International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 07 | July - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

A Resilient Back Propagation Based Deep Learning Model for Predicting Customer Churn Rate

Khushi Soni[1], Prof. Manisha Kadam[2]

ABSTRACT

Data Science and Machine Learning are being used extensively for business analytics. One of the major applications happens to be estimating churn and attrition rates. In today's competitive market landscape, retaining customers is as crucial as acquiring new ones. Churn rate, which measures the proportion of customers who discontinue their relationship with a business over a specific period, is a critical metric for companies across industries. Forecasting churn enables businesses to proactively address customer dissatisfaction and refine their strategies to retain valuable clients. By understanding the likelihood of churn, companies can make informed decisions to sustain growth and profitability. The proposed approach employs Resilient Backpropagation to train a deep neural network. The results clearly indicate that the proposed approach outperforms existing baseline approaches in terms of forecasting accuracy.

Keywords: Data Analytics, Machine Learning, Churn Rate, Particle Resilient Backpropagation, Mean Absolute Percentage Error, Regression.

I. Introduction

Machine Learning is being used extensively to evaluate market conditions and develop optimize approaches. With the rise of machine learning and big data analytics, companies can now forecast churn with greater accuracy [1]. These technologies analyze vast datasets to detect subtle patterns and indicators of churn that may be overlooked by traditional methods. For instance, a sudden drop in product usage or a spike in support ticket submissions can signal dissatisfaction [2]. By integrating these insights into their customer relationship management systems, businesses can automate responses and take timely actions to mitigate churn risk [3]. One important application happens to be estimating attrition or churn rates. Churn not only signifies lost revenue but also reflects potential issues in product quality, service satisfaction, or customer engagement [4]. High churn rates can cripple a company's growth, especially in subscription-based industries like telecom, SaaS, and streaming services, where recurring revenue is vital [5]. Moreover, the cost of acquiring new customers often surpasses that of retaining existing ones, making churn a significant financial concern. Therefore, accurately forecasting churn helps in identifying weak points in the customer journey that need immediate attention [6].

Forecasting churn involves using historical data, customer behavior patterns, and predictive analytics to estimate the likelihood of a customer leaving [7]. This insight allows businesses to tailor personalized interventions—such as targeted promotions, loyalty programs, or enhanced support services—before customers decide to leave. In industries such as banking and insurance, where long-term relationships are valuable, churn forecasting serves as an essential tool for maintaining customer satisfaction and trust [8].

II. Existing Machine Learning Models for Churn Prediction

Traditional machine learning models like logistic regression, decision trees, random forests, and support vector machines have been widely used in churn prediction tasks [8]. These models rely on structured data such as customer demographics, transaction history, service usage, and feedback. These models require careful feature engineering and selection to achieve optimal results, but they are generally efficient and easy to implement [10].

Deep learning models, especially artificial neural networks (ANNs), have gained traction due to their ability to learn hierarchical representations from raw data [11]. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are particularly effective in handling time-series data, making them ideal for analyzing customer behavior over time [12]. Convolutional Neural Networks (CNNs), though traditionally used in image processing, have also been adapted for churn prediction by treating sequential data as matrices [13]. Deep learning models can automatically extract features and capture nonlinear relationships, although they typically require large datasets and more computational resources [14].

While both ML and DL models can effectively forecast churn, their applicability depends on the data and business context [15]. Machine learning models are preferable when interpretability and lower computational costs are critical. In contrast, deep learning models excel in scenarios involving complex and high-dimensional data, such as user activity logs and behavioral sequences [16].

III. Proposed Methodology

The proposed approach employs the Resilient Backpropagation (Rprop). The discussion starts with the fundamental mathematical model of the artificial neural network.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 07 | July - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

The ANN Model:

The ANN model is one of the most powerful regression models which has been used multiple times for traffic speed forecasting [17]. The mathematical model of the ANN is depicted in figure 1.

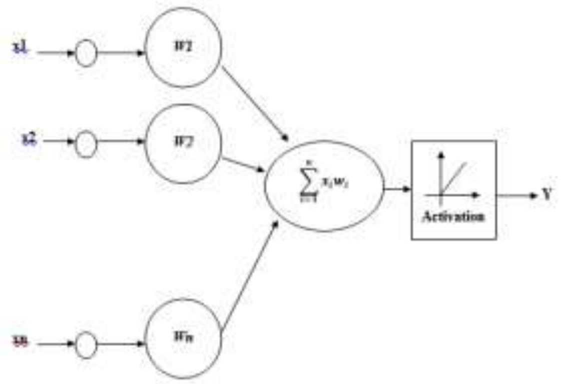

Fig.1 Mathematical Model of Neural Network

The output of the neural network is given by:

$$y = f\left(\sum_{i=1}^{n} \mathbf{XiWi} + \Theta\right) \quad \mathbf{(1)}$$

Where,

Xi represents the signals arriving through various paths, Wi represents the weight corresponding to the various paths and

ϴ is the bias.

In this approach, the Resilient Backpropagation (Rprop) is employed. The training algorithms play a critical role in determining the speed and accuracy of learning. One of the most well-known training algorithms is backpropagation, which uses gradient descent to adjust the network's weights in the direction of error minimization. However, traditional gradient descent methods often suffer from slow convergence and sensitivity to the magnitude of gradients, especially when gradients are very small or vary greatly in scale. To address these challenges, Resilient Backpropagation (Rprop) was introduced as an improved alternative that enhances the stability and speed of the learning process [18].

Given $\theta_1, \theta_2, \ldots\ldots\ldots\ldots\ldots\ldots\ldots.\theta_{nu}$, to learn the value of features x_i

The following cost function is to be minimized:

$$\min\left[\sum_{j:r(i,j)=1}^{n} {\left[\theta_j^T x^i - y^{i,j}\right]^2}\Big/_{n}\right] + \frac{\lambda}{2}\sum_{i=1}^{n}[x_i]^2 \quad \mathbf{(2)}$$

Here,

i is the number of features

j is the number of features which are labelled

$\sum_{j:r(i,j)=1}^{n} {\left[\theta_j^T x^i - y^{i,j}\right]^2}\Big/_{n}$ is the mean square error (mse) of actual class and predicted class.

$\frac{\lambda}{2}\sum_{i=1}^{n}[x_i]^2$ is the regularizing term to limit the number of features to become too large, to avoid overfitting [19].

The fundamental idea behind Rprop is to decouple the weight update process from the magnitude of the gradient. Unlike standard gradient descent, which uses the actual value of the gradient to scale the weight updates, Rprop only uses the sign of the gradient to determine the direction of the weight adjustment. This means that the algorithm is not influenced by the steepness of the error surface, making it more robust to issues like vanishing or exploding gradients. Each individual weight in the network has its own adaptive learning rate, or update value, which is adjusted during training based on the consistency of the gradient's sign. The training rule is given by [20]:

$$\boldsymbol{w_{t+1} = w_t - \Delta w_{i,j}^t \chi_t g_t} \quad \mathbf{(3)}$$

The learning rate is updated as**:**

$$\boldsymbol{\chi_{t+1} = \frac{\chi_t}{E(g_t^2)+\epsilon}} \quad \mathbf{(4)}$$

During the training process, historical data is used to feed the network, and the algorithm calculates the error between the predicted and actual values. This error is then propagated backward through the network, adjusting the weights and biases of the connections to minimize the prediction error. This iterative process continues until the network converges to a state where the error is minimized.

The training is stopped based on the mean square error or mse given by:

$$\boldsymbol{mse = \frac{\sum_{i=1}^{n} e_i^2}{n}} \quad \mathbf{(5)}$$

Proposed Algorithm:

The algorithm of the proposed approach is presented subsequently:

Start

{

Step.1 *Extract dataset and divide data into the ratio of 70:30 for training : testing.*

Step.2 *Assign input and target variables.*

Step.3 *Initialize weight matrix randomly.*

Step.4 *To train the network, employ the following training rule:*

$$\boldsymbol{w_{t+1} = w_t - \Delta w_{i,j}^t \chi_t g_t}$$

Step.*5 If (cost function stabilizes)*

Truncate training

Else if (max. iterations are over)

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 07 | July - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Truncate Training
Else
Feedback errors as inputs to subsequent iteration.
Step.6 *if (error is stable through validation checks i.e. consecutive iterations)*
Stop training
else if (maximum iterations are over even without error stabilization)
Stop Training
else
{
Feed next training vector
Back propagation of error
}
Step.7: *Simulate model to forecast samples.*
Step.8 *Compute performance metrics.*
}
Stop

The final computation of the performance metric is the mean absolute percentage error given by:

$$MAPE = \frac{100}{M}\sum_{i=1}^{N}\frac{E-E_i|}{i} \quad (6)$$

The accuracy of prediction is computed as:

$$Ac = 100 - \frac{100}{M}\sum_{i=1}^{N}\frac{E-E_i|}{i}\,\% \quad (7)$$

Here,
n is the number of errors
i is the iteration number
E is the actual value
E_i is the predicted value

IV. RESULTS AND DISCUSSIONS

The proposed model is implemented on MATLAB due to the availability of in built mathematical functions for analysis of traffic volume. The data parameters used are: Age, gender, tenure, usage frequency, support calls, payment delay, subscription type, contract length, total spend, and last interaction.
The target variable is churn:
(1: Yes, 0: No)
While other parameters may also be important, the limited set of parameters are chosen to design a streamlined model The results are presented next.

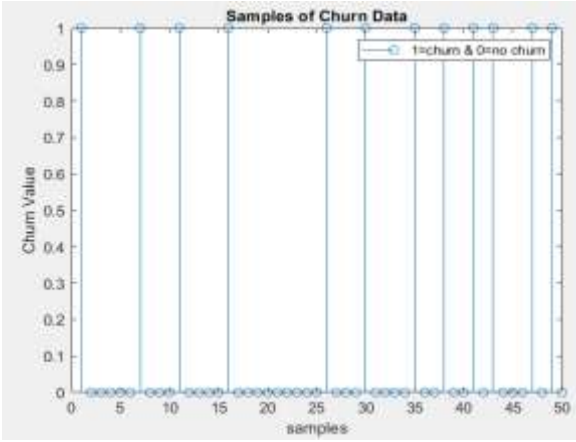

Fig.2. Raw Data

Figure 2 depicts the raw data with:
1= Churn
0= No Churn

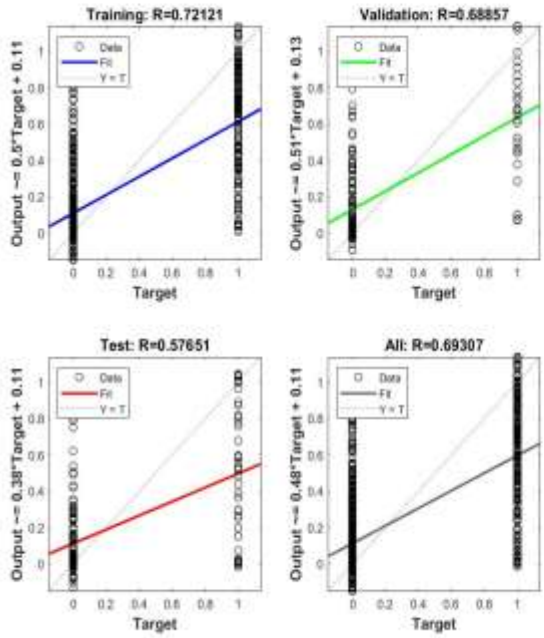

Fig.3 Regression

The figure above depicts the regression obtained in the proposed approach which is a sort of similarity among two random variables. The maximum allowable regression is unity depicting complete similarity.

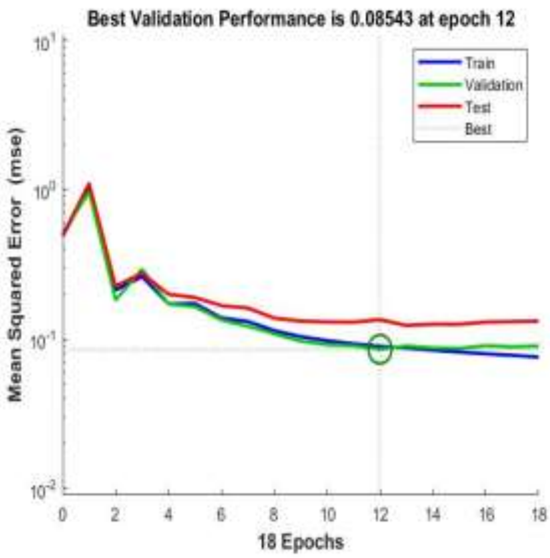

Fig.4 Performance Function

The performance function that decides the culmination of training is the mean squared error or mse.

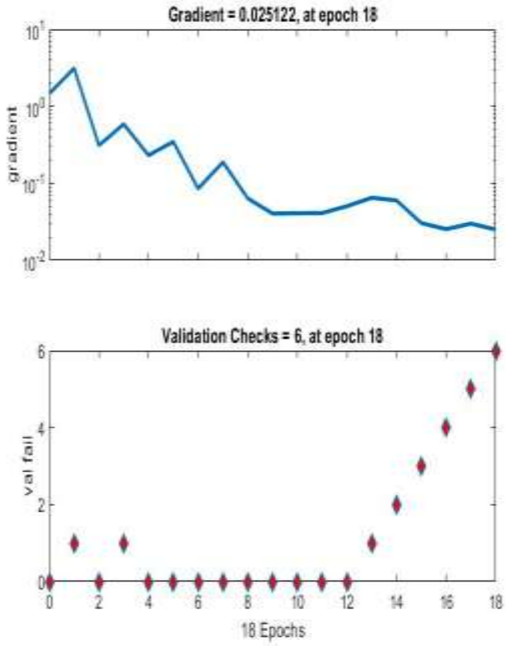

Fig.5 Training States

The training state parameters such as gradient, combination co-efficient and validations checks are depicted in the figure above.

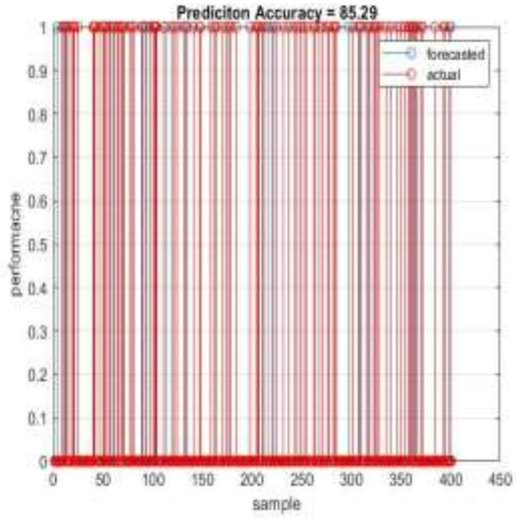

Fig.6 Accuracy of Proposed Work

The above figure shows that the accuracy of the proposed work is 85.29%

The summary of results with comparison with previous work is presented in table 1.

Table. 1 Summary of Results

S.No	PARAMETER	VALUE
1.	Model	Neural Networks with Resilient Backpropagation (Rprop
2.	Target	1=Churn 0=No Churn
3.	Regression	0.9389
4.	Accuracy (Proposed Work)	85.29%
5.	Accuracy (Previous Work)	81%
6.	Approach (Proposed Work)	Resilient Backpropagation (Rprop)
7.	Approach (Previous Work) [21]	Gradient Boosting (BGM)

The summary of results is presented in table 1. The performance of the proposed approach (Accuracy of 85.29%) is found better compared to previously existing technique [1] which attains a Accuracy of 81%% using the Gradient Boosting (GBM) model.
The improvement in the results can be attributed to the optimization in the training process

V. CONCLUSION

Forecasting churn is not just a reactive tool but a strategic asset in today's customer-centric economy. It empowers organizations to stay ahead of customer

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 07 | July - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

behavior, reduce attrition, and foster lasting relationships. As competition intensifies and customer expectations evolve, the ability to predict and prevent churn will increasingly define a company's success. Industries that invest in churn forecasting capabilities position themselves better for resilience and growth in a dynamic marketplace. The proposed work employs the Resilient Backpropagation (Rprop) in neural networks and obtains improved forecasting accuracy for churn compared to existing approaches in the domain.

References:

[1] S. Wu, W. -C. Yau, T. -S. Ong and S. -C. Chong, "Integrated Churn Prediction and Customer Segmentation Framework for Telco Business," in IEEE Access, vol. 9, pp. 62118-62136, 2021

[2] V. Chang, K. Hall, Q. A. Xu, F. O. Amao, M. A. Ganatra, and V. Benson, "Prediction of Customer Churn Behavior in the Telecommunication Industry Using Machine Learning Models," Algorithms, vol. 17, no. 6, p. 231, 2024.

[3] A. Manzoor, M. Atif Qureshi, E. Kidney and L. Longo, "A Review on Machine Learning Methods for Customer Churn Prediction and Recommendations for Business Practitioners," in IEEE Access, vol. 12, pp. 70434-70463, 2024

[4] S. Wu, "Customer Churn Prediction in Telecom Based on Machine Learning," in Proc. 2nd Int. Conf. on Computer, Machine Learning and Artificial Intelligence (CMLAI 2024), Highlights in Science, Engineering and Technology, vol. 94, pp. 113–118, 2024

[5] R. Tyagi and K. Sindhu, "Customer Churn Analysis Using Machine Learning," in Proc. Int. Joint Conf. on Advances in Computational Intelligence, Algorithms for Intelligent Systems, Springer, Singapore, pp. 423–432, 2022.

[6] Y. Zhang, S. He, S. Li, and J. Chen, "Intra-Operator Customer Churn in Telecommunications: A Systematic Perspective," IEEE Transactions on Vehicular Technology, vol. 69, no. 1, pp. 948–957, Jan. 2020

[7] P Sharmila, J Baskaran, C Nayanatara, "A hybrid technique of machine learning and data analytics for optimized distribution of renewable energy resources targeting smart energy management", Procedia in Computer Science, Elsevier 2019, vol.165, pp. 278-284.

[8] Y. Deng, D. Li, L. Yang, J. Tang, and J. Zhao, "Analysis and Prediction of Bank User Churn Based on Ensemble Learning Algorithm," in Proc. IEEE Int. Conf. on Power Electronics, Computer Applications (ICPECA), 2021, pp. 288–291.

[9] R. K. V. L. N. and P. Deeplakshmi, "Dynamic Churn Prediction Using Machine Learning Algorithms—Predict Your Customer Through Customer Behaviour," in Proc. IEEE Int. Conf. on Computer Communication and Informatics (ICCCI), 2021, pp. 1–6.

[10] R. Yahaya, O. A. Abisoye, and S. A. Bashir, "An Enhanced Bank Customers Churn Prediction Model Using a Hybrid Genetic Algorithm and K-Means Filter and Artificial Neural Network," in Proc. IEEE 2nd Int. Conf. on Cyberspace (CYBER NIGERIA), 2021, pp. 52–58.

[11] Y. E. Özköse, A. Haznedaroğlu, and L. M. Arslan, "Customer Churn Analysis with Deep Learning Methods on Unstructured Data," in Proc. Innovations in Intelligent Systems and Applications Conf. (ASYU), 2021, pp. 1–5.

[12] S. De, P. P., and J. Paulose, "Effective ML Techniques to Predict Customer Churn," in Proc. 3rd Int. Conf. on Inventive Research in Computing Applications (ICIRCA), 2021, pp. 895–902.

[13] V. Geetha, A. Punitha, A. Nandhini, T. Nandhini, S. Shakila, and R. Sushmitha, "Customer Churn Prediction in Telecommunication Industry Using Random Forest Classifier," in Proc. Int. Conf. on System, Computation, Automation and Networking (ICSCAN), 2020, pp. 1–5.

[14] P. Bhuse, A. Gandhi, P. Meswani, R. Muni, and N. Katre, "Machine Learning Based Telecom-Customer Churn Prediction," in Proc. 3rd Int. Conf. on Intelligent Sustainable Systems (ICISS), 2020, pp. 1297–1301.

[15] I. Jahan and T. Farah Sanam, "An Improved Machine Learning Based Customer Churn Prediction for Insight and Recommendation in E-commerce," 2022 25th International Conference on Computer and Information Technology (ICCIT), Cox's Bazar, Bangladesh, 2022, pp. 1-6.

[16] H. Zhao, X. Zuo and Y. Xie, "Customer Churn Prediction by Classification Models in Machine Learning," 2022 9th International Conference on Electrical and Electronics Engineering (ICEEE), Alanya, Turkey, 2022, pp. 399-407.

[17] B. Ghavami, M. Sadati, Z. Fang and L. Shannon, "FitAct: Error Resilient Deep Neural Networks via Fine-Grained Post-Trainable Activation Functions," 2022 Design, Automation & Test in Europe

Conference & Exhibition (DATE), Antwerp, Belgium, 2022, pp. 1239-1244

[18] C. D. Schuman et al., "Resilience and Robustness of Spiking Neural Networks for Neuromorphic Systems," 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 2020, pp. 1-10.

[19] W Xiaosen, K Tong, K He, “Rethinking the backward propagation for adversarial transferability”, Advances in Neural Information proceedings.neurips.cc

[20] M Almiani, A Abughazleh, Y Jararweh, “Resilient back propagation neural network security model for containerized cloud computing”, Simulation Modelling Practice and Theory, Elsevier 2022, vol.118, 102544.

[21] SS Poudel, S Pokharel, M Timilsina, “Explaining customer churn prediction in telecom industry using tabular machine learning models”, Machine Learning with Applications, Elsevier 2024, vol.17, 100567.