IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

GAN-Driven Synthetic Data Generation for Enhanced Customer Attrition Prediction

[1]Mr. Pritam Ahire, [2]Atharva Wankar, [3]Yashashree Mahajan, [4]Vaibhav Zope

[1]Professor, [2,3,4]Student
[1]Computer Engineering,
[1]Nutan Maharashtra Institute of Engineering and Technology, Pune, India

Abstract : Customer attrition, or churn, is a critical challenge faced by businesses across various industries, leading to significant revenue loss and increased customer acquisition costs. Traditional methods for predicting customer churn often struggle with data imbalance and fail to capture complex customer behavior patterns. Research proposes a novel approach to customer attrition prediction by leveraging Generative Adversarial Networks (GANs) and machine learning (ML) techniques. GANs are employed to generate synthetic data, addressing the issue of imbalanced datasets, while robust ML models such as Random Forests and Artificial Neural Networks (ANNs) are used to predict churn with high accuracy. The system aims to improve data quality, boost prediction accuracy, and offer useful insights for better customer retention strategies. The methodology involves data collection, preprocessing, synthetic data generation using GANs, feature selection, model training, and evaluation. The results demonstrate that the integration of GANs with ML models significantly improves churn prediction accuracy, offering businesses a scalable and adaptable solution to reduce customer attrition. Research contributes to the field by providing a comprehensive framework for predicting customer churn, which can be applied across various industries, including telecommunications, banking, and e-commerce.

***IndexTerms* - Customer Attrition, Churn Prediction, Generative Adversarial Networks (GANs), Machine Learning, Random Forest, Artificial Neural Networks (ANN), Data Augmentation, Feature Selection, Imbalanced Data, Customer Retention, Predictive Analytics, Synthetic Data Generation.**

INTRODUCTION

Customer attrition, often known as customer churn, is a major problem for companies in a number of sectors, such as banking, e-commerce, telecommunications, and subscription-based services [25]. When consumers leave a business, it's known as attrition, and it may result in lost income, a smaller market share, and harm to the company's reputation. For organizations looking to sustain profitability and development, churn prediction is a crucial area of study because getting new customers is sometimes far more expensive than keeping existing ones. Conventional approaches to customer churn prediction include statistical methods and simple machine learning models like support vector machines (SVM), logistic regression, and decision trees [21][27].

Even while these techniques have shown some degree of success, they frequently have trouble with datasets that are unbalanced, meaning that the proportion of consumers who churn is significantly lower than that of those who do not. An imbalance may result in skewed models that are unable to correctly identify clients who are at danger. Additionally, traditional models may not capture the complex, non-linear relationships between customer behavior and churn, resulting in suboptimal prediction accuracy [12].

To tackle these issues, a new method that combines Generative Adversarial Networks (GANs) with sophisticated machine learning models like Artificial Neural Networks (ANNs) and Random Forests. GANs are a type of deep learning model that can generate synthetic data, which can be used to augment the original dataset and address the issue of data imbalance. By combining GAN-generated synthetic data with robust ML models, the proposed system aims to significantly improve the accuracy of churn prediction and provide actionable insights for customer retention strategies.

The primary objectives of research are to:

- Generate synthetic data using GANs to address data imbalance and enhance dataset quality [19].
- Train machine learning models, including Random Forests and ANNs, on the augmented dataset to improve churn prediction accuracy [11][22].
- Use measures like accuracy, precision, recall, and F1 score to assess the models' performance.
- Give organizations insight into the main causes of client attrition so they can create focused strategies to retain customers.

Research contributes to the field of customer churn prediction by offering a scalable and adaptable solution that can be applied across various industries. The integration of GANs with ML models not only improves prediction accuracy but also provides a deeper understanding of customer behavior, helping businesses proactively retain at-risk customers and reduce churn rates [18].

LITERATURE SURVEY

Customer churn prediction has been a widely researched topic in the fields of data science and machine learning. Various studies have explored different methodologies and algorithms to improve the accuracy of churn prediction models. The following section offers a thorough analysis of the existing research on predicting customer attrition, with an emphasis on machine learning and deep learning methods.

2.1Traditional Machine Learning Approaches

Predicting customer attrition is a major problem faced by firms in several kinds of industries. Accurately identifying customers at risk of leaving allows for targeted retention strategies, minimizing revenue loss and maximizing customer lifetime value. Early research in the area often employed traditional statistical methods, such as logistic regression, to model the relationship between customer attributes and churn probability. These methods, while interpretable, often struggle with the complexities and non-linear relationships present in modern customer data.

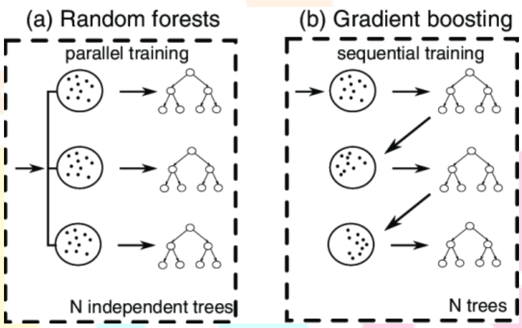
- **Logistic Regression:** One of the most popular techniques for churn prediction is logistic regression. The statistical model uses a linear combination of input characteristics to estimate the likelihood of churn. Despite being simple to understand, logistic regression could not work well on datasets with intricate patterns. [9].
- **Decision Trees:** Decision trees use customer characteristics to describe churn as a sequence of decision-making phases. Although they are good at identifying non-linear correlations, they can overfit, particularly when dealing with big datasets [11].
- **Support Vector Machines (SVM):** For classification tasks, including churn prediction, SVM is an effective method. Finding the hyperplane that best divides churners from non-churners is how it operates [21]. SVM may be computationally demanding, nevertheless, especially when dealing with big datasets that have a lot of features [23].

2.2 Advancements in Machine Learning for Churn Prediction

As machine learning techniques advanced, researchers began exploring more sophisticated algorithms for churn prediction. Methods like Random Forest, Support Vector Machines (SVM), and XGBoost [31] offered improved predictive performance compared to traditional statistical models. Because of its capacity to manage intricate datasets and offer feature relevance rankings, XGBoost in particular became well-known. The rise of ensemble methods like XGBoost also led to increased focus on model explain ability, with techniques like Shapley values being employed to understand the factors driving churn predictions [18].

Random Forest: Several decision trees are constructed using the Random Forest ensemble learning technique, which then aggregates the results to provide the final forecast. It is renowned for handling high-dimensional data and for being resilient to overfitting [16].

Gradient Boosting Machines (GBM): Another ensemble technique is GBM, which creates decision trees repeatedly, fixing the mistakes of the previous tree as it goes. Because of its great accuracy and capacity to manage unbalanced datasets, XGBoost, a well-known GBM implementation, has been extensively utilized for churn prediction [11].



2.3 Deep Learning Approaches

Churn prediction has also been studied in relation to the use of deep learning, most especially Artificial Neural Networks (ANNs) and Convolutional Neural Networks (CNNs) [8]. While these models have shown promise, particularly in domains with rich data like retail, their adoption has been somewhat limited. A key challenge lies in the lack of research on integrating these deep learning architectures with adversarial networks or hybrid models, which could potentially enhance their performance and robustness. Furthermore, comparisons between deep learning and traditional machine learning models often lack a comprehensive evaluation of factors beyond simple accuracy. Future research could explore improving ANN-based models further, such as experimenting with hybrid models or integrating additional deep learning techniques for even higher accuracy [15].

Artificial Neural Networks (ANNs): ANNs are computational simulations that draw inspiration from the composition and operations of the human brain. In order to process incoming data and provide predictions, they are made up of several layers of linked neurons [26]. ANNs are appropriate for churn prediction tasks because of their exceptional ability to capture non-linear interactions [17].

Convolutional Neural Networks (CNNs): Although CNNs are typically used for image data, recent studies have demonstrated their effectiveness in handling tabular data for churn prediction. CNNs can learn spatial patterns within the data, leading to improved prediction accuracy.

2.4 Generative Adversarial Networks (GANs)

Despite these advancements, a significant challenge in churn prediction remains the issue of limited or imbalanced datasets. Real-world customer data often suffers from scarcity, particularly for churned customers, leading to biased models and poor generalization performance. While some studies have explored the use of Multi-Layer Perceptron (MLP) for churn prediction

[13][17], the potential of leveraging advanced techniques like Generative Adversarial Networks (GANs) for synthetic data generation to address system data scarcity and imbalance remains relatively unexplored. Specifically, the application of GANs to generate synthetic customer data for improved training of Random Forest and ANN models for churn prediction represents a significant research gap [19]. In situations when the dataset is unbalanced, GANs have shown to be an effective tool for data augmentation. The generator and discriminator neural networks, which make up a GAN, collaborate to produce artificial data that closely mimics the original dataset [19].

- Generator: By understanding the statistical characteristics of the actual dataset, the generator generates fake data samples.
- Discriminator: The discriminator recognizes the difference between synthetic and actual data and gives the generator input to enhance the quality of the latter.

By generating synthetic data, GANs can address the issue of data imbalance, enabling machine learning models to learn more effectively and improve prediction accuracy [11][19].

2.5 Proposed Research and its Contributions

Research aims to address system gap by investigating the application of GANs to generate synthetic customer data for enhanced training of Random Forest and ANN models. By generating synthetic data that augments the original dataset, we seek to overcome the limitations posed by data scarcity and imbalance, leading to more robust and accurate customer attrition prediction [11][19]. The following approach will not only improve predictive performance but also enable a comparative analysis of Random Forest and ANN models trained on both real and GAN-augmented datasets, providing valuable insights into the effectiveness of synthetic data generation for churn prediction [18].

Research contributes to the field by exploring a novel application of GANs in the context of churn prediction, specifically for enhancing the performance of Random Forest and ANN models, which are widely used in practice. Furthermore, while hybrid neural networks have been explored for churn prediction [10], the integration of GANs for data augmentation in such hybrid architectures is an area requiring further investigation.

PROBLEM DEFINATION

Client attrition, or churn, is a significant concern for organizations across sectors, resulting in revenue loss and decreased client lifetime value. Traditional churn prediction models, such as logistic regression and decision trees, often struggle with imbalanced datasets and fail to capture complex, non-linear relationships in customer behavior. Thus, results in biased models that inaccurately classify at-risk customers. Moreover, existing models lack the ability to generate synthetic data to address data scarcity, limiting their effectiveness in real-world scenarios. To overcome these challenges, there is a need for an advanced churn prediction framework that can enhance model performance, handle data imbalance, and provide actionable insights for customer retention.

PROPOSED METHODOLOGY

The proposed method for customer attrition prediction involves a multi-step process that integrates Generative Adversarial Networks (GANs) with machine learning models to improve prediction accuracy. The approach involves multiple stages like gathering data, preprocessing, synthetic data synthesis, selecting features, training of models, and evaluations.
Module 1 : Data Collection and Pre-processing
Module 2 : Synthetic Data Generation Using GANs
Module 3 : Model Training
Module 4 : Model Evaluation

4.1Data Collection and Pre-processing

The initial stage in the suggested strategy is gathering data. The dataset used in research is obtained from Kaggle and contains 1000 records of customers with 14 attributes, including demographic details, financial metrics, and customer behaviour. To guarantee data consistency and accuracy, the dataset is already processed.
Data Cleaning: Duplicates, missing values, or unrelated features could be present in the raw dataset.
Addressing missing entries, eliminating duplicates, and deleting superfluous columns like "RowNumber" and "CustomerId" are all part of data cleansing.
Feature Engineering: To deepen the analysis, new traits are developed based on preexisting ones. For example, age groups and balance ranges are created to categorize customers, and interaction features such as credit score and age are derived to enhance the model's performance [2][5][7].
Data Transformation: Categorical variables such as 'Geography' and 'Gender' are converted into numerical formats using techniques like one-hot encoding. Continuous variables such as 'CreditScore' and 'EstimatedSalary' are normalized to ensure that all features contribute equally during the training process.

4.2 Data and Sources of Data

To address the issue of data imbalance, a Generative Adversarial Network (GAN) is employed to generate synthetic data. The discriminator and generator in GANs neural networks, which make up GANs, cooperate to produce lifelike data that is synthetic [19].
Generator: The generator learns the distribution of the actual dataset to produce fake data examples. It converts input of unpredictability into data that is similar to the primary dataset.
Discriminator: By differentiating between actual and synthetic data, the discriminator gives the generator suggestions to enhance the level of accuracy of the data generated.
Training Process: an iterative training is used to train the discriminator and generator. The generator originally generates data of low quality, but it progressively improves what it produces after getting input through the discriminator.
Learning keeps up until the discriminator can no more consistently tell the generator's data apart from actual information [19].
The original dataset and the synthetic data produced by the GAN are merged to create an enhanced dataset, which is then utilized to train the models using machine learning.

4.3 Feature Selection

Selection of Features is a crucial phase in the suggested approach is feature selection, which lowers the dataset's complexity and enhances the effectiveness of the model. The most pertinent traits that have a major impact on churn prediction are found via Recursive Feature Elimination (RFE). [21][23].

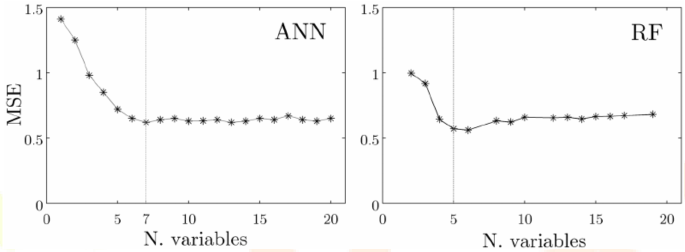
- Recursive Feature Elimination (RFE): RFE is a continuous process that entails training a model, prioritizing characteristics, and eliminating the least significant ones until the ideal feature set is achieved. The process helps reduce model complexity and improve computational efficiency [21][23].
- Feature Analysis: Additional analysis, such as correlation matrices, is performed to understand the relationships between features and identify any multicollinearity. It will ensures that the models are trained on diverse and informative features.

4.4 Model Training

Both machine learning models—Random Forest and Artificial Neural Network—are trained using the enhanced information.

Random Forest: Random an method of ensemble learning called Random Forest creates several decision trees and aggregates their results to get the most accurate forecast. It is selected due to its resilience to overfitting and capacity to manage data with multiple dimensions [11][16].

Artificial Neural Network (ANN): An artificial neural network (ANN) is a deep learning model made up of numerous layers of linked neurons. It is appropriate for churn prediction jobs because it is very good at identifying complicated patterns and unpredictable correlations in customer data [17][22][26].



4.5 Model Evaluation

Numerous measures, including accuracy, precision, recall, F1 score, and AUC-ROC, are used to assess the effectiveness of the models that have been trained [27][28][29]. These metrics ensure that the models not only predict churn accurately but also minimize false positives and false negatives.

- Accuracy: Calculates the percentage of churn versus non-churn cases that were accurately forecasted.
- Precision: calculates the percentage of churn instances that were accurately predicted out of all those that were forecasted.
- F1 Score: A balanced indicator of model performance is the harmonic mean of accuracy and recall.
- AUC-ROC: evaluates how well the model can differentiate between churners versus non-churners.

Based on these assessment measures, the top-performing model is chosen and applied to forecast customer attrition in actual situations.

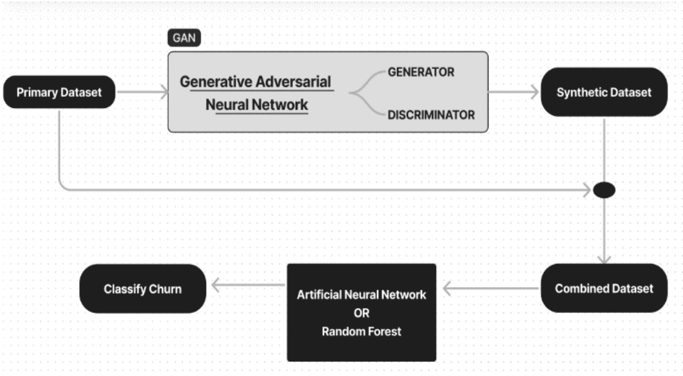
PROPOSED ALGORITHM

The proposed algorithm for customer attrition prediction involves the integration of Generative Adversarial Networks (GANs) with machine learning models such as Random Forest and Artificial Neural Networks (ANN). The algorithm is divided into several steps, including data preprocessing, synthetic data generation, feature selection, model training, and evaluation.

Key Steps:

1. Data Augmentation : A GAN will be trained to generate synthetic customer data to address data imbalance[19].
2. Data Preprocessing : To get the data ready for modeling purposes, feature construction and data cleansing will be done.
3. Feature Selection : Select the most significant features for predicting churn, improving model performance and reducing computational complexity[23].
4. Model Building and Evaluation : Random Forest and Artificial Neural Network models will be trained and evaluated using metrics like accuracy, precision, recall, F1-score, and ROC-AUC[27][28].

By combining GANs and ML, system aims to improve the accuracy and robustness of customer churn prediction models.



ADVANTAGES

1.Better Handling of Imbalanced Data:

- GANs generate synthetic samples to balance the dataset, ensuring the model learns from both churned and non-churned customers effectively.
- ANNs, on the other hand, can struggle with class imbalance, often favouring the majority class unless additional techniques like SMOTE or weighted loss functions are applied [11].

2.Robustness Against Overfitting:

- Random Forest (RF), as an ensemble method, mitigates overfitting by averaging multiple decision trees, providing better generalization [16].
- ANNs are highly flexible but prone to overfitting, especially with small datasets or noisy data, requiring careful regularization (dropout, L2 penalty) [22].

3.Lower Computational Complexity:

Random Forest is computationally efficient compared to deep ANNs, which require extensive training time and hyperparameter tuning. Training a deep ANN, especially with multiple hidden layers, demands more processing power and time compared to training an RF model [11].

EXISTING SYSTEM DISADVATAGES

1. Poor Performance on Imbalanced Data: Traditional models struggle with class imbalance, often failing to accurately predict churned customers due to bias toward the majority class.
2. Limited Ability to Capture Complex Patterns: Simple models like logistic regression or decision trees fail to capture nonlinear relationships in customer behavior [9][11][12], leading to lower prediction accuracy.
3. Lack of Data Augmentation and Generalization: Simple models like logistic regression or decision trees fail to capture nonlinear relationships [9][11][12], in customer behavior, leading to lower prediction accuracy.

APPLICATION

The proposed system has wide-ranging applications across various industries, including telecommunications, banking, e-commerce, and subscription-based services [25][30]. In the telecommunications industry, the system can be used to predict churn based on factors such as network quality, billing rates, and customer service interactions [25][30]. In the banking sector, the system can help identify customers at risk of churning due to better offers from competitors or dissatisfaction with services. E-commerce platforms can use the system to predict churn based on customer purchase behavior, product reviews, and engagement metrics [8]. The system can also be applied to subscription-based services, such as streaming platforms, to predict churn based on usage patterns and customer feedback [8]. Overall, the proposed system provides a comprehensive solution for predicting customer churn and developing targeted retention strategies, helping businesses maintain profitability and growth.

CONCLUSION

The proposed research presents a novel approach to customer attrition prediction by integrating Generative Adversarial Networks (GANs) with machine learning models such as Random Forests and Artificial Neural Networks (ANNs). The methodology addresses the challenges of data imbalance and complex customer behavior patterns, leading to improved prediction accuracy. The quality of the dataset is improved by using GANs to generate synthetic data, which helps machine learning models train more efficiently. By offering useful information for client retention tactics, the suggested approach assists companies in lowering attrition rates and improving their profitability. By providing a flexible and scalable approach that can be used in a variety of sectors, the research advances the area.

FUTURE SCOPE

The future scope of the research includes exploring the application of the proposed methodology to other industries, such as healthcare and retail, where customer churn is a significant concern. Additionally, further research can be conducted to improve the interpretability of the models using explainable AI techniques. It is also possible to investigate the combination of sophisticated deep learning models, such as Long Short-Term is Memory (LSTM) networks and Recurrent Neural Networks (RNNs), to identify trends in how customers behave over time. Lastly, real-time churn prediction may be added to the suggested system, allowing companies to take preventative action to keep at-risk clients.

REFERENCES

[1]Ahire, Pritam Ramesh, and K. Ulaga Priya. "Monitoring Body Mass Index (BMI) Pre & Post Covid-19 Outbreak: A Comprehensive study in Healthcare." 2024 MIT Art, Design and Technology School of Computing International Conference (MITADTSoCiCon). IEEE, 2024.
https://www.semanticscholar.org/paper/A-Framework-for-Healthcare-Everywhere%3A-BMI-Using-on-Tai-Lin/d4c19def868241a0e94471b9a14547e1e1bbc81b

[2]Ahire, Pritam. "Predictive and Descriptive Analysis for Healthcare Data, A Hand book on Intelligent Health Care Analytics-Knowledge Engineering with Big Data" https://www. wiley. com/enus/Handbook+ on+ Intelligent+ Healthcare+ Analytics% 3A+ Knowledge+ Engineering+ with+ Big+ Data-p-9781119792536 Published by Scrivener Publishing." (2021).
https://www.researchgate.net/publication/368277313_Predictive_and_Descriptive_Analysis_for_Healthcare_Data

[3]Ahire, Pritam, et al. "LSTM based stock price prediction." International Journal of Creative Research Thoughts 9.2 (2021): 5118-5122.

[4]Ahire, Pritam R., and Preeti Mulay. "Discover compatibility: Machine learning way." Journal of Theoretical & Applied Information Technology 86.3 (2016).

[5]Ahire, Pritam R., Rohini Hanchate, and Vijayakumar Varadarajan. "Indigenous Knowledge in Smart Agriculture." Advanced Technologies for Smart Agriculture. River Publishers, 2024. 241-258.

[6]Hanchate, R., & Anandan, R. (2023). Medical Image Encryption Using Hybrid Adaptive Elliptic Curve Cryptography and Logistic Map-based DNA Sequence in IoT Environment. IETE Journal of Research, 1–16. https://doi.org/10.1080/03772063.2023.2268578

[7]Ahire, Pritam Ramesh, Rohini Hanchate, and K. Kalaiselvi. "Optimized Data Retrieval and Data Storage for Healthcare Applications." Predictive Data Modelling for Biomedical Data and Imaging. River Publishers 107-126

[8]Vidya Rajasekaran, et al. (2023). Predicting Customer Churn in E-Commerce Using Statistical and Machine Learning Methods . International Journal on Recent and Innovation Trends in Computing and Communication, 11(9), 3968–3973.
https://doi.org/10.17762/ijritcc.v11i9.9738

[9]Sandhya Rani K., Thaslima S., Prasanna N.G.L., Vindhya R., and Srilakshmi P., "Analysis of Customer Churn Prediction in Telecom Industry Using Logistic Regression", International Journal of Innovative Research in Computer Science & Technology (IJIRCST), Vol. 9, Issue 4, July2021
https://doi.org/10.21276/ijircst.2021.9.4.6

[10]Tsaia C.-F. and Lub Y.-H., "Customer Churn Prediction by Hybrid Neural Networks", Elsevier Ltd., 2009
https://cmapspublic3.ihmc.us/rid=1MSXYH555-HXXB67-14YC/Tsai_Lu_Hybrid_Neural_Network_2009.pdf

[11]X. Hu, Y. Yang, L. Chen and S. Zhu, "Research on a Customer Churn Combination Prediction Model Based on Decision Tree and Neural Network," 2020 IEEE 5th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA), Chengdu, China, 2020, pp. 129-132, doi: 10.1109/ICCCBDA49378.2020.9095611.
https://sci-hub.se/https://ieeexplore.ieee.org/document/9095611

[12]D. Chandrakala, "A Survey on Customer Churn Prediction using Machine Learning Techniques", vol. 154, no. 10, pp. 13-16, 2016.
https://www.researchgate.net/profile/Saran-A/publication/310757545_A_Survey_on_Customer_Churn_Prediction_using_Machine_Learning_Techniques/links/5bb5fb8a299bf13e605e2ae9/A-Survey-on-Customer-Churn-Prediction-using-Machine-Learning-Techniques.pdf

[13]Ismail M.R., Awang M.K., Rahman M.N.A., and Makhtar M., "A Multi-Layer Perceptron Approach for Customer Churn Prediction", Journal of Telecommunication, Electronic and Computer Engineering, 2024, 12(3),pp.49-54. http://dx.doi.org/10.14257/ijmue.2015.10.7.22

[14]Xiao, J., Jiang, X., He, C., & Teng, G. (2016). Churn Prediction in Customer Relationship Management via GMDH-Based Multiple Classifiers Ensemble. IEEE Intelligent Systems, 31(2), 37–44. doi:10.1109/mis.2016.16
https://sci-hub.se/https://ieeexplore.ieee.org/document/7412625

[15]Seymen O.F., Olmez E., Dogan O., Er O., and Hiziroglu A., "Customer Churn Prediction Using Ordinary Artificial Neural Network and Convolutional Neural Network Algorithms: A Comparative Performance Assessment", Gazi University Journal of Science, 2023, 36(2), pp. 720-733
https://www.researchgate.net/publication/360529236

[16]V. Umayaparvathi and K. Iyakutti, "Automated Feature Selection and Churn Prediction using Deep Learning Models", International Research Journal of Engineering and Technology (IRJET), vol. 4, no. 3, 2017. https://www.irjet.net/archives/V4/i3/IRJET-V4I3422.pdf

[17] A. Sharma, "A Neural Network based Approach for Predicting Customer Churn in Cellular Network Services," International Journal of Computer Science, 2011
https://arxiv.org/pdf/1309.3945

[18]Maan J. and Maan H., "Customer Churn Prediction Model using Explainable Machine Learning", International Journal of Computer Science Trends and Technology (IJCST), 2023.
https://arxiv.org/pdf/2303.00960

[19] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Nets," ,2014
https://arxiv.org/abs/1406.2661

[20] P. Datta, B. Masand, D. R. Mani and B. Li, "Automated Cellular Modeling and Prediction on a Large Scale," Artificial Intelligence Review, vol. 14 no. 6, (2000), 2018
https://etasr.com/index.php/ETASR/article/view/2108

[21] C. Kang and S. P. Ji, "Customer Churn Prediction Based on SVM-RFE," Paper presented at the International Seminar on Business and Information Management, (2008).
https://www.researchgate.net/publication/224535114_Customer_churn_prediction_based_on_SVM-RFE

[22] S. Haykin, "Neural Network: A Comparehsive Foundation. New Jersey Prentice Hall International, vol. 2, (1999).
Neural Networks - A Comprehensive Foundation - Simon Haykin.pdf

[23] X.-w. Chen and J. C. Jeong, "Enhanced recursive feature elimination," in Sixth International Conference on Machine Learning and Applications (ICMLA 2007), 2007.
https://www.researchgate.net/publication/4321531_Enhanced_recursive_feature_elimination

[24] P. Rothenbuehler, J. Runge, F. Garcin and B. Faltings, "Hidden Markov Models for churn prediction," in 2015 SAI Intelligent Systems Conference (IntelliSys), 2015.
https://www.researchgate.net/publication/308191153_Hidden_Markov_models_for_churn_prediction

[25] A. Amin, F. Al-Obeidat, B. Shah, A. Adnan, J. Loo and S. Anwar, "Customer churn prediction in telecommunication industry using data certainty," Journal of Business Research, vol. 94, pp. 290-301, 2019.
https://ideas.repec.org/a/eee/jbrese/v94y2019icp290-301.html

[26] A. Krogh, "What are artificial neural networks?," Nature Publishing Group, vol. 26, no. 2, pp. 195-197, 2008.
https://www.researchgate.net/publication/5593394_What_are_artificial_neural_networks

[27] D. M. Levine, P. P. Ramsey and R. K. Smidt, Applied statistics for engineers and scientists: using Microsoft Excel and Minitab. Applied Statistics for Engineers

[28]I. Kaur and J. Kaur, "Customer Churn Analysis and Prediction in Banking Industry using Machine Learning," 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC), Waknaghat, India, 2020, pp. 434-437, doi:10.1109/PDGC50313.2020.9315761.
https://sci-hub.se/https://ieeexplore.ieee.org/abstract/document/9315761

[29]K. G. M. Karvana, S. Yazid, A. Syalim and P. Mursanto, "Customer Churn Analysis and Prediction Using Data Mining Models in Banking Industry," 2019 International Workshop on Big Data and Information Security (IWBIS), Bali, Indonesia, 2019, pp. 33-38,doi:10.1109/IWBIS.2019.8935884.
https://sci-hub.se/https://ieeexplore.ieee.org/abstract/document/8935884

[30]Shaaban, Essam & Helmy, Yehia & Khedr, Ayman & Nasr, Mona. (2012). A Proposed Churn Prediction Model. International Journal of Engineering Research and Applications (IJERA. 2. 693-697.
https://www.researchgate.net/publication/236625937_A_Proposed_Churn_Prediction_Model

[31]Xin Hu, Lanhua Chen, "Research on a Customer Churn Combination Prediction Model Based on Decision Tree and Neural Network," 2020 IEEE 5th International Conference on Cloud Computing and Big Data Analytics.
https://sci-hub.se/10.1109/ICCCBDA49378.2020.9095611