COGNITIVE INSIGHTS INTO CUSTOMER BEHAVIOR: ANALYZING INTERNATIONAL BANK CUSTOMERS USING ARTIFICIAL NEURAL NETWORKS

Imran Hussain, Sherien Zafar, Richa Gupta

Jamia Hamdard, Department of Computer Science & Engineering, School of Engineering Sciences & Technology, New Delhi, INDIA-110062

ihussain@jamiahamdard.ac.in,

sherin.zafar@jamiahamdard.ac.in,

richagupta@jamiahamdard.ac.in

ABSTRACT

This research paper aims to address the problem of customer attrition in an international bank and propose a solution using artificial neural networks. The problem statement revolves around predicting whether a customer will leave the bank after one year of using its services. The study focuses on a sample of 10,000 customers, employing demographic segmentation to analyze factors such as age, salary, gender, and the type of account (loan, saving, fixed deposit, or random). By utilizing an artificial neural network, the proposed solution aims to predict customer churn within a six-month timeframe. The main objective is to create a demographic segmentation model that can identify customers at the highest risk of leaving the bank. The outcome of this research will provide valuable insights for the bank to develop effective customer retention strategies.

KEYWORDS: *Customer behavior, International bank, Artificial neural networks, Customer attrition, Demographic segmentation*

1. INTRODUCTION

In the dynamic landscape of the banking industry, understanding customer behavior [1] is of utmost importance for organizations striving to enhance customer satisfaction and retention. This research delves into the realm of an international bank, facing a pressing challenge that demands a comprehensive analysis of customer behavior. With millions of customers spread across three European countries—France, Spain, and Germany—this fictional bank embarked on a crucial endeavor six months ago.

Recognizing the need to address the issue at hand, the bank carefully selected a sample of 10,000 customers for in-depth analysis. Various attributes such as customer ID, surname, credit score, geography, gender, age, and tenure with the bank were meticulously recorded. Additionally, information about the balance of products, including savings accounts, credit cards, and loans, was collected. The bank also sought to gauge customer activity, whether they were active members based on their online banking usage or recent transactions. While the actual salary of the customers remained unknown, estimates were derived using other available data points.

Over the subsequent six months, the bank monitored the selected customers closely, observing their interactions and financial behaviors. The key focus was to identify which customers remained loyal and who decided to sever ties with the bank during this period. Each customer was marked with a flag indicating their retention or attrition status, providing valuable insights into customer churn patterns.

The ultimate objective of this research is to develop a robust demographic segmentation model that can effectively identify customers at the highest risk of leaving the bank. While the implications of this study hold significance for customer-centric organizations, the potential applications of demographic

DOI: 10.5281/zenodo.10935279

segmentation [2] extend beyond the banking sector. By leveraging prior experiences and constructing accurate models, organizations can make informed decisions regarding customer reliability, loan approvals, or even detecting fraudulent transactions.

In scenarios where binary outcomes are involved, and a multitude of independent variables are present, the knowledge gained from this research holds promise for diverse applications. The findings obtained through this study can be applied in various domains to comprehend the factors that significantly influence outcomes.

By undertaking this research, we aim to contribute to the advancement of customer-centric strategies and provide a valuable framework for organizations to proactively address customer attrition challenges.

2. Literature Review

Understanding customer behavior and predicting customer attrition are crucial areas of research in the banking industry. Numerous studies have explored the application of artificial neural networks (ANNs) to gain cognitive insights into customer behavior and improve customer retention strategies.

One relevant study by Fujo, S. W., Subramanian et al. investigated the use of ANNs for predicting customer churn in the telecommunications sector [3]. The authors employed demographic and usage data to train an ANN model, achieving high accuracy in predicting customer attrition. Their findings emphasized the effectiveness of ANNs in identifying key factors influencing customer behavior and churn.

In the banking context, Khan, Y., Shafiq, S et al. conducted a comprehensive analysis of customer attrition using ANNs [4]. Their study focused on various customer segments and incorporated factors such as account types, transaction history, and customer demographics. The results demonstrated the capability of ANNs to accurately predict customer churn and provided insights into the most influential factors driving attrition.

Furthermore, a study by Lin, C. P., Huang et al. examined the impact of demographic segmentation on customer behavior in the banking industry [5]. The authors utilized ANNs to analyze customer data and identify patterns related to customer attrition. Their research highlighted the significance of demographic variables, such as age, income, and occupation, in understanding customer behavior and predicting churn.

Another relevant contribution by Lin, C. P. explored the integration of ANNs with customer relationship management (CRM) systems in the banking sector [6]. Their study emphasized the importance of leveraging ANNs to enhance customer segmentation and develop personalized retention strategies. The authors highlighted the potential of ANNs to uncover hidden patterns and improve customer targeting.

Overall, the existing literature demonstrates the potential of ANNs in analyzing customer behavior and predicting customer attrition in the banking industry. These studies provide a foundation for the current research, which aims to leverage ANNs to gain cognitive insights into customer behavior and develop an effective demographic segmentation model for predicting customer churn in an international bank. By incorporating relevant demographic variables and utilizing advanced ANN techniques, this research aims to contribute to the existing body of knowledge and provide valuable insights for banks to implement proactive customer retention strategies.

3. Artificial Neural Network (ANN)

The proposed Artificial Neural Network (ANN) model in this research paper consists of interconnected nodes or artificial neurons. The model follows a feedforward architecture, where information flows from the input layer through one or more hidden layers to the output layer.

The input layer of the ANN receives customer-related features, including demographic information such as age, salary, and gender, as well as information about the type of account (loan, saving, fixed deposit, or random). These input features are processed and propagated through the network.

The hidden layers, located between the input and output layers, perform the internal computations and feature extraction. Each neuron in the hidden layers receives weighted inputs from the previous layer and applies an activation function, such as the rectifier function, to introduce non-linear properties and capture complex relationships between the input variables.

The output layer represents the final layer of the ANN and provides the prediction or classification of interest, which in this case is whether a customer is likely to leave the bank within the next six months. The activation function for the output layer is chosen as the sigmoid function, which maps the network's weighted inputs to a probability value between 0 and 1.

During the training phase, the ANN model learns to adjust the weights associated with each input to minimize the discrepancy between the predicted outputs and the actual customer churn data. This is achieved using the backpropagation algorithm, which iteratively updates the weights based on the computed errors.

By utilizing this ANN model, the research aims to effectively analyze customer behavior, identify patterns, and predict customer churn in the international bank under investigation. The ANN's ability to capture non-linear relationships and process complex data makes it a suitable approach for addressing the problem of customer attrition in the banking industry.

4. METHODOLOGY

4.1. Dataset

The dataset used in this study comprises 10,000 customers who were randomly selected from a Fictional International bank operating in three European countries, namely France, Spain, and Germany. This dataset was meticulously curated by conducting a comprehensive study on customer behavior and their utilization of various bank services. By tracking the churn rates over 6 months, the dataset provides valuable insights into customer attrition patterns.

In our research paper, we have incorporated a comprehensive set of attributes for data analysis in the context of customer behavior and attrition prediction in an international bank. These attributes include Customer-ID, Gender, Age, tenure, credit score, Estimated Salary, Exit rate, and more.

By including these attributes in our analysis, we aimed to capture a wide range of factors that could potentially influence customer behavior and contribute to the prediction of customer churn. Customer ID serves as a unique identifier for each customer, allowing us to track individual behavior patterns. Gender and Age provide demographic information that can help uncover trends and preferences within different customer segments.

Tenure represents the duration of the customer's relationship with the bank, which can be a crucial factor in assessing customer loyalty. CreditScore and Estimated Salary offer insights into the customer's financial profile and can contribute to understanding their propensity to leave the bank. The Exit rate attribute specifically measures the rate at which customers have discontinued their services.

By analyzing and considering these attributes collectively, we aimed to gain a deeper understanding of the customer dynamics and uncover significant patterns that could aid in predicting customer churn and developing effective retention strategies for the bank.

It is important to note that these attributes were carefully chosen based on their relevance and potential impact on customer behavior. The inclusion of such attributes enables a comprehensive analysis and provides valuable insights into the factors that drive customer attrition in the international banking context.

DOI: 10.5281/zenodo.10935279

dataset - DataFrame

Index	RowNumber	CustomerId	Surname	CreditScore	Geography	Gender	Age	Tenure	Balance	NumOfProducts	HasCrCard	IsActiveMember	EstimatedSalary	Exited
0	1	15634602	Hargrave	619	France	Female	42	2	0	1	1	1	101349	1
1	2	15647311	Hill	608	Spain	Female	41	1	83807.9	1	0	1	112543	0
2	3	15619304	Onio	502	France	Female	42	8	159661	3	1	0	113932	1
3	4	15701354	Boni	699	France	Female	39	1	0	2	0	0	93826.6	0
4	5	15737888	Mitchell	850	Spain	Female	43	2	125511	1	1	1	79084.1	0
5	6	15574012	Chu	645	Spain	Male	44	8	113756	2	1	0	149757	1
6	7	15592531	Bartlett	822	France	Male	50	7	0	2	1	1	10062.8	0
7	8	15656148	Obinna	376	Germany	Female	29	4	115047	4	1	0	119347	1
8	9	15792365	He	501	France	Male	44	4	142051	2	0	1	74940.5	0
9	10	15592389	H?	684	France	Male	27	2	134604	1	1	1	71725.7	0
10	11	15767821	Bearce	528	France	Male	31	6	102017	2	0	0	80181.1	0
11	12	15737173	Andrews	497	Spain	Male	24	3	0	2	1	0	76390	0
12	13	15632264	Kay	476	France	Female	34	10	0	2	1	0	26261	0
13	14	15691483	Chin	549	France	Female	25	5	0	2	0	0	190858	0
14	15	15600882	Scott	635	Spain	Female	35	7	0	2	1	1	65951.6	0
15	16	15643966	Goforth	616	Germany	Male	45	3	143129	2	0	1	64327.3	0
16	17	15737452	Romeo	653	Germany	Male	58	1	132603	1	1	0	5097.67	1

Format Resize Background color Column min/max OK Cancel

Figure 1. Dataset of 10,000 Customers of an International Bank (Europe): A

dataset - DataFrame

Index	RowNumber	CustomerId	Surname	CreditScore	Geography	Gender	Age	Tenure	Balance	NumOfProducts	HasCrCard	IsActiveMember	EstimatedSalary	Exited
9983	9984	15656710	Cocci	613	France	Male	40	4	0	1	0	0	151325	0
9984	9985	15696175	Echezonachukwu	602	Germany	Male	35	7	90602.4	2	1	1	51695.4	0
9985	9986	15586914	Nepean	659	France	Male	36	6	123841	2	1	0	96833	0
9986	9987	15581736	Bartlett	673	Germany	Male	47	1	183580	2	0	1	34047.5	0
9987	9988	15588839	Mancini	606	Spain	Male	30	8	180308	2	1	1	1914.41	0
9988	9989	15589329	Pirozzi	775	France	Male	30	4	0	2	1	0	49337.8	0
9989	9990	15605622	McMillan	841	Spain	Male	28	4	0	2	1	1	179437	0
9990	9991	15798964	Nkemakonam	714	Germany	Male	33	3	35016.6	1	1	0	53667.1	0
9991	9992	15769959	Ajuluchukwu	597	France	Female	53	4	88381.2	1	1	0	69384.7	1
9992	9993	15657105	Chukwualuka	726	Spain	Male	36	2	0	1	1	0	195192	0
9993	9994	15569266	Rahman	644	France	Male	28	7	155060	1	1	0	29179.5	0
9994	9995	15719294	Wood	800	France	Female	29	2	0	2	0	0	167774	0
9995	9996	15606229	Obijiaku	771	France	Male	39	5	0	2	1	0	96270.6	0
9996	9997	15569892	Johnstone	516	France	Male	35	10	57369.6	1	1	1	101700	0
9997	9998	15584532	Liu	709	France	Female	36	7	0	1	0	1	42085.6	1
9998	9999	15682355	Sabbatini	772	Germany	Male	42	3	75075.3	2	1	0	92888.5	1
9999	10000	15628319	Walker	792	France	Female	28	4	130143	1	1	0	38190.8	0

Format Resize Background color Column min/max OK Cancel

Figure 2. Dataset of 10,000 Customers of an International Bank (Europe): B

4.2. Classification Approach

Classification involves the identification of classifiers or models that can assign records to distinct classes [7]. In the proposed model, this process entails the utilization of two types of datasets: the training set and the testing set, each serving a specific purpose in constructing and evaluating the model. The training dataset consists of customer behavior records with predefined class labels, which are used to train the model and determine parameters or characteristics that differentiate between the two classes: "Exited" and "Non-exited." This phase is known as the learning process.

Following that, the testing dataset, comprising pre-classified labels, is employed to assess the model's performance. If the outcomes of the testing phase fall short of expectations, further training iterations might be necessary to enhance the model's accuracy. Conversely, if the results are considered satisfactory, the trained model can be utilized to make predictions regarding the likelihood of a customer leaving the bank within the upcoming 6 months. By adopting this classification approach, the research

DOI: 10.5281/zenodo.10935279

aims to effectively distinguish between customers who are more likely to churn and those who are not, contributing to the development of a robust predictive model for customer attrition.

5. INTRODUCTION TO ARTIFICIAL NEURONS

Artificial neurons serve as fundamental building blocks within artificial neural networks [8], representing a mathematical function that emulates the behavior of biological neurons. These neurons play a vital role in the processing and transmission of information in the network. An artificial neuron is capable of receiving multiple inputs, each representing excitatory and inhibitory postsynaptic potentials at the neural dendrites. These inputs are weighted individually, and their sum is then passed through an activation function.

The activation function introduces non-linearity to the output of the artificial neuron [9]. Commonly, activation functions exhibit a sigmoid shape, allowing for a smooth transition between different states. However, alternative non-linear functions such as piecewise linear functions can also be employed as activation functions.

By leveraging the capabilities of artificial neurons and their activation functions, artificial neural networks can effectively model complex relationships and make non-linear transformations on input data. This enables neural networks to learn and adapt to various types of data, making them powerful tools for tasks such as classification, regression, and pattern recognition.

5.1 Basic Structure

Consider an artificial neuron with m + 1 inputs, represented by signals x0 through xm and weights w0 through wm. To introduce a bias input, the value +1 is typically assigned to x0, resulting in wk0 being equivalent to the bias term bk. As a result, the neuron effectively processes m actual inputs, ranging from x1 to xm. By incorporating this basic structure, the neuron can capture and process relevant information to perform its intended computations.

The output of the kth neuron is: The output of the *k*th neuron is:

$$y_k = \varphi\left(\sum_{j=0}^{m} w_{kj} x_j\right)$$

Where φ (phi) is the transfer function.

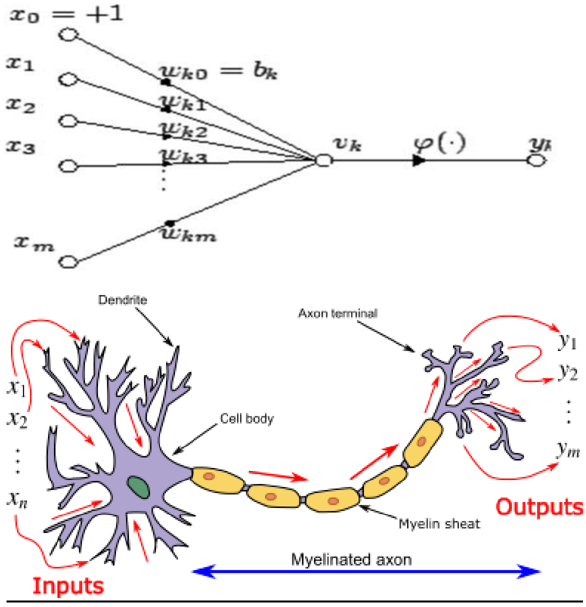


Figure 3. Neuron and myelinated axon, with signal flow from inputs at Dendrites to outputs at axon terminals

DOI: 10.5281/zenodo.10935279

6. Activation Functions

Activation functions play a critical role in the learning process of an Artificial Neural Network (ANN) by enabling the comprehension of complex and non-linear relationships between input variables and the response variable. They introduce non-linear characteristics to our network, allowing for the transformation of input signals from one node into output signals. These output signals then serve as inputs for the subsequent layer within the network.

In the context of ANN, the activation function, denoted as f(x), is applied to the summation of the products of inputs (X) and their corresponding weights (W). This process facilitates the computation of the output for a given layer, which is subsequently used as input for the subsequent layer [10]. By employing activation functions, ANNs can effectively capture and represent intricate patterns and nonlinear mappings, enhancing the network's ability to understand and process complex data.

Most popular types of Activation functions -

1. Sigmoid or Logistic
2. Threshold Function
3. Tanh—Hyperbolic tangent
4. ReLu -Rectified linear units

- In our problem we'll use the rectifier function for hidden layers and the Sigmoid function for output layer.

6.1 Sigmoid Function

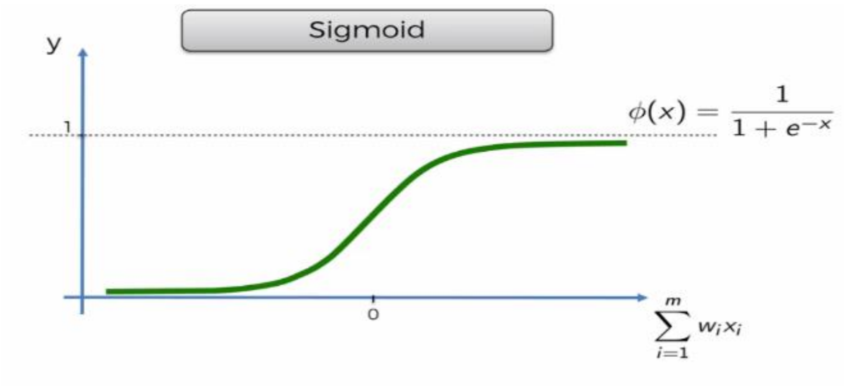


Figure 4. Sigmoid Function

6.2 Threshold Function

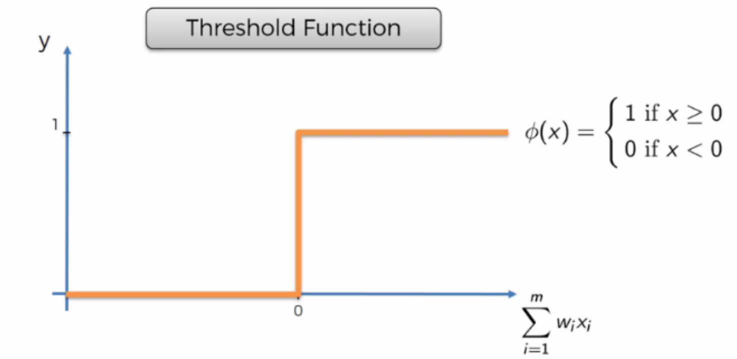


Figure 5. Threshold Function

DOI: 10.5281/zenodo.10935279

6.3 Rectifier Function

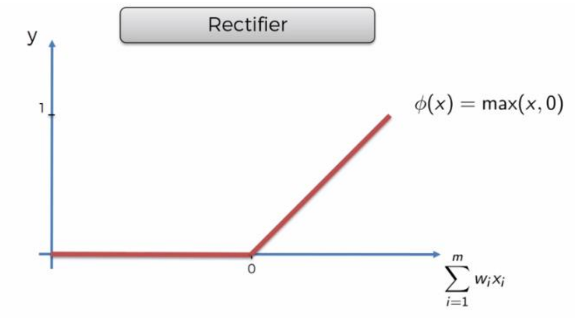


Figure 6. Rectifier Function

6.4 Tanh—Hyperbolic tangent

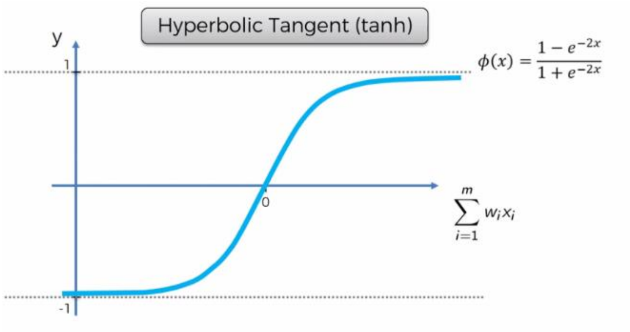


Figure 7. Tanh—Hyperbolic tangent

For our specific use case, the rectifier function [11] will be employed for the hidden layers, while the output layer will utilize the Sigmoid function.

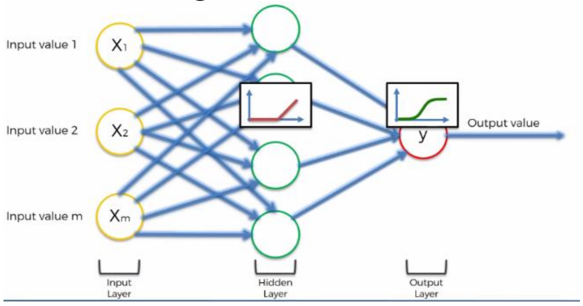


Figure 8. Rectifier function for hidden layers and sigmoid function for the output layer

7. Stochastic Gradient Descent

Stochastic gradient descent (SGD), also known as incremental gradient descent, is an iterative optimization technique specifically developed for differentiable objective functions. It acts as a stochastic approximation of the gradient descent optimization algorithm. The foundations of SGD can be attributed to the work of Herbert Robbins and Sutton Monro, who introduced the concept of stochastic approximation in their seminal 1951 article. Additional information on stochastic approximation can be found in the referenced publication [12]. The term "stochastic" in SGD stems from the random selection or shuffling of samples during the optimization process. This contrasts with standard gradient descent, which processes the entire training set as a single group, or with a sequential approach that adheres to the order of appearance in the training set.

DOI: 10.5281/zenodo.10935279

By randomly selecting samples, SGD introduces variability into the optimization process, allowing it to navigate complex landscapes and potentially escape local optima. This stochastic nature makes SGD particularly useful in scenarios with large datasets, as it enables efficient computation by considering only a subset of samples in each iteration.

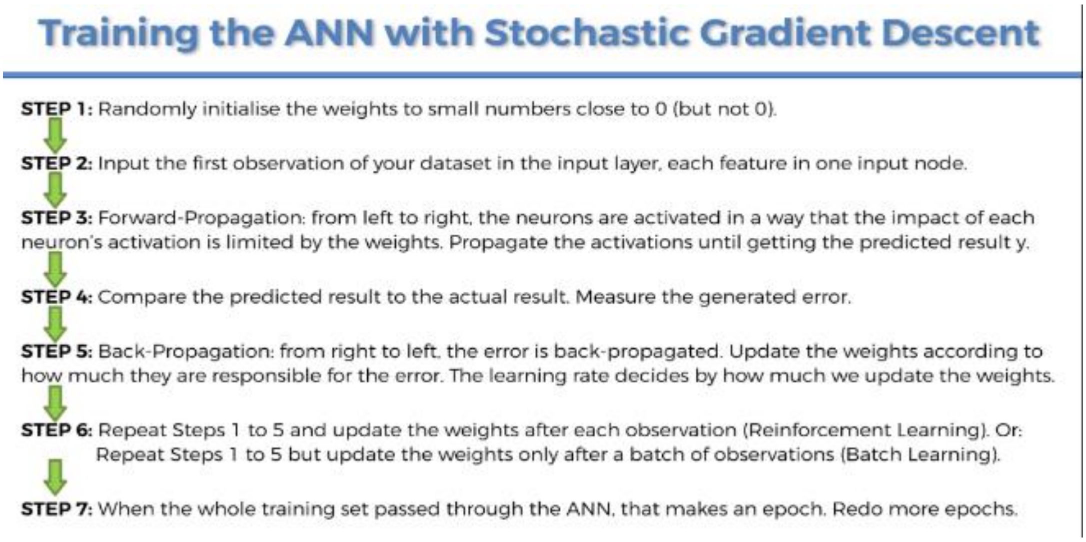


Figure 9. Step-by-step procedure of training the ANN with Stochastic gradient Descent

7.1 Back Propagation

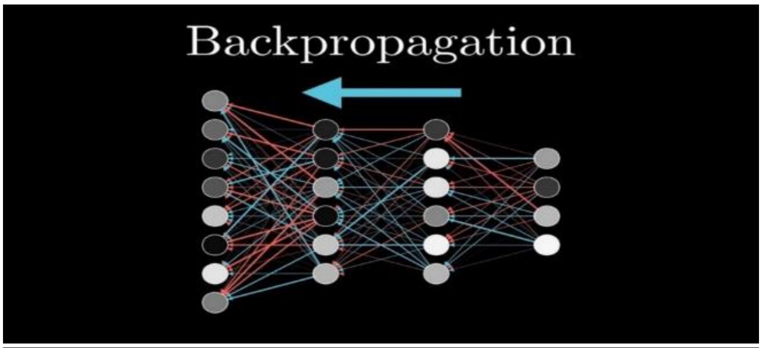


Figure 10. Back Propagation

Ever since the introduction of non-linear recursive functions in Machine Learning, such as Artificial Neural Networks, their applications have experienced significant growth. In this context, training a Neural Network properly emerges as the pivotal factor in creating a dependable model. However, the concept of "Back-propagation" [13] often perplexes newcomers to Deep Learning, and even many industry professionals are unfamiliar with its inner workings.

Back-propagation constitutes the fundamental principle behind neural network training. It involves the iterative adjustment of the network's weights based on the error rate, or loss, computed in the previous iteration, also known as an epoch. Precise fine-tuning of the weights aims to minimize the error rate, ultimately enhancing the model's reliability and improving its ability to generalize to unseen data.

8. ANN MODEL DESIGN

A neural network comprises interconnected nodes, including input layer neurons, one or more hidden layers, and an output layer, as depicted in Figure 11.

DOI: 10.5281/zenodo.10935279

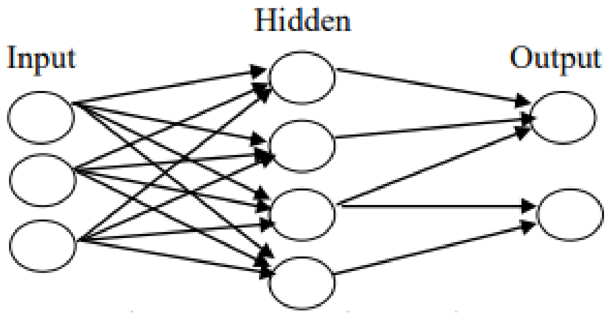


Figure 11. Neural Network Structure

The middle layer, situated between the input layer and the output layer, serves as an internal representation of the input pattern. Within this layer, patterns are extracted by distilling the input characteristics and subsequently transmitted to the output layer. The output layer generates a classification judgment based on the input patterns. Hence, the middle layer can be referred to as the distilling character layer. In the present study, the input and output variables were initially classified. The input variables encompassed customer satisfaction factors, such as geographic position, product variety, price level, product quality, payment waiting time, service attitude, and shopping environment. The output variable represented a single measurement value.

The construction of the ANN model [14] followed a six-step procedure. Firstly, the data was defined and presented to the ANN as input data patterns along with their corresponding desired targets. Secondly, the data was divided into training and validation sets. The training set was utilized to facilitate the model's learning process and development, while the validation set served to assess the model's predictive capability and determine the appropriate stopping point for training. Thirdly, the structure of the ANN was defined by specifying the number of hidden layers to be created and the number of neurons within each hidden layer. Fourthly, all ANN parameters were set before commencing the training process. Subsequently, the training process was initiated, involving the computation of outputs from the input data and weights. The backpropagation algorithm was employed to train the ANN by adjusting the weights to minimize the discrepancy between the current ANN output and the desired output. During training, a threshold value of 0.2 was applied to prevent predicted values exceeding 0.8, ensuring that the output does not fall below 0.2. Finally, an evaluation process was conducted to assess the ANN's learning capability. This process involved periodically pausing the training and testing the ANN's performance until an acceptable result was achieved. Once an acceptable result was obtained, the ANN model was considered fully trained and ready for implementation. The model approach adopted, aiming for clarity and ease of implementation, is illustrated in the figure.

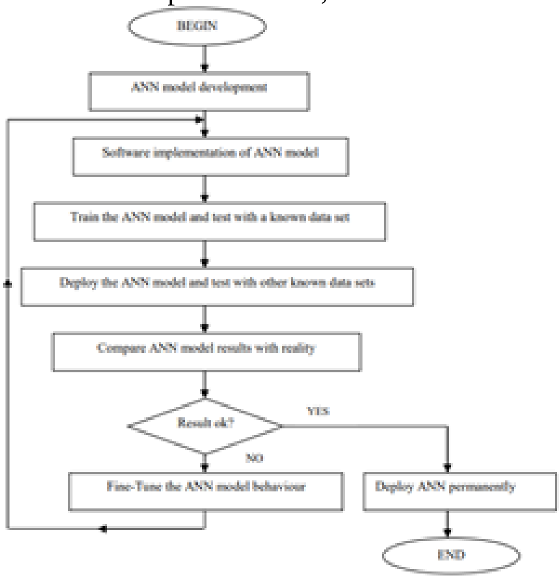


Figure 12. Steps in ANN model design

DOI: 10.5281/zenodo.10935279

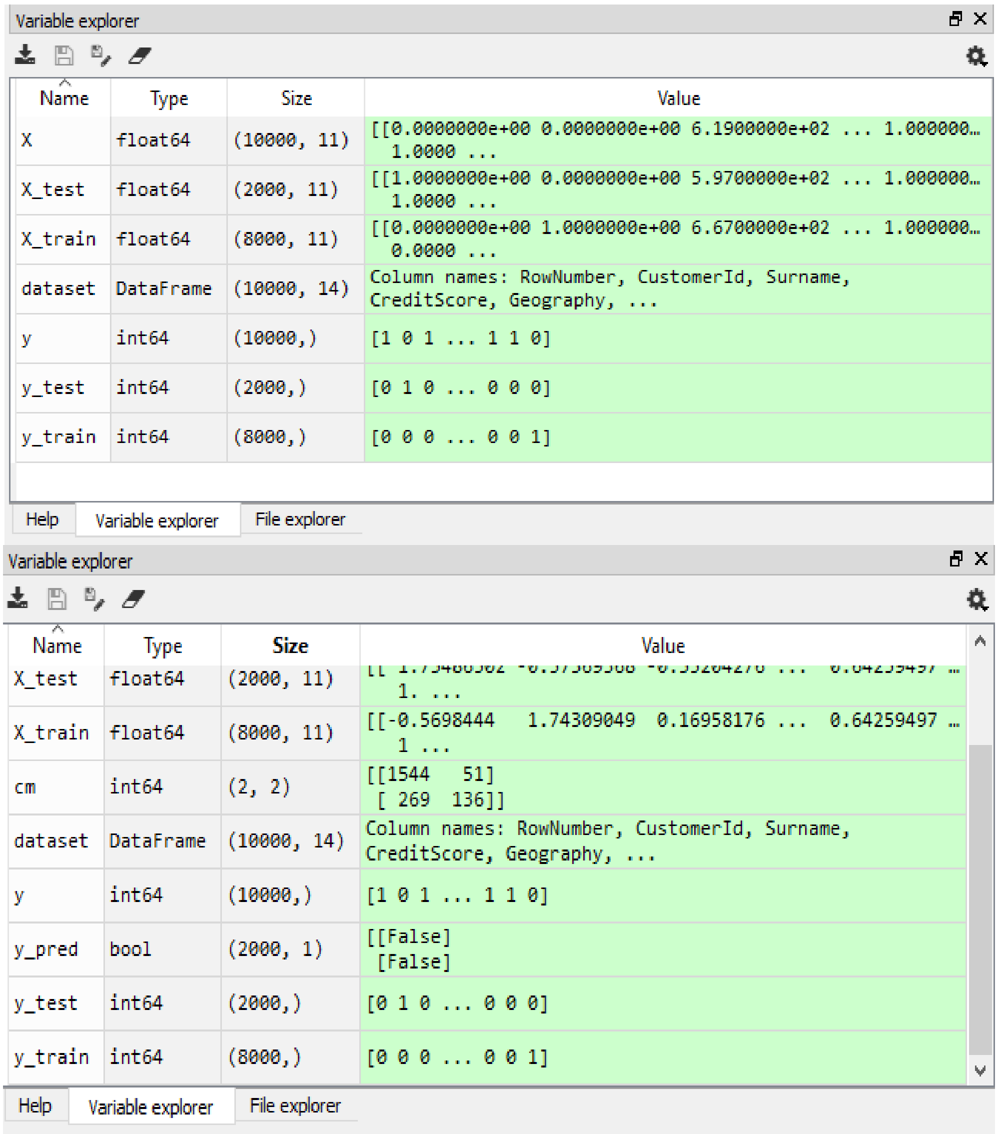
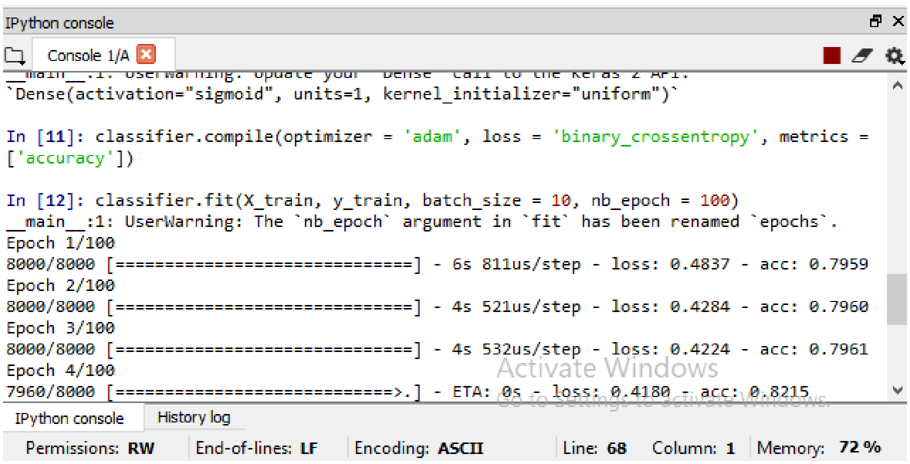


Figure 13. Dataset processed and divided into different labels for compiling the ANN

9. TRAINING THE ANN



DOI: 10.5281/zenodo.10935279

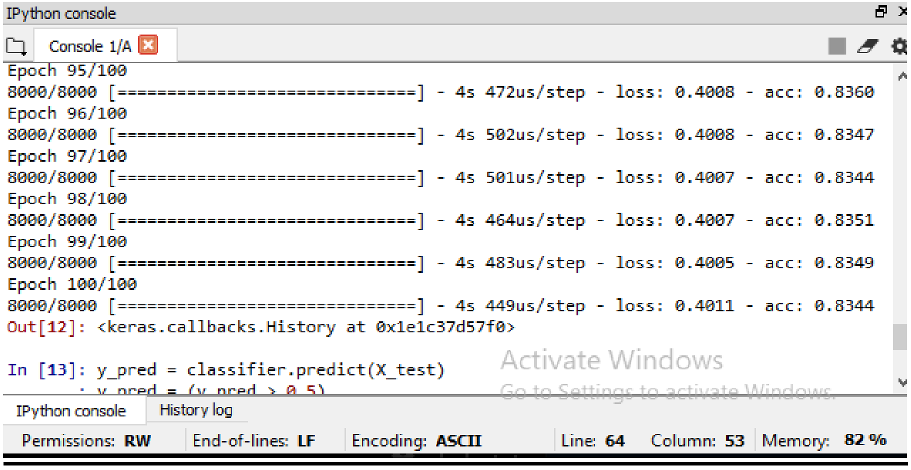


Figure 14. Training ANN

We trained our data on 8,000 customers and tested it on 2,000 customers, achieving an accuracy of 83.4%. The training process [15] involved 100 epochs, with a batch size of 10. The hidden layer activation function was set as the rectifier function, while the output layer utilized the sigmoid function as the activation function.

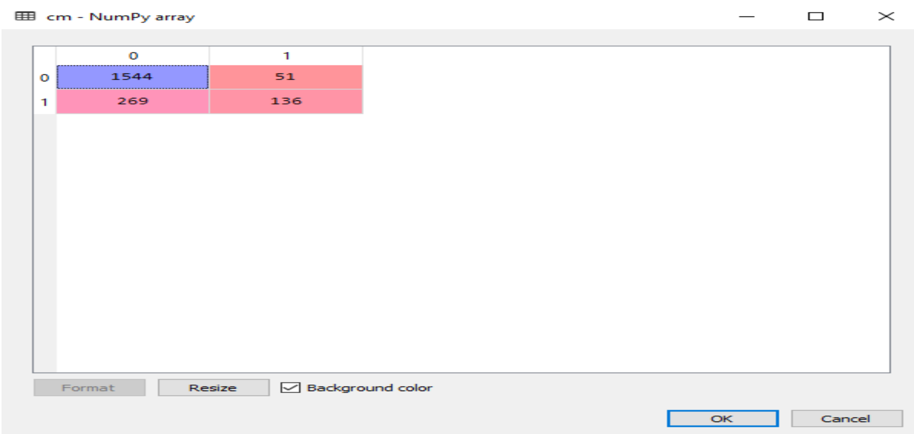
To ensure a comprehensive evaluation, we divided our dataset into two distinct sets: a training set and a testing set. The training set, which consisted of a substantial portion of the data, was used to train the ANN model by iteratively adjusting the weights and optimizing its performance. This iterative process allowed the model to learn from the training data and improve its ability to predict customer attrition.

After the training phase, we evaluated the performance of the trained model using the testing set, which was kept completely separate and not involved in the training process. The purpose of the testing set was to assess how well the model generalized to unseen data and to provide an unbiased estimate of its predictive accuracy.

We rigorously evaluated the performance of the ANN model on both sets. The accuracy of the training set serves as an indication of how well the model learned from the provided data, while the accuracy of the testing set provides insights into the model's ability to generalize and make accurate predictions on new, unseen data.

10.CONCLUSION

A confusion matrix is a widely employed table for assessing the effectiveness of a classification model. It allows for the visualization and identification of class confusion, where one class is mislabelled as another. Performance measures are typically derived from the confusion matrix.



DOI: 10.5281/zenodo.10935279

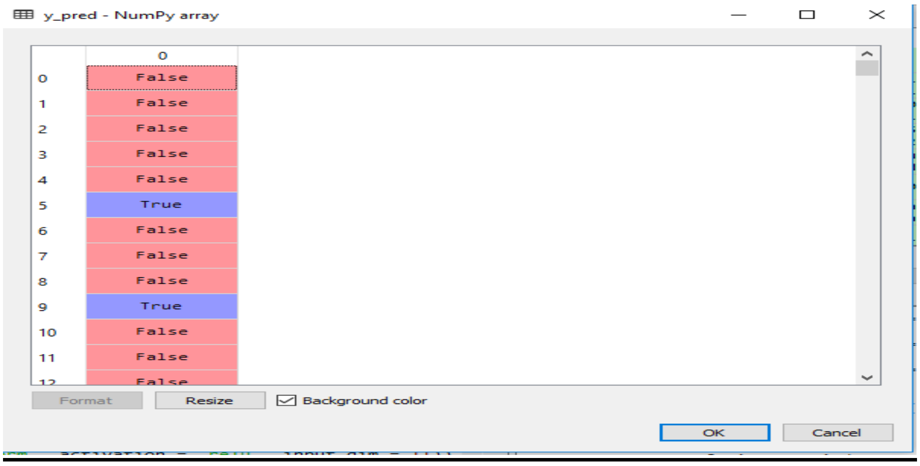


Figure 15. *Y_Pred (True or False value)*

In our study, we trained an artificial neural network on a training set and now it's time to make predictions on the test set. Fortunately, we don't need to make any changes to the classifier object we created with scikit-learn. We can use the same prediction method to make predictions on the test set. By executing this method, we obtain the predicted probabilities for each of the 2,000 customers in the test set.

For example, if we look at the first probability of 20 percent, it indicates that the first customer in the test set (indexed as zero) has a 20 percent chance of leaving the bank. Making predictions on the test set allows us to evaluate the model's performance on new observations. We aim to assess if the accuracy achieved on the training set holds up on unseen data. To evaluate this accuracy, we utilize the confusion matrix.

If we achieve good accuracy on the test set, close to the 86 percent obtained on the training set, it would validate the model. This validated model can then be applied to all the customers of the bank, ranking them based on the probabilities of leaving the bank. For instance, the bank can focus on the top 10 percent of customers with the highest probabilities and analyze them further using data mining techniques to understand why they are more likely to leave. This neural network provides added value to the bank by enabling targeted measures to prevent customer churn.

To analyze the results using the confusion matrix, we need to convert the predicted probabilities into predicted results in the form of "True" or "False." We accomplish this by applying a threshold, such as 0.5, where values above the threshold are predicted as "True" and values below as "False."

After updating our vector Y_pred with the converted results, we can observe the predictions in the form of "False" or "True." For example, the first customer in the test set is predicted to stay with the bank, while the sixth customer is predicted to leave. We can then calculate the accuracy of new observations by comparing the number of correct predictions to the total number of predictions.

In our case, out of the 2,000 new observations, we have 1,550 correct predictions and 230 incorrect predictions. This accuracy confirms the validation of our model. The bank can now utilize this model to rank their customers based on their probabilities of leaving, from the highest to the lowest.

To generate models using unsupervised learning methods, neural networks can be applied to the data. This involves techniques such as profiling, classification, clustering, and transaction rules, particularly when there are no prior sets of normal and abnormal observations.

In summary, to achieve a holistic Customer Relationship Management (CRM) solution in banking, it is crucial to divide the CRM implementation into manageable segments. All business units and departments within the bank need to be seamlessly integrated using an adaptable approach. Factors like safety and health support, systems compatibility, contract structure, distribution flexibility, and effective troubleshooting and problem-solving should be taken into account to ensure the successful implementation of an efficient CRM solution. Future research can delve into exploring the implementation of internet banking and ATM transactions within the banking industry, further enhancing CRM practices and customer service.

DOI: 10.5281/zenodo.10935279

REFERENCES

[1] Sethna, Z., & Blythe, J. (2019). Consumer behaviour. Sage.

[2] Elena-Bucea, A., Cruz-Jesus, F., Oliveira, T., & Coelho, P. S. (2021). Assessing the role of age, education, gender and income on the digital divide: Evidence for the European Union. Information Systems Frontiers, 23, 1007-1021.

[3] **T1** Fujo, S. W., Subramanian, S., & Khder, M. A. (2022). Customer churn prediction in telecommunication industry using deep learning. *Information Sciences Letters*, *11*(1), 24.

[4] T2 Chan, Y., Shafiq, S., Naeem, A., Ahmed, S., Safwan, N., & Hussain, S. (2019). Customers churn prediction using artificial neural networks (ANN) in telecom industry. *International journal of advanced computer science and applications*, *10*(9).

[5] T3 Cruz-Cárdenas, J., Zabelina, E., Deyneka, O., Guadalupe-Lanas, J., & Velín-Fárez, M. (2019). Role of demographic factors, attitudes toward technology, and cultural values in the prediction of technology-based consumer behaviors: A study in developing and emerging countries. *Technological Forecasting and Social Change*, *149*, 119768.

[6] T4 Lin, C. P., Huang, Y. H., & Chen, C. T. (2009). Customer relationship management in Taiwan's banking industry. *International Journal of Management and Decision Making*, *10*(1-2), 53-68.

[7] Hassani, H., Huang, X., Silva, E., & Ghodsi, M. (2020). Deep learning and implementations in banking. Annals of Data Science, 7, 433-446.Machine Learning (Super Data Science) [Kiril Eremenko and Hadelin de Pontavis]

[8] Paule-Vianez, J., Gutiérrez-Fernández, M., & Coca-Pérez, J. L. (2020). Prediction of financial distress in the Spanish banking system: An application using artificial neural networks. Applied Economic Analysis, 28(82), 69-87.

[9] Feng, J., & Lu, S. (2019, June). Performance analysis of various activation functions in artificial neural networks. In Journal of physics: conference series (Vol. 1237, No. 2, p. 022030). IOP Publishing.

[10] Manning, C. D., Clark, K., Hewitt, J., Khandelwal, U., & Levy, O. (2020). Emergent linguistic structure in artificial neural networks trained by self-supervision. Proceedings of the National Academy of Sciences, 117(48), 30046-30054.

[11] Kaloev, M., & Krastev, G. (2021, June). Comparative analysis of activation functions used in the hidden layers of deep neural networks. In 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA) (pp. 1-5). IEEE.

[12] Mustapha, A., Mohamed, L., & Ali, K. (2020). An overview of gradient descent algorithm optimization in machine learning: Application in the ophthalmology field. In Smart Applications and Data Analysis: Third International Conference, SADASC 2020, Marrakesh, Morocco, June 25–26, 2020, Proceedings 3 (pp. 349-359). Springer International Publishing.

[13] Alsaawy, Y. A. Z. E. D., Alkhodre, A. H. M. A. D., Benaida, M. O. H. A. M. M. E. D., & Khan, R. A. (2020). A comparative study of multiple regression analysis and back propagation neural network approaches for predicting financial strength of banks: an Indian perspective. WSEAS Transactions on Business and Economics, 17, 627-637.

[14] Nosratabadi, S., Pinter, G., Mosavi, A., & Semperger, S. (2020). Sustainable banking; evaluation of the European business models. Sustainability, 12(6), 2314.

[15] Teles, G., Rodrigues, J. J. P. C., Rabê, R. A., & Kozlov, S. A. (2020). Artificial neural network and Bayesian network models for credit risk prediction. Journal of Artificial Intelligence and Systems, 2(1), 118-132.

DOI: 10.5281/zenodo.10935279