed datasets. This hybrid approach enhances traditional SMOTE by aligning it with deep feature representations, leading to improved recall and precision in deep learning models by effectively addressing class imbalance challenges in complex data environments.

3. Methodology

In this section, brief descriptions of key parameters and classification techniques that were employed to create a model to forecast loss of customers are explained and the classification techniques that were used to build a model to predict customer churn are described. The Framework of this study is depicted in Fig 1.

A. Dataset

The dataset used in this study is downloaded from Git Hub [9]. The dataset consists of 10,000 instances with 14 features.

B. Data pre-processing

The preprocessing methods are employed on banking dataset to mitigate class imbalance. The dataset is balanced using the SMOTE oversampling technique to address class imbalance, enabling more effective deep learning model training. After applying SMOTE, the dataset is examined for any outliers or anomalies to enhance data quality. The final balanced dataset is then split into training and testing sets to ensure model generalization during deep learning. The balanced dataset is used for further deep learning analysis. The size of the balanced dataset is shown in is shown in TABLE I.

Table I. Size of The Dataset

Method	Class Exited (0)	Class Exited (1)
Before SMOTE	7963	2037
After SMOTE	7963	7963

C. Parameter Tuning

Hyperparameter tuning plays a vital role in optimizing deep learning models, as it significantly affects their performance. The key parameters for different models are outlined in Table II.

D. Classification Techniques

a) Convolutional Neural Network: The Convolutional Neural Network (CNN) operates through a sequence of four key steps namely convolution, non-linearity, pooling, and classification. Initially, the convolution step is responsible for extracting essential features from the input data, preserving the spatial relationships between the features and the class labels [10]. This process helps the model learn important patterns in the data. Next, the non-linearity step introduces activation functions, such as the sigmoid or Rectified Linear Unit (ReLU), to introduce non-linear transformations that enable the model to capture complex relationships between the input features and the hidden layers. Following this, the pooling operation reduces the dimensionality of the feature maps, simplifying the data while retaining the most relevant information. This reduction in complexity helps speed up training and prevents overfitting by focusing on the most significant features. Finally, the classification step takes the high-level features learned from the previous stages and feeds them into the output layer, where the model makes its predictions. Each of these stages plays a crucial role in enabling CNNs to efficiently learn hierarchical features and make accurate predictions based on the input data.

Table II. Key Parameters of Different Models

PARAMETERS	DESCRIPTION
Learning Rate	Controls how much the models' weight are updated during training.
Batch Size	Number of training examples used to compute the gradient for a single update.
Epoch	How many times the entire training data set is passed through the model.
Activation Function	Affects how the model transforms the input.
Optimizer	Determines how the model updates weights based on gradients.
Dropout Rate	Prevents over fitting by randomly dropping a fraction of neurons.

b) Long Short-Term Memory: LSTM (Long Short-Term Memory) is a widely used architecture in recurrent neural networks (RNNs) designed to address the vanishing gradient problem, which impedes traditional RNNs from effectively learning long-term dependencies [11]. LSTM has feature specialized units called memory cells that can retain information for extended periods, making them well-suited for tasks requiring the modeling of long sequences. The memory cell's information flow is regulated by three distinct gates: The Forget Gate, which decides what information from the previous cell state should be retained and what should be discarded; the Input Gate, which controls the flow of new information that enters the cell state; and the Output Gate, which determines the data that is passed on as the output. These gates work in conjunction to selectively manage memory and prevent the loss of important information over time, allowing LSTM networks to overcome the challenges faced by standard RNNs. Due to their ability to handle long-range dependencies and mitigate issues in training, LSTMs are considered one of the most successful and powerful forms of RNNs.

c) Deep Neural Networks: Deep neural networks (DNNs) are an advanced extension of conventional artificial neural networks (ANNs), characterized by their increased depth meaning they contain many more hidden layers compared to traditional networks, which typically have one or two hidden layers [12]. These networks are built to automatically learn hierarchical representations of data by gradually extracting more abstract features from raw input. DNNs are a fundamental part of deep learning, a branch of machine learning that leverages multi-layered neural networks to tackle complex tasks.

d) Artificial Neural Networks: A computational model called an artificial neural network (ANN) is modeled after the biological neural networks seen in the human brain. Layers of interconnected nodes, referred to as neurons, make up this structure. Each neuron in the network receives inputs, processes them using activation functions, and then transmits the results to neurons in the layer below. This is how information moves through the network. Like perceptron, ANNs can be made simple, or they can be made complex, like deep neural networks with plenty of hidden layers. They are skilled at tasks like clustering, regression, classification, and pattern recognition. They are trained via algorithms like backpropagation, which modifies the connection weights in response to variations between the expected and actual outputs. These parameters are adjusted by optimization techniques like adam or gradient Descent in order to minimize the loss function, which gauges prediction accuracy.

e) Recurrent Neural Networks: A recurrent neural network (RNN) is an extension of a conventional feedforward neural network, which is able to handle a variable-length sequence input [12]. Recurrent Neural Networks (RNNs) are a type of artificial neural network designed to process sequential data by retaining a memory of previous inputs through recurrent connections. Unlike traditional feedforward neural networks, RNNs feature loops within their architecture, allowing information to persist across time steps. This enables them to excel in tasks where the sequence and order of inputs are important, such as time-series forecasting, natural language processing, and speech recognition. In an RNN, the output of a neuron is influenced not only by the current input but also by its previous output, allowing the network to "remember" past information. This capability to process sequences of varying lengths allows RNNs to capture time-dependent patterns. However, RNNs can face challenges in handling long-term dependencies due to issues like the vanishing gradient problem, which can hinder the training of the network over long sequences.

E. Evaluation Metrics

The evaluation metrics are used to assess the performance or effectiveness of a model. The evaluation metrics used are as follows.

a) Accuracy: Accuracy measures the proportion of correct predictions out of the total predictions made. The formula to find the accuracy is given in the equation (1).

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (1)$$

Precision: Precision measures the proportion of true positive predictions out of all positive predictions made. The formula to find the precision is given in the equation (2).

$$\text{Precision}= \frac{Tp}{TP+FP} \quad (2)$$

Recall: The recall function quantifies the percentage of accurate positive predictions among all real positive occurrences. The formula to find the recall is given in the equation (3).

$$\text{Recall}= \frac{TP}{TP+FN} \quad (3)$$

F1-Score: F1 Score is the harmonic mean of precision and recall, balancing the two metrics. The formula to find the F1 score is given in the equation (4).

$$\text{F1 – Score} = \frac{2*Precision*Recall}{Precision+Recall} \quad (4)$$

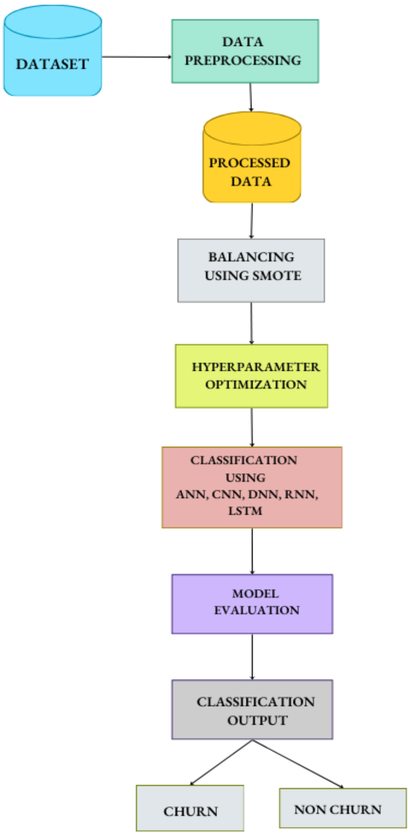

Fig. 1 Framework

4. Result and Discussion

The implementation of the deep learning models is carried out using Python that facilitate data analysis using various models on the banking dataset. The results of the evaluation metrics of all the algorithms are shown in Table III. DNN model was developed with three dense layers, using the ReLU activation function, a 0.4 dropout rate to reduce overfitting, and batch normalization for stable training. The model was trained with the RMSprop optimizer for 25 epochs and a batch size of 128, aiming to balance efficiency and accuracy for improved performance. The results reveal that deep learning models, particularly DNN, outperforms traditional machine learning algorithms in terms of recall when applied to customer churn prediction. Specifically, these models demonstrated higher recall rates, indicating a greater capability to identify true churners compared to the machine learning models. SMOTE was applied to address class imbalance in deep learning models by generating synthetic samples for the minority class. This technique helps balance the dataset and provides more diverse data for better

Table III. Results of The Evolution Metrics

Algorithm	Accuracy	Precision	Recall	F1-Score
ANN	84.6%	81.2%	89.1%	85%
CNN	84.5%	84.2%	83.9%	84.1%
RNN	84.1%	80.5%	88.9%	84.5%
DNN	85.5%	81.4%	91.1%	86%
LSTM	79.3%	80%	76.7%	78.3%

generalization. By preventing bias toward the majority class, SMOTE enhances the model's ability to classify minority class instances, improving performance metrics like recall and precision, which are crucial for handling imbalanced datasets effectively. In the process of training deep learning models on imbalanced datasets, the Synthetic Minority Over-sampling Technique (SMOTE) was applied to address class imbalance. By introducing synthetic samples, SMOTE ensures that the model is not biased toward the majority class, thereby improving its ability to detect and classify minority class instances. This leads to enhanced model performance, particularly in metrics like recall and precision, which are crucial for evaluating the model's ability to handle imbalanced class distributions effectively. Deep learning models remained superior in recall performance, highlighting their adaptability to complex patterns in customer data. Emphasizing recall enables companies to identify a greater number of at-risk customers, allowing them to implement targeted retention strategies proactively, before churn takes place. Although traditional models typically achieved higher precision, they frequently missed a significant portion of actual churners, leading to lower recall rates. The graphical representation of a overall comparison of different machine learning algorithms, deep learning algorithms, comparison of different machine learning in terms of recall and comparison of different machine learning in terms of recall is given in Fig. 2, Fig. 3, Fig. 4 and Fig. 5.

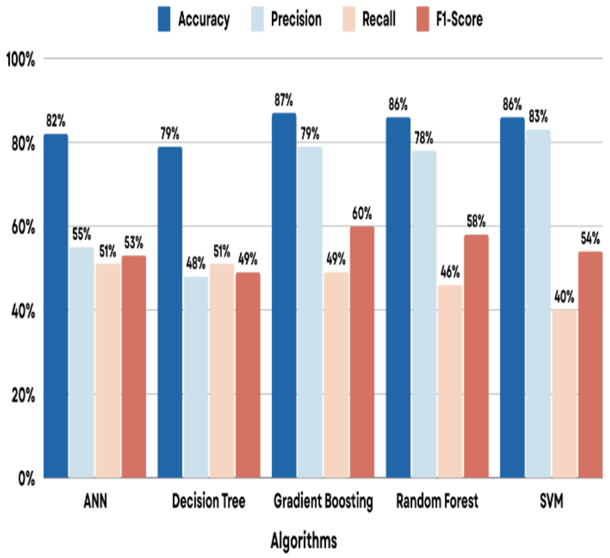

Fig 2. Comparison of Metrics for Different Machine Learning Algorithms

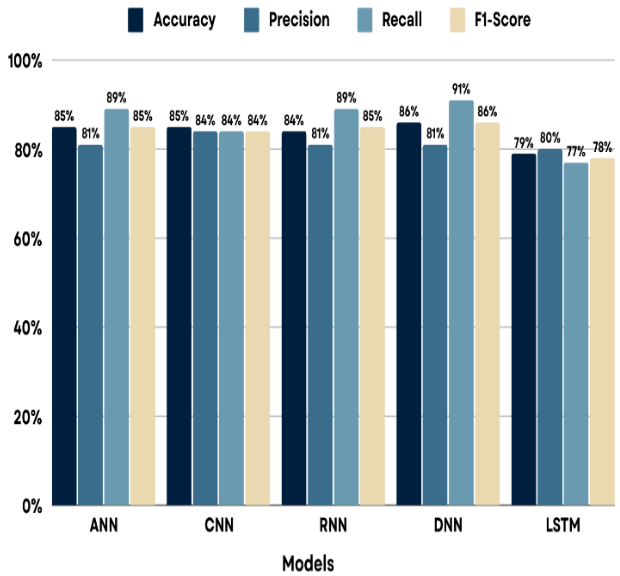

Fig 3. Comparison of different Deep Learning Models

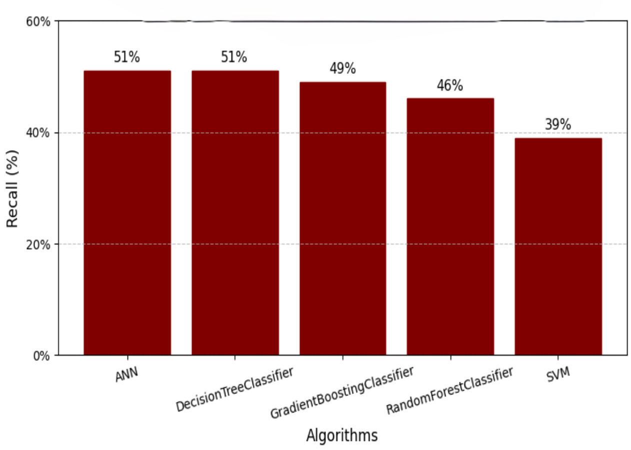

Fig 4. Recall of different Machine Learning Algorithms

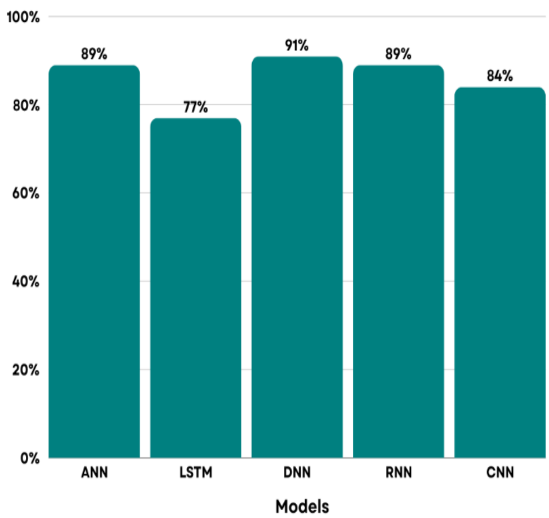

Fig 5. Recall of different Deep Learning Algorithms

5. Conclusion

Effective customer churn prediction is essential for businesses aiming to reduce revenue loss and boost customer retention. Deep learning models, particularly Deep Neural Networks (DNNs), have proven significantly advantageous over traditional machine learning methods in this domain. A key strength of DNNs lies in their high recall, enabling them to identify a greater proportion of customers likely to churn. By prioritizing recall, businesses can capture more at-risk customers, providing opportunities for timely intervention. Deep learning models enhance recall further by employing techniques like SMOTE (Synthetic Minority Over-sampling Technique), which balances datasets by creating synthetic examples of churn cases. Balanced, diverse datasets enable DNNs to generalize more effectively, improving their ability to detect churn-indicating patterns across different customer behaviours. This heightened recall directly supports proactive retention strategies, ensuring that most potential churners are identified, which enables companies to take targeted steps to retain these customers. This recall-centric approach is especially valuable in sectors where each retained customer holds significant financial value, such as telecom, finance, and subscription services. Retaining customers in these fields often costs less than acquiring new ones, making high-recall models vital for maintaining revenue and lowering churn.

Moreover, recall-focused deep learning models facilitate more personalized retention frameworks. These models offer insights not only into which customers might churn but also into specific behaviours and triggers that lead to churn. With this knowledge, businesses can implement tailored interventions that address individual customer needs, building loyalty and enhancing satisfaction. Ultimately, prioritizing recall in churn prediction allows companies to address churn risks early, creating a more dynamic approach to customer retention. By capturing potential churners before they leave, businesses can strengthen customer relationships, supporting long-term growth and stability. This recall-driven strategy goes beyond churn prediction; it enables robust retention frameworks that enhance customer engagement and reinforce resilience in competitive markets.

References

1. Agrawal, Sanket, Amit Gaikwad, Sudhir Dhage. "Customer churn prediction modelling based on behavioural patterns analysis using deep learning." 2018 International conference on smart computing and electronic enterprise (ICSCEE). IEEE, 2018, (pp. 1-6)
2. Saha, S., Saha, C., Haque, M. M., Alam, M. G. R., & Talukder, A. (2024). ChurnNet: Deep Learning Enhanced Customer Churn Prediction in Telecommunication Industry. IEEE Access.
3. Rudd, D. H., Huo, H., & Xu, G. (2021, December). Causal analysis of customer churn using deep learning. In 2021 International Conference on Digital Society and Intelligent Systems (DSInS) (pp. 319-324). IEEE.
4. Mahalekshmi, A., & Chellam, G. H. (2022). Analysis of customer churn prediction using machine learning and deep learning algorithms. International Journal of Health Sciences, 6(S1), 11684–11693.
5. Cao, Shulin, Wei Liu, Yuxing Chen, and Xiaoyan Zhu. "Deep learning-based customer churn analysis." In 2019 11th International Conference on Wireless Communications and Signal Processing (WCSP), pp. 1-6. IEEE, 2019.
6. Spanoudes, Philip, and Thomson Nguyen. "Deep learning in customer churn prediction: unsupervised feature learning on abstract company independent feature vectors." arXiv preprint arXiv:1703.03869 (2017).
7. Liu, Yunjie, Mu Shengdong, Gu Jijian, and Nadia Nedjah. "Intelligent prediction of customer churn with a fused attentional deep learning model." Mathematics 10, no. 24 (2022): 4733.
8. Dablain, Damien, Bartosz Krawczyk, and Nitesh V. Chawla. "DeepSMOTE: Fusing deep learning and SMOTE for imbalanced data." IEEE Transactions on Neural Networks and Learning Systems 34.9 (2022): 6390-6404.
9. https://github.com/DataThinkers/Datasets/blob/main/DS/Churn_ Modelling.csv
10. Mishra, Abinash, and U. Srinivasulu Reddy. "A novel approach for churn prediction using deep learning." 2017 IEEE international conference on computational intelligence and computing research (ICCIC). IEEE, 2017.
11. Sarker, Iqbal H. "Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions." SN computer science 2, no. 6 (2021): 420.
12. Chan KY, Abu-Salih B, Qaddoura R, Ala'M AZ, Palade V, Pham DS, Del Ser J, Muhammad K. Deep neural networks in the cloud: Review, applications, challenges and research directions. Neurocomputing. 2023 Aug 7; 545:126327.