International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

IJIEMR Transactions, online available on 11th Apr 2024. Link
:http://www.ijiemr.org/downloads.php?vol=Volume-13&issue=Issue4

10.48047/IJIEMR/V13/ISSUE 04/10

Title AI Resonance: Forecasting Customer Churn for Future Stability

Volume 13, ISSUE 4, Pages: 80-87

Paper Authors **Dr. UdayaSri Kompalli, 2Gandham Mani Saketh 3Abdul Faheem**

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

AI Resonance: Forecasting Customer Churn for Future Stability

[1]Dr. UdayaSri Kompalli [0009-0008-2110-8631],
[2]Gandham Mani Saketh [0009-0007-6602-3933] , [3]Abdul Faheem

[1]Assistant Professor, Dept. of Computer Science, PB Siddhartha College of Arts & Science, Vijayawada, A.P., India,
[2]Student, Department of Computer Science, P.B. Siddhartha College of Arts & Science, Vijayawada, AP, India,
[3]Student, Department of Computer Science, P.B. Siddhartha College of Arts & Science, Vijayawada, AP, India
kudayasri@pbsiddhartha.ac.in, gandhammani2421@yahoo.com, abdulfaheemaf11@gmail.co

Abstract
This study delves into the dynamic realm of customer churn prediction using Artificial Neural Networks (ANN) and deep learning methodologies. In an era where customer retention is paramount, understanding and forecasting churn patterns is a critical aspect of business strategy. Leveraging a comprehensive dataset, we employ advanced neural networks to decode intricate patterns within customer behaviour. The study explores the intricate interplay of features contributing to churn, harnessing the power of AI to enhance predictive analytics. Our findings shed light on the underlying factors influencing customer loyalty, providing businesses with actionable insights for crafting effective retention strategies. This study not only contributes to the growing field of predictive analytics but also presents a practical guide for businesses seeking to navigate the seas of customer churn with confidence and precision.

Keywords: Customer Churn, Deep Learning, ANN(Artificial Neural Network).

Introduction
Customer analytics has become a pervasive buzzword in the industry. It is employed to make critical business decisions to model customer behaviours using predictive analytics. Lifetime value modelling, market / customer segmentation and churn analysis are the most popular topics of customer analytics in the industry and academia. Information obtained from these application domains is used for different purposes such as customer relationship management and direct marketing. Yet, companies have still much to obtain from different business applications of customer analytics [1].

The rapidly evolving landscape of customer analytics has witnessed a growing interest in the application of advanced artificial intelligence techniques, particularly Artificial

Neural Networks (ANN), to address the critical challenge of predicting customer churn. Customer churn, defined as the phenomenon where customers discontinue their association with a service or product, holds immense significance across various industries. The intricate nature of churn prediction, characterized by issues such as imbalanced churn rates, data quality concerns, and the need for precise classifications, has prompted the exploration of innovative methodologies. Our study embarks on a journey into the realm of churn prediction, leveraging the capabilities of ANN in deciphering complex patterns inherent in customer behaviour data. As traditional machine learning methods grapple with challenges related to accuracy and precision, the paper delves into the potential of deep learning techniques, specifically ANN, to overcome these limitations. The utilization of ANN promises a more efficient and accurate churn prediction process, with the ability to autonomously identify crucial features without exhaustive manual intervention.

We are now Exploring into the unique advantages offered by ANN in processing vast datasets, particularly in sectors characterized by diverse and extensive customer attributes. By adopting ANN, we aim to enhance the accuracy and effectiveness of churn prediction models, presenting a novel approach that contributes to the ongoing discourse in the field of customer analytics. This paper unfolds the methodology, challenges, and implications of employing ANN in churn prediction, shedding light on the transformative potential of advanced artificial intelligence in deciphering the intricate dynamics of customer retention.

Review of Literature

Exploring the ascent of customer analytics and CMOs' doubts about its impact on firm performance, Frank Germann's study employs a multi-method approach, encompassing global surveys and interviews. It validates a positive association between customer analytics adoption and firm success while emphasizing the critical role of top management team (TMT) advocacy in its implementation and effectiveness.

The study, authored by Baby, B. et al.'s investigates the impact of customer churn on banking industries. It develops an ANN model to predict churn, achieving an 86% accuracy rate through optimization techniques, surpassing logistic regression. The research provides insights into enhancing client retention strategies, aiding banks in proactive customer retention efforts.

Omer Faruk Seymen et al.'s research introduces ANN and CNN models to forecast customer churn in the retail industry, crucial for profitability and customer relationships. The study compares these models with common machine learning techniques, highlighting the CNN model's exceptional performance with a 97.62% accuracy rate. This underscores its superiority in classification and prediction, providing valuable insights for optimizing churn prediction strategies.

The study by Abinash Mishra and U. Srinivasulu Reddy addresses the critical issue of customer retention in the telecommunications sector. With a focus on churn prediction, the research employs Convolutional Neural Networks (CNNs) to achieve an impressive 86.85% accuracy rate. These findings underscore the potential of CNNs to enhance customer retention strategies in CRM.

Methodology

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

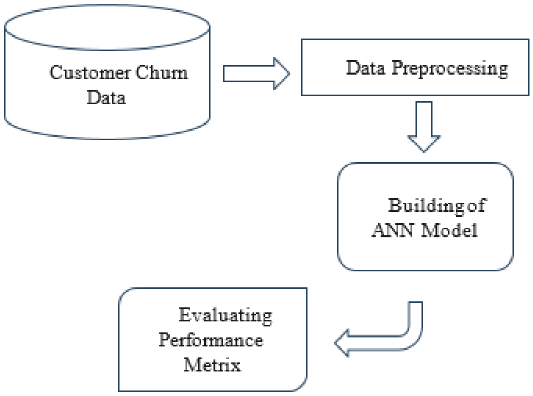


Fig.1. Framework of Our Study

Table 1. Summary of the Attributes

RowNumber:	Row Numbers from one to ten thousand
CustomerId:	It is a unique ID for the customer identification
Surname:	Customer Last Name
CreditScore:	Credit score of the customer
Geography:	location of the customers
Gender:	Male or Female
Age:	The age of the customer
Tenure:	Number of years the customer joined the bank
Balance:	Customer balance
NoOfProdcucts:	Number of products the customer is using
HasCrCard:	Binary flag, if the customer holds a credit card or not
IsActiveMember:	Binary flag, the customer is an active member with the bank or not
EstimatedSalary:	Estimated salary of the customer salary in Dollars
Exited:	Binary flag 1 for closing the account and 0 if the customer is retained

Data Collection

The churn dataset, curated by Ayushi Sharma, encapsulates a diverse set of features shedding light on customer behaviour. It includes essential attributes such as Credit Score, Geography, Gender, Age, Tenure, Balance, Num of Products, Has Cr Card, Is Active Member, Estimated Salary, and the binary target variable Exited. These features span demographic, financial, and behavioral dimensions, offering a comprehensive view of customers associated with a certain entity. The dataset aims to facilitate the prediction of customer churn, with the "Exited" variable serving as the focal point for identifying customers who have exited (1) or not (0). This amalgamation of features provides a valuable foundation for developing predictive models and extracting meaningful insights into the dynamics of customer retention and attrition.

Data Preprocessing

In this phase, a meticulous examination of the churn dataset revealed a commendable state of cleanliness, with the absence of any null values. This inherent data quality eliminates the need for extensive imputation or cleansing procedures, ensuring the dataset's reliability for subsequent analyses. Additionally, a crucial step involved standardizing the values, contributing to a consistent and uniform scale across diverse features. This standardization process enhances the model's robustness by preventing certain features from disproportionately influencing the predictive outcomes. The combination of a pristine dataset and standardized values establishes a solid foundation for the forthcoming churn prediction modeling using Artificial Neural Networks (ANN).

ANN : Artificial Neural Network

Artificial Neural Networks (ANNs) represent a cornerstone in the landscape of deep learning, designed to mimic the intricate workings of the human brain. In the realm of ANNs, neurons play a pivotal role, functioning as interconnected nodes arranged in layers. This architecture comprises an input layer, one or more hidden layers, and an output layer, each contributing to the network's ability to process information. The

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

connections between neurons are governed by weights, akin to the strengths of synapses in biological systems.

The magic of ANNs unfolds in their ability to learn and adapt through a training process. During forward propagation, input data traverses the layers, generating predictions. The subsequent assessment of these predictions against actual outcomes leads to the computation of a loss. The pivotal backpropagation phase adjusts weights through optimization algorithms, aligning the model with the desired outcomes. This iterative process occurs over multiple epochs, allowing the network to refine its understanding of complex patterns in the data.

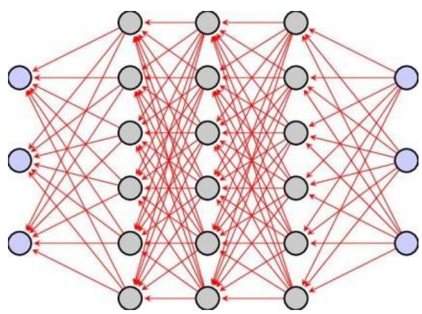


Fig.2. Artificial Neural Networks Architecture

Crucially, ANNs aren't a one-size-fits-all paradigm; they manifest in various types tailored to specific tasks.

• Convolutional Neural Networks (CNNs) excel in image-related tasks.
• Recurrent Neural Networks (RNNs) navigate sequential data.
• Generative Adversarial Networks (GANs) create synthetic data.

The versatility of Multi-Layer Perceptron's (MLPs) spans classification and regression tasks. While their applications are vast, ANNs also face challenges, from mitigating overfitting to grappling with computational complexity and interpretability.

MLP: Multi-Layer Perceptron's

Imagine a neural network as a team of interconnected nodes working together to solve intricate puzzles. At the heart of this network is the Multi-Layer Perceptron (MLP), a dynamic architecture capable of unravelling complex patterns in data.

An MLP consists of an input layer, one or more hidden layers, and an output layer. Each layer comprises interconnected nodes, or neurons, forming a network. Neurons within layers are connected with weights, and each connection has an associated weight denoted as wij, representing the strength of the connection from neuron i in the previous layer to neuron j in the current layer.

Forward Propagation

The process begins with forward propagation, where input data X passes through the layers, and transformations occur at each neuron. The output of a neuron is calculated using the activation function f as follows:

$$j = f\left(\sum_i w_{ij} \times \text{Input}_i + b_j\right) \tag{1}$$

Here, Output *j* is the output of neuron *j*, Input *i* is the input from neuron *i* in the previous layer, *wij* is the weight, and *bj* is the bias term.

Activation Function

The activation function introduces non-linearity, enabling the network to learn complex patterns. Common choices include the sigmoid function ,hyperbolic tangent function tanh(x), and rectified linear unit (ReLU) max(0,x).

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{2}$$

Backpropagation

The crux of learning in an MLP lies in backpropagation. The predicted output is compared to the actual target, and a loss function quantifies the error. Gradient Descent or its variants are then employed to minimize this error by adjusting the weights and biases. The chain rule of calculus is pivotal in computing gradients and updating parameters:

$$\frac{\partial Loss}{\partial w_{ij}} = \frac{\partial Loss}{\partial Output_j} \times \frac{\partial Output_j}{\partial Net_j} \times \frac{\partial Net_j}{\partial w_{ij}} \tag{3}$$

This process iterates over multiple epochs until the model converges to an optimal state. MLPs are trained using labelled datasets, where the model learns to map inputs to desired outputs. Optimization algorithms like Stochastic Gradient Descent (SGD) or Adam fine-tune the model parameters during training.

MLP is an adaptable explorer in the realm of neural networks, deciphering intricate patterns and providing a valuable toolkit for tackling diverse problems.

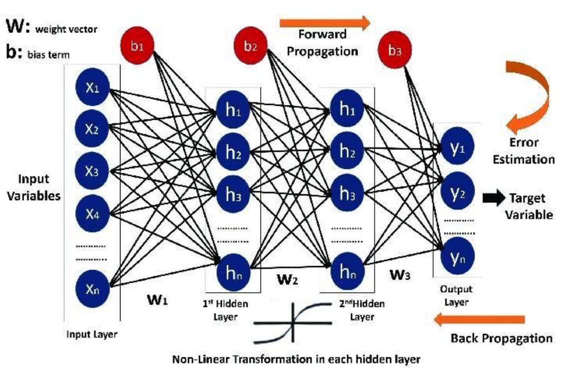


Fig.3. Working of Multi-Layer Neural Network (MLP)

Proposing ANN Model On Churn data

Embarking on the construction of our neural navigator for churn prediction, we carefully assemble the layers and specifications of the neural network using the Keras library. Here's a detailed breakdown of the implementation:

Setting with Sequential

We initiate the construction with the `Sequential` model, which serves as the foundation for creating a linear stack of layers. This intuitive architecture allows us to add one layer at a time, ensuring a streamlined progression.

Crafting the Neural Framework

Constructing our MLP for churn prediction involves a systematic design. The input layer, equipped with 10 units and 'he_uniform' kernel initializer, ushers in non-linearity through 'relu' activation with an input dimension of 12. Progressing, the second layer adopts 8 units, 'he_uniform' initializer, and 'relu' activation, fostering nuanced pattern recognition. The third layer unveils 6 units, persisting with 'he_uniform' and 'relu,' enhancing the network's ability to decipher complex data structures. Culminating our design, the output layer features 1 unit with 'glorot_uniform' kernel initializer and sigmoid activation, tailored for binary churn prediction. This compact MLP architecture ensures a structured and adaptive learning approach.

Activating Layers for Stability

To enhance the stability of our neural navigator, activation functions play a pivotal role. We opt for 'relu' in the initial layers, known for its effectiveness in learning complex representations. The output layer, tailored for binary classification, embraces the 'sigmoid' activation function.

Dealing with the Initializers

The selection of kernel initializers is a strategic move to set the tone for learning. 'he_uniform' and 'glorot_uniform' are chosen for their suitability in different layers,

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

contributing to smoother convergence and preventing issues like vanishing gradients.

Ending the with Dropout

Dropout layers are introduced strategically to prevent overfitting. These layers randomly deactivate a fraction of neurons during training, promoting robustness and preventing reliance on specific neurons.

Our neural navigator, guided by this meticulously crafted architecture, is poised to navigate the seas of churn prediction. The chosen configurations, activations, and initializers collectively contribute to a model capable of discerning patterns and making informed predictions. The sails are set, and the neural odyssey unfolds.

Optimizing the model

By setting these parameters, we define how the neural network will be trained, what algorithm will be used for optimization, how the performance will be measured, and how the model will adjust its weights to minimize the defined loss.

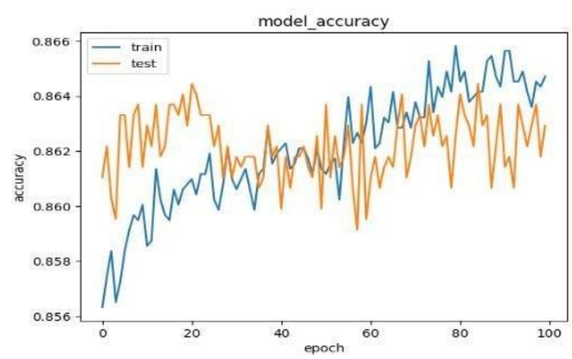
Fig.4. Parameters for compiling the model

- The 'Adamax' optimizer is an extension of the Adam optimizer, which adapts the learning rates for each parameter.
- The 'binary_crossentropy' the loss function used to measure the difference between the true labels and the predicted labels during training.
- This 'metrics=['accuracy']' parameter specifies the evaluation metric(s) to be used, and in this case, we are tracking accuracy.

Results

Examining the accuracy curve plot

The resulting chart provides a visual representation of how the training and validation accuracy change over epochs. Monitoring these curves helps in assessing the model's performance and understanding whether it is learning well from the training data and generalizing to new, unseen data.



Evaluating the Classification Report and Confusion Matrix

The evaluation metrics presented in the table(2.1) offer a comprehensive assessment of the classification model's performance. Precision, which measures the accuracy of positive predictions, is 0.88 for class 0 and 0.73 for class 1. Recall, indicating the ability to capture actual positive instances, stands at 0.95 for class 0 and 0.51 for class 1. The F1score, a harmonic mean of precision and recall, reflects the balanced performance with values of 0.92 for class 0 and 0.60 for class 1. The overall accuracy of the model is 0.86, showcasing the proportion of correctly predicted instances. Macro average metrics

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

provide a broader view, yielding precision of 0.81, recall of 0.73, and an F1-score of 0.76. Weighted averages, considering class prevalence, show precision at 0.85, recall at 0.86, and an F1-score of 0.85. These metrics collectively offer insights into the model's effectiveness across different aspects, enabling a nuanced understanding of its classification performance.

Table 2. Classification Report

	precision	recall	f1-score	support
0	0.88	0.95	0.92	1591
1	0.73	0.51	0.60	409
accuracy			0.86	2000
macro avg	0.81	0.73	0.76	2000
weighted avg	0.85	0.86	0.85	2000

The confusion matrix summarizes the performance of a binary classification model. In this specific case, the model correctly identified 207 instances as positive (True Positives) and accurately recognized 1516 instances as negative (True Negatives). However, it misclassified 75 instances as positive when they were actually negative (False Positives) and erroneously labeled 202 instances as negative when they were positive (False Negatives). The overall accuracy of the model on the given dataset is approximately 86%, reflecting its ability to make correct predictions, but there is room for improvement in reducing both false positive and false negative rates for a more robust classification.

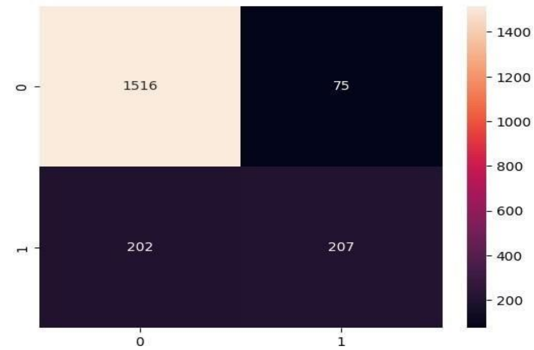


Fig.6. Confusion Matrix

Conclusion

At Last our study embraced the use of artificial neural networks (ANN), specifically a multi-layer perceptron (MLP), to enhance customer churn prediction models. The MLP, designed to autonomously learn from extensive datasets, demonstrated promising outcomes, emphasizing the potential of ANN in unravelling the complexities of customer retention. This study contributes to the evolving landscape of customer analytics by advocating for the adoption of sophisticated artificial intelligence techniques. Our findings underscore the importance of continuous refinement and adaptation for more accurate churn predictions, positioning ANN as a transformative force in navigating contemporary business challenges.

Future Expansion

In terms of our future directions, delving into advanced neural network architectures like recurrent neural networks (RNNs) or long short-term memory networks (LSTMs) can be pivotal for capturing temporal dependencies in customer behaviours. Furthermore, incorporating diverse data sources and employing unsupervised learning techniques may enhance the model's depth of understanding. Exploring methods for enhancing the interpretability of neural networks and addressing concerns related to data privacy and ethics represents an essential path forward. The model should be continuously refined and adapted to stay attuned to evolving business landscapes, ensuring a dynamic approach to customer churn prediction within the realm of artificial intelligence.

Acknowledge

We extend our heartfelt gratitude to Dr. Kompalli Udaya Sri, M.B.A., M.Sc.,(IT), MTech.(CSE), MA(Astro), Ph.D., Gold Medal, World Book of Records, FSIESRP, for her invaluable guidance and mentorship. Her vast knowledge and unwavering support have illuminated our path throughout this research journey.

This research is not just an academic milestone; it's a celestial event in our lives, and Dr. Kompalli Udaya Sri has been the guiding North Star, steering us toward intellectual horizons we hadn't even dared to dream of. We extend our deepest appreciation for her unwavering dedication and limitless encouragement.

References

1. Germann, F., Lilien, G., Moorman, and C., Fiedler, L., "Driving customer analytics from the top. Customer Needs and Solutions", Customer Needs and Solutions, 7(3):43–61, (2021).

2. Seymen, Omer & Hiziroglu, Abdulkadir & Dogan, Onur & Er, Orhan & Ölmez, Emre. (2022). Customer Churn Prediction Using Ordinary Artificial Neural Network and Convolutional Neural Network Algorithms: A Comparative Performance Assessmen t.GAZI UNIVERSITY JOURNAL OF SCIENCE. 36. 10.35378/gujs.992738.

3. Baby, B., Dawod, Z., Sharif, S. and Elmedany, W. 2023. Customer Churn Prediction Model Using Artificial Neural Networks (ANN): A Case Study in Banking. 3ICT 2023: International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies. University of Bahrain, Bahrain 20 - 21 Nov 2023 IEEE.

4. Germann, F, Lilien, G.L, Fiedler, L, and Kraus, M., "Do retailers benefit from deploying customer analytics?", Journal of Retailing, 90(4):587-593, (2014).

5. Mishra, A. and Reddy, U. S., "A novel approach for churn prediction using deep learning", Proceedings of the IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Coimbatore India, 1–4, (2017).

6. Saha, L., Tripathy, H.K., Gaber, T., El-Gohary, H., & El-kenawy, E.M. (2023). Deep Churn Prediction Method for Telecommunication Industry. Sustainability.

7. Aspect Software. Aspect Consumer Index 2020. https://www.aspect.com/the-aspect-consumer-experience-index-2020, 2020. Online, accessed 10 April 2021.

8. Uzair, A., Khan, A., Saddam, H. K., Basit, A., Haq, I., and Lee, Y. S., "Transfer learning and meta classification based deep churn prediction system for telecom industry", arXiv preprint arXiv:1901.06091, (2019).

9. Ladyzynski, P., Zbikowski, K. and Gawrysiak, P., "Direct marketing campaigns in retail banking with the use of deep learning and random forests", Expert Systems with Applications, 134: 28–35, (2019).

10. Umayaparvathi, V. and Iyakutti, K., "Automated feature selection and churn prediction using deep learning models", International Research Journal of Engineering and Technology (IRJET), 4(3): 1846–1854, (2017).

11. Kurani, Akshit, Pavan Doshi, Aarya Vakharia and Manan Shah. "A Comprehensive Comparative Study of Artificial Neural Network (ANN) and Support Vector Machines (SVM) on Stock Forecasting." Annals of Data Science 10 (2021): 183-208.