An Optimized Regression Learning Approach for Accurate Estimation of Customer Attrition

Yati Siyota[1] Prof. Arjun Singh Parihar[2]
Research Scholar[1], Head of the Department[2]
Department of CSE, SDBCE, Indore[1,2]

ABSTRACT
In today's competitive market landscape, retaining customers is as crucial as acquiring new ones. Customer attrition or customer churn is a negative attribute for maintaining business profitability over a long period of time. 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 combines swarm intelligence and neural networks to forecast churn rates. In terms of the prediction accuracy, the proposed model clearly beats existing approaches.

Keywords: Customer Attrition or Churn, Machine Learning, Swarm Intelligence, Artificial Neural Network (ANN), 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 methodology presents an amalgamation of the following two approaches:

1. Particle Swarm Optimization (PSO)
2. Artificial Neural Networks (ANN)

Each of the approaches are explained next.

The PSO:

The PSO algorithm is an evolutionary computing technique, modeled after the social behavior of a flock of birds. In the context of PSO, a swarm refers to a number of potential solutions to the optimization problem, where each potential solution is referred to as a particle [17]. The aim of the PSO is to find the particle position that results in the best evaluation of a given fitness function. In the initialization process of PSO, each particle is given initial parameters randomly and is 'flown' through the multi-dimensional search space [18]. During each generation, each particle uses the information about its previous best individual position and global best position to maximize the probability of moving towards a better solution space that will result in a better fitness. When a fitness better than the individual best fitness is found, it will be used to replace the individual best fitness and update its candidate solution according to the following equations [19]:

$$\mathbf{v_{id}(t) = w \times v_{id}(t-1) + c_1 Ø_1(p_{id}\text{-}x_{id}(t\text{-}1)) + c_2 Ø_2(p_{gd}\text{-}x_{id}(t\text{-}1))} \quad (1)$$

$$\mathbf{x_{id}(t) = x_{id}(t-1) + v_{id}(t)} \quad (2)$$

Table. 1 List of variables used in PSO equations.

v	The particle velocity
x	The particle position
t	Time
c_1,c_2	Learning factors
Φ_1 ,Φ_2	Random numbers between 0 and 1
p_{id}	Particle's best position
p_{gd}	Global best position
w	Inertia weight

The PSO is used to adaptively update the weights of the neural network based on the minimization of the performance function.

The ANN Model:

The ANN model is one of the most powerful regression models which has been used multiple times for traffic speed forecasting [20]. The mathematical model of the ANN is depicted in figure 1.

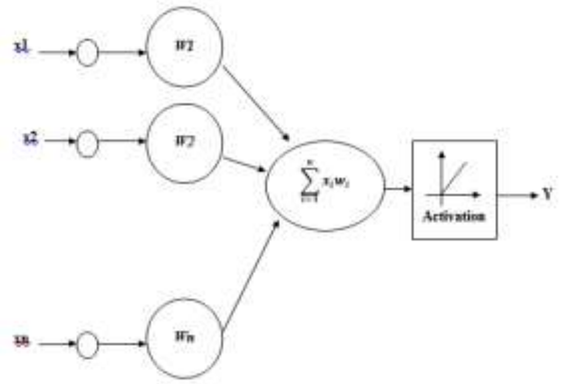

Fig.1 Mathematical Model of Neural Network

The output of the neural network is given by:

$$y = f(\sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{XiWi} \ + \ \Theta) \quad \mathbf{(4)}$$

Where,

Xi represents the signals arriving through various paths,
Wi represents the weight corresponding to the various paths and
Ɵ is the bias.

In this approach, the back propagation based neural network model has been used. A backpropagation neural network for customer attrition forecasting typically consists of an input layer, one or more hidden layers, and an output layer [21]. The number of nodes in the input layer corresponds to the features used for prediction. The hidden layers contain nodes that learn and capture the intricate patterns within the data, while the output layer provides the predicted value [22]. The training of a backpropagation neural network involves the iterative application of the backpropagation algorithm [23].

The training rule for the back propagation used in this approach is given by:

Considering the loss/cost function as the mean squared error, the weight update algorithm for PSO-ANN hybrid is [24]:

$$\boldsymbol{w_{k+1} = \vartheta_k w_k - \left[J_k J_k^T + \mu I\right]^{-1} J_k^T e_k * X_k} \quad \mathbf{(5)}$$

Here,
k is the iteration number
w_{k+1} is weight of next iteration,
w_k is weight of present iteration
J_k is the Jacobian Matrix and is given by the terms $J_k = \frac{\partial^2 e}{\partial w^2}$ i.e. the second order rate of change of errors with respect to weights
J_k^T is Transpose of Jacobian Matrix

e_k is error of Present Iteration
μ is step size i.e. amount by which weight changes in each iteration
I is an identity matrix, with all diagonal elements equal to 1 and other elements 0.

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 [25]. This iterative process continues until the network converges to a state where the error is minimized [26].
The training is stopped based on the mean square error or mse given by:

$$mse = \frac{\sum_{i=1}^{n} e_i^2}{n} \tag{6}$$

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:*

$$w_{k+1} = \vartheta_k w_k - [J_k J_k^T + \mu I]^{-1} J_k^T e_k * X_k$$

Step.5 *If (cost function stabilizes)*
Truncate training
Else if (max. iterations are over)
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:

$$MAPE = \frac{100}{M} \sum_{i=1}^{N} \frac{E - E_i|}{i} \tag{7}$$

The accuracy of prediction is computed as:

$$Ac = 100 - \frac{100}{M} \sum_{i=1}^{N} \frac{E - E_i|}{i} \% \tag{8}$$

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.

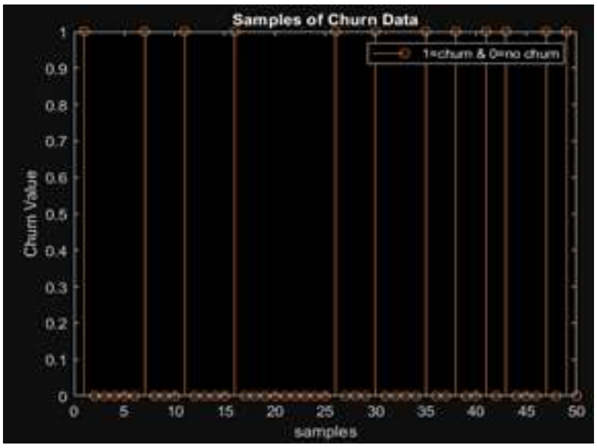

Fig.2. Raw Data

Figure 2 depicts the raw data with both churn and non-churn cases.

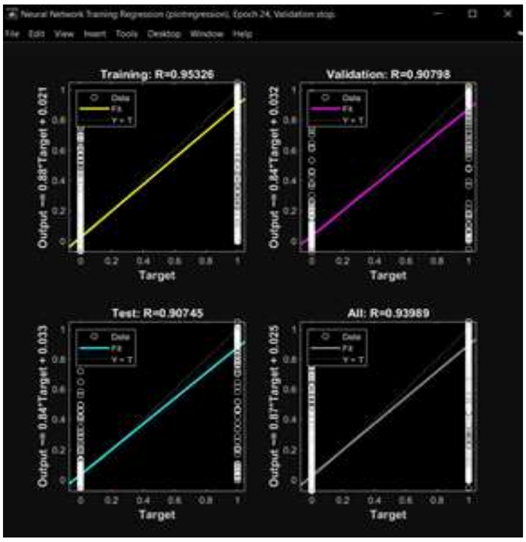

Fig.3 Regression

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.

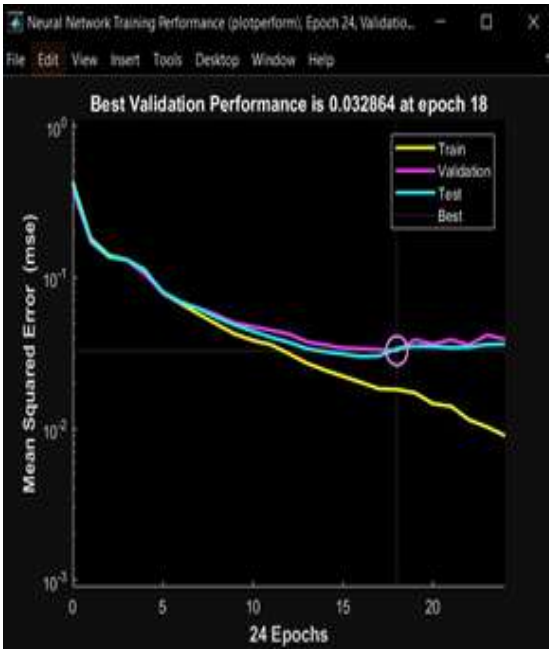

Fig.4 Performance Function

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

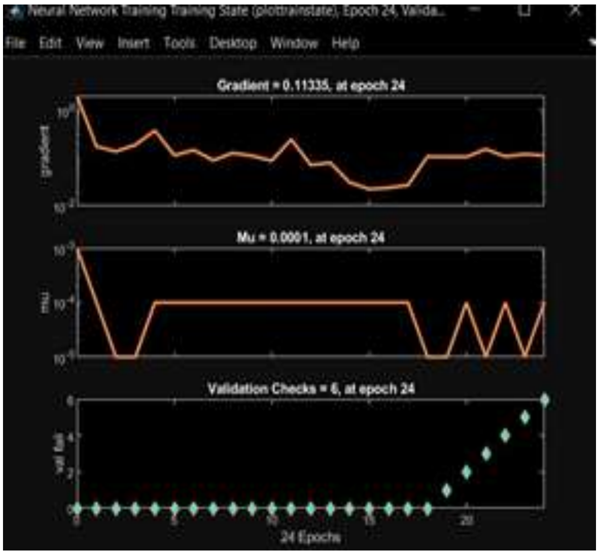

Fig.5 Training States

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

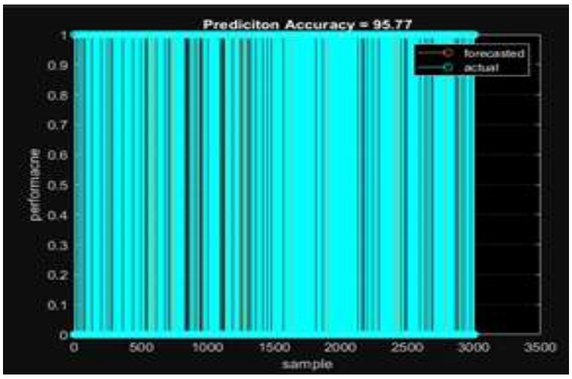

Fig.6 Accuracy of Proposed Work

The above figure shows that the accuracy of the proposed work is 95.77%

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

Table. 1 Parameter Table

S.NO	PARAMETER	VALUE
1.	Software Platform	MATLAB
2.	ML Model	PSO-ANN
3.	Regression	0.9389
4.	Iterations	24
5.	Accuracy (Previous Work [18])	81%
6.	Accuracy (Proposed Work)	95.779%
7.	Cost Function at Convergence	0.032
8.	Learning Rate at Convergence	0.0001

The parameter table clearly shows that the proposed approach outperforms the existing work in the domain in terms of accuracy of prediction.

V. CONCLUSION

It can be conclude that customer attrition or churn, which is the loss of customers from a service provider—is a major challenge across industries such as telecommunications, banking, e-commerce, and subscription-based platforms. Companies invest heavily in predicting churn so they can intervene early with targeted retention strategies. However, churn behavior is complex and influenced by multiple nonlinear factors such as customer activity, service usage patterns, satisfaction levels, pricing, and demographic attributes. Traditional statistical models often struggle to capture these interactions, which creates a need for advanced hybrid approaches that combine swarm intelligence with neural networks to achieve higher predictive accuracy and robustness. The proposed approach combines PSO and Neural Networks to predict customer attrition or churn and clearly outperforms previous work in the domain.

References:

[1] Y Ahn, "Predicting customer attrition using binge trading patterns: Implications for the financial services industry", Journal of the Operational Research Society, Taylor & Francis, 2023, vol.74, no.8, pp. 1878-1891.
[2] P. Sobreiro, D. D. S. Martinho, J. G. Alonso and J. Berrocal, "A SLR on Customer Dropout Prediction," in IEEE Access, 2022, vol. 10, pp. 14529-14547
[3] 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.
[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] S. Wu, W.-C. Yau, T.-S. Ong, and S.-C. Chong, "Integrated Churn Prediction and Customer Segmentation Framework for Telco Business," IEEE Access, vol. 9, pp. 62118–62136, 2021
[8] 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.
[9] 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.
[10] 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.
[11] 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.
[12] 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.
[13] 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.
[14] 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.
[15] 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.
[16] 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.
[17] 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.
[18] 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.
[19] B. A. S. Emambocus, M. B. Jasser and A. Amphawan, "A Survey on the Optimization of Artificial Neural Networks Using Swarm Intelligence Algorithms," in IEEE Access, vol. 11, pp. 1280-1294, 2023.
[20] A. Ashraf, W. H. Bangyal, H. T. Rauf, S. Pervaiz and J. Ahmad, "Training of Artificial Neural Network Using New Initialization Approach of Particle Swarm Optimization for Data Classification," 2020 International Conference on Emerging Trends in Smart Technologies (ICETST), 2020, pp. 1-6.
[21] M. G. M. Abdolrasol, R. Mohamed, M. A. Hannan, A. Q. Al-Shetwi, M. Mansor and F. Blaabjerg, "Artificial Neural Network Based Particle Swarm Optimization for Microgrid Optimal Energy Scheduling," in IEEE Transactions on Power Electronics, vol. 36, no. 11, pp. 12151-12157, Nov. 2021.
[22] T. Lawrence, L. Zhang, C. P. Lim and E. -J. Phillips, "Particle Swarm Optimization for Automatically Evolving Convolutional Neural Networks for Image Classification," in IEEE Access, vol. 9, pp. 14369-14386, 2021.
[23] U. R. Muduli, K. Al Jaafari and R. K. Behera, "Optimized Neural Network Based Predictive Maintenance for Five-Phase Induction Motor Failure," 2021 IEEE 12th Energy Conversion Congress & Exposition - Asia (ECCE-Asia), Singapore, Singapore, 2021, pp. 1537-1541.
[24] . Khan et al., "Design of Neural Network With Levenberg-Marquardt and Bayesian Regularization Backpropagation for Solving Pantograph Delay Differential Equations," in IEEE Access, vol. 8, pp. 137918-137933, 2020.
[25] H. R. Madvar, M. Dehghani, R. Memarzadeh, E. Salwana, A. Mosavi and S. S., "Derivation of Optimized Equations for Estimation of Dispersion Coefficient in Natural Streams Using Hybridized ANN With PSO and CSO Algorithms," in IEEE Access, 2020, vol. 8, pp. 156582-156599.
[26] A. Ansari, I. S. Ahmad, A. A. Bakar and M. R. Yaakub, "A Hybrid Metaheuristic Method in Training Artificial Neural Network for Bankruptcy Prediction," in IEEE Access, vol. 8, pp. 176640-176650,.