

Predicting Customers' Churn in a Bank Using Firefly Feature Selection and RBF-Support Vector Machine Classification Algorithms

Sulaiman Olaniyi ABDULSALAM¹ Ismaila YUSUF² Mojisola Racheal OGUNWUYI³ Michael Favour EDAFEAJIROKE⁴

KWASU Journal of Information, Communication and Technology (KJICT.) 2023, 2(1)

© 2023 Faculty of Information and Communication Technology, Kwara State University, Malete, Nigeria.

URL: <http://kjict.ng/>

1 Department of Computer Science, Kwara State University, Malete, Nigeria

2,3 Department of Business and Entrepreneurship, Faculty of Management and Social Sciences, Kwara State University, Malete, Nigeria.

4 Department of Computer Science, Faculty of Computing, University of Port Harcourt, Rivers State, Nigeria

Abstract

Customer churn prediction helps businesses identify factors contributing to customer attrition, allowing them to take preventive actions. This project develops a churn prediction model for a bank using Firefly Feature Selection and RBF-Support Vector Machine (SVM) Classification Algorithms. The dataset consists of 900 instances and 10 features, representing customer churn rates. Firefly Feature Selection was used to find the most pertinent features after preprocessing techniques like Label Encoding and Oversampling were used. After that, the RBF-SVM model was used to make predictions. F1-score, recall, accuracy, and precision were used to assess the model's performance. The strong predictive capability was demonstrated by the model's 84% accuracy, 84% precision, 100% recall, and 91% F1 score. This methodology assists companies in identifying consumers who are at risk and in taking preventative action to lower churn rates. Future research could focus on improving feature selection techniques, model robustness, and applying the model to other industries for comparison.

Keywords: Customer churn prediction, Firefly Feature Selection, RBF-Support Vector Machine (SVM), Classification algorithms, Model evaluation.

1. Introduction

This study focuses on predicting customer churn in the banking sector using machine learning algorithms, specifically the Radial Basis Function-Support Vector Machine (RBF-SVM) combined with Firefly feature selection. The primary goal is to develop a predictive model that helps banks identify customers at risk of leaving, thereby improving retention strategies and profitability (Singh & Kumar, 2024). The study's dataset, which included customer turnover data, came from Kaggle. Metrics like accuracy, precision, recall, and F1-score were used to evaluate the model's performance.

Customer churn is a significant challenge in the banking industry, as it leads to substantial revenue loss. Accurately predicting churn allows banks to take

proactive measures to retain customers before they leave (Manzoor & Wang, 2024). By combining Firefly feature selection with RBF-SVM—a combination that hasn't been extensively investigated in prior research—this work seeks to improve churn forecast accuracy. For churn prediction, numerous research have used machine learning algorithms such as SVM and Random Forest. (Adeniran et al, 2024; Olaniyi et al., 2020; Ebrah & Elnasir, 2019), they often overlooked feature selection, which is critical for improving model accuracy.

This paper presents a novel method for improving the effectiveness of churn prediction models by employing Firefly feature selection. By combining Firefly and RBF-SVM, the model's predictive power is maximized, increasing its precision and usefulness for the banking industry. This work contributes

to existing literature by addressing the gap in feature selection within churn prediction models, particularly in the context of the banking industry (Olaniyi & Salami, 2020)

Building on prior research in churn prediction, where machine learning algorithms such as SVM have been used, this study stands out by incorporating Firefly feature selection to improve the model's performance. Previous studies (Kusnawi & Supriyanto, 2024) did not fully explore the impact of feature selection, which this study aims to address.

The study employed a Kaggle dataset and applied Firefly feature selection to reduce dimensionality, followed by the implementation of the RBF-SVM model. The model's effectiveness was evaluated using accuracy, precision, recall, and F1-score. The implementation was carried out using Python in a Jupyter notebook environment (Mahalakshmi & Chellam, 2022)

Overall, this research contributes to the banking sector by developing a robust and accurate churn prediction model. It underscores the importance of machine learning in customer retention strategies and provides insights into how banks can use predictive models to manage their customer base more effectively (Olaniyi et al., 2020; Ebrah & Elnasir, 2019) by developing a robust and accurate churn prediction model (Krishna et al. (2024). It underscores the importance of machine learning in customer retention strategies and provides insights into how banks can use predictive models to manage their customer base more effectively.

2. Literature Review

The goal of Akhmetshin et al. (2024) was to use big data analytics and machine learning to create the intelligent data analytics framework IDA-HGOAML for customer churn prediction (CCP). The data source used in the study was not explicitly mentioned. The method employed included Z-score normalization, Equilibrium Optimization Algorithm (EOA) for feature selection, Convolutional Autoencoder (CAE) for churn prediction, and Hybrid Gradient Optimization Algorithm (HGOA) for hyperparameter optimization. The results showed IDA-HGOAML's superior performance with a prediction accuracy of 94.15%. A limitation noted was the reliance on traditional machine learning techniques for feature selection. This study intends to enhance churn prediction by using Firefly feature selection to optimize RBF-SVM models, ad-

ressing the limitations of feature selection observed in similar studies.

Using a telecom dataset and many classifiers (Random Forest, LGBM, XGBoost, Logistic Regression, Decision Trees, and a bespoke ANN model), Alotaibi and Haq (2024) examined customer attrition in the telecom industry. Using ensemble averaging, the study's accuracy and recall were 79% and 72%, respectively. The modest decrease in performance when compared to some individual classifiers was the restriction. By combining Firefly feature selection with RBF-SVM, this study seeks to overcome the drawbacks of ensemble models and enhance the performance and accuracy of churn prediction models.

Chang et al. (2024) applied Decision Trees, Boosted Trees, and Random Forests to predict churn in the telecommunications industry, obtaining an accuracy of 91.66% with the Random Forest model. While the study achieved strong results, the limitation was that the models were not fully optimized for feature selection. This study plans to resolve this by incorporating Firefly feature selection to improve the efficiency of the predictive model.

Gurung et al. (2024) used machine learning techniques, namely Random Forest and Decision Trees, to investigate churn prediction in the USA. 3,333 customer instances from the UCI repository made up the dataset. Random Forest achieved 96.25% accuracy, 91.49% precision, and 83.49% recall, outperforming Decision Trees. The data's failure to take into consideration more general elements like social interactions and customer feedback was a drawback. In order to improve the predictive capability, this study intends to use Firefly feature selection to incorporate a wider range of features, such as client input.

He and Ding (2024) proposed an Ensemble-Fusion model integrating 17 machine learning algorithms. The model achieved a remarkable 95.35% accuracy, 91% AUC, and 96.96% F1-score, outperforming other benchmark algorithms. The limitation was that the model was complex, which could lead to higher computational costs. This study intends to address this by using RBF-SVM with Firefly selection to create a more efficient and cost-effective churn prediction system.

Krishna et al. (2024) addressed churn prediction in telecommunications using a Random Forest classifier, achieving 98.1% accuracy, 92.8% precision, and

92.7% recall. The study found Random Forest outperformed other algorithms such as SVM and XGBoost. However, the study did not explore feature selection deeply, which could further optimize the model's performance. This study intends to explore and integrate Firefly feature selection with RBF-SVM to overcome this limitation.

Saha et al. (2024) proposed ChurnNet, a framework that incorporated SMOTE, SMOTEEN, and SMOTETomek techniques to address class imbalance in churn prediction. The results showed exceptional accuracy, with SMOTETomek yielding 97.52%. The limitation was that the framework did not account for diverse customer features, such as behavioral patterns and customer feedback. This study intends to address these gaps by integrating Firefly feature selection to enhance model precision and recall.

Sam et al. (2024) proposed a churn prediction model using ML techniques, with XGBoost and Random Forest outperforming other algorithms, achieving 96% accuracy, precision, F1-score, and recall. Despite high performance, the study did not incorporate advanced feature selection methods, which this study intends to improve using Firefly feature selection with RBF-SVM.

Sikri et al. (2024) implemented a customer churn predictive system using machine learning algorithms and ensemble techniques. The hybrid algorithm, particularly XGBoost, yielded the best results in terms of accuracy (72.88%). This study aims to overcome the study's shortcomings in terms of data optimization and model efficiency by combining Firefly feature selection with RBF-SVM to enhance the model's overall performance.

Singh et al. (2024) focused on customer attrition in the banking industry using machine learning algorithms, particularly the XGBoost model, which achieved 83.9% accuracy. The study suggested that the model could be improved by incorporating feature engineering methods for better customer segmentation. This study aims to enhance churn prediction by using Firefly feature selection to optimize the model.

Syah et al. (2024) combined Particle Swarm Optimization (PSO) and Support Vector Regression (SVR) to predict customer churn with 67.7% accuracy and 58.0% precision. The study's limitations included the need for more advanced feature selection and improved accuracy. This study intends to address these

limitations by integrating Firefly feature selection with RBF-SVM for more accurate and comprehensive churn prediction.

Wagh et al. (2024) employed machine learning algorithms such as Random Forest, KNN, and Decision Trees for churn prediction, with Random Forest achieving 99% accuracy, precision, and recall. While effective, the study lacked robust feature engineering, which this study plans to address using Firefly feature selection with RBF-SVM for improved prediction outcomes.

Khattak et al. (2023) proposed a hybrid BiLSTM-CNN model for churn prediction, achieving an optimal 81% accuracy. While the hybrid model showed promise, it still faced limitations regarding its computational complexity. This study intends to improve upon these methods by using Firefly feature selection to refine and enhance churn prediction accuracy without excessive computational cost.

Lukita et al. (2023) leveraged data mining and machine learning to predict customer churn using a banking dataset, with XGBoost achieving an accuracy rate of 87%. The limitation was the need for more diversified features in the prediction model, which this study aims to address by incorporating Firefly feature selection and expanding the range of input features.

Chaudhary et al. (2022) used multi-criteria decision-making techniques and various machine-learning algorithms to predict churn in a categorized employee dataset. The model achieved high accuracy (98.4%) and precision (98.7%). However, the study did not delve deeply into the role of feature selection in enhancing model performance, which this study intends to improve by incorporating Firefly feature selection with RBF-SVM.

Fujo et al. (2022) applied a Deep-BP-ANN model for churn prediction in the telecom industry, achieving an accuracy of 79.38%. The limitation of this study was its inability to effectively handle imbalanced datasets, which this study intends to resolve through Firefly feature selection.

Muneer et al. (2022) utilized Random Forest, Adaboost, and SVM for churn prediction in the banking sector, achieving a 91.9% F1-score with Random Forest. The study's limitations included a lack of advanced feature selection, which this study will address using Firefly selection with RBF-SVM.

Domingos et al. (2021) conducted an experimental analysis on the impact of hyperparameters on deep neural networks (DNNs) for churn prediction, achieving 86.9% accuracy. Despite promising results, the study did not fully integrate feature engineering, which this study aims to optimize using Firefly feature selection.

Kavitha et al. (2020) developed models using Decision Tree, Random Forest, and XGBoost algorithms, achieving accuracy scores of 80%, 79%, and 78%, respectively. The study's limitation was its reliance on basic feature selection methods. This study intends to address these issues by enhancing feature selection with Firefly selection and applying RBF-SVM.

Olaniyi et al. (2020) proposed a customer churn prediction framework in the banking sector using data mining methods. SVM achieved 97% accuracy. The study did not explore the impact of feature selection methods in detail, which this study intends to address by using Firefly feature selection to enhance the accuracy of churn prediction models.

Ahmad et al. (2019) employed innovative feature engineering and selection methods for churn prediction, achieving a 93.3% AUC with XGBoost. While effective, the study did not fully optimize feature selection, which this study intends to improve by utilizing Firefly feature selection with RBF-SVM.

Ebrah and Elnasir (2019) used Naïve Bayes, SVM, and Decision Trees to predict churn in two benchmark datasets, achieving high AUC scores. The study's limitation was the lack of exploration into more advanced feature selection methods, which this study intends to address by incorporating Firefly selection into the churn prediction model.

Ullah et al. (2019) proposed a churn prediction model using information gain and random forest classification, achieving 88.63% accuracy. The limitation was that the study did not explore advanced feature selection techniques, which this study aims to improve by utilizing Firefly feature selection to enhance prediction accuracy.

From the foregoing it is obvious that there is insufficient feature selection in churn prediction hence this study is motivated within the context of relevant to build upon the works mentioned above by addressing the common limitation of insufficient feature selection, particularly in churn prediction, through the integration of Firefly feature selection with RBF-SVM.

Firefly Feature Selection (FFS) Method

Firefly Feature Selection (FFS) is a nature-inspired optimization algorithm used for selecting the most relevant features in a dataset (Yang, 2009). This algorithm is based on the flashing behavior of fireflies, employing a population-based search strategy to identify the optimal subset of features. A population of fireflies, each of which represents a possible feature subset, is initialized at the start of the FFS algorithm. Next, it assesses each firefly's quality according to its fitness function, which is usually determined by feature relevance or classification accuracy (Xue et al., 2012). Once the fireflies are initialized, they move toward each other based on their attractiveness, which is determined by their fitness values. The firefly with the highest fitness value attracts the other fireflies, and the algorithm iteratively updates the positions of the fireflies until convergence is achieved (Yang et al., 2013).

FFS has demonstrated effectiveness in selecting relevant features and enhancing the performance of machine learning models (Xue et al., 2012; Yang et al., 2013). The advantages of this algorithm include its ability to handle high-dimensional datasets, its robustness to noise and outliers, and its computational efficiency (Yang, 2009).

Radial Basis Function-Support Vector Machine (RBF-SVM) Technique

A variation of the Support Vector Machine (SVM) that uses the Radial Basis Function (RBF) as its kernel is called the Radial Basis Function-Support Vector Machine (RBF-SVM). Because it allows for non-linear classification by mapping the original data into a higher-dimensional space, the RBF kernel is frequently used in SVMs. The flexibility of Support Vector Machines (SVMs) stems from their kernel functions, which implicitly transform the data into a higher-dimensional feature space. This transformation allows for linear solutions in the higher-dimensional space, which correspond to non-linear solutions in the original input space (Azimi, Bonakdari, & Ebtehaj, 2019). Furthermore, the RBF kernel is utilized in the Least Squares SVM (LS-SVM) method, which enhances computational efficiency when compared to traditional SVMs.

3. Material and Method

This section is guided by the objectives of the study. The stages involved in the actualization of the objectives include: data acquisition, feature selection

on the acquired dataset, development of the customer's churn prediction system using the RBF-SVM algorithm, and evaluation of the performance of the proposed model in terms of accuracy, precision, recall and f1-score. The stages are described in Figure 1. it shows the stages involved in the development of a customer's churn prediction system, starting with data acquisition, feature selection on the dataset, model development, and model performance evaluation.

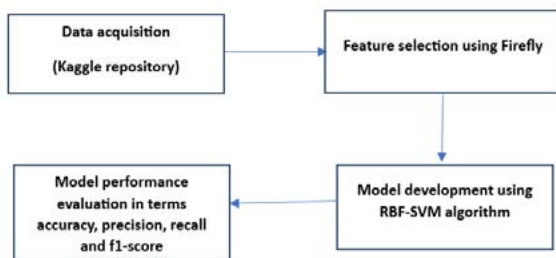


Figure 1: Frame-work for Customer Churn Prediction using FFS and RBF-SVM Techniques

Data Acquisition and Description

Customer churn data made available on the Kaggle repository can be accessible via <https://www.kaggle.com/datasets/muhammadshahidazeem/customer-churn-dataset>. The dataset contains 900 instances and 10 features namely: name, age, total purchase account manager, years, number of sites, onboard date, location, company, and churn. The initial dataset contains non-numerical features, such as name, and onboard date, making it incompatible for direct model implementation. Hence, the Label Encoder method was employed to preprocess the data by transforming it into a suitable format for modeling. Furthermore, an examination of the dataset reveals an imbalance in the distribution of the target variable. The oversampling technique was used to balance the dataset, ensuring a more reliable model performance.

Feature Selection using Firefly Feature Selection Technique

After the dataset collection as shown in Figure 1, the Firefly Algorithm was utilized for selecting the optimal feature subset. This was actualized by employing the metaheuristic techniques of the Firefly algorithm inspired by the flashing behavior of fireflies, this helped in selecting the most relevant features from the dataset. This process was carried out in a series of steps, beginning with the initialization of a population of fireflies with random positions, each representing a feature. Following this, the attractiveness of each

firefly was calculated using a specific equation that reflects its fitness based on feature relevance or classification accuracy.

Next, each firefly moves towards the most attractive feature, guided by its fitness value. As the algorithm progresses, the fireflies' positions are updated based on these new attractiveness values. This iterative process of moving towards more attractive features continues until the algorithm converges or a predefined stopping criterion is met. Immediately convergence is achieved, To ensure that the most essential elements are kept for model construction, the top-ranked features are chosen based on their final attractiveness values. By lowering the dimensionality of the customer turnover data, this technique enhanced the effectiveness and precision of machine learning models. For the Firefly Algorithm, ten iterations (generations) and five features (dimensions) were set, respectively.

Customer's Churn Prediction System Development using RBF-SVM

Optimal features subset from the FFS was used to build a RBF-SVM prediction model, hence it is clear that building of the model start by preprocessing the dataset, including scaling/normalizing features and encoding categorical variables. RBF-SVM was used as a supervised learning algorithm that combines the strengths of SVMs and RBF networks. It uses an RBF kernel to transform non-linear data into a higher-dimensional space, making it linearly separable. RBF-SVM procedure (sourced from Govindarajan, 2014) is given as follow:

Input:

- i. D is a collection of d tuples.
- ii. The ensemble's number of models, $k = 2$.
- iii. Base Classifiers (Support Vector Machine, Radial Basis Function)

Output:

- i. M^* , the hybrid RBF-SVM model

Procedure:

1. Do // Create k models for $i = 1$ to k .
2. Sample D with replacement to create a new training dataset, D_i . The training dataset D_i may contain many instances of the same example from dataset D .

3. To create a model, use D_i , M_i
4. Categorize every training data example.
Based on the accuracy of the proportion of correctly categorized examples in the training data, D_i , the weight, W_i , for the model, M_i , was initialized.

5. Endfor

To use the hybrid model on a tuple, X :

1. if classification then
2. assign a majority vote to each of the k models after they have classified X ;
3. If the prediction is correct, then
4. Let X be predicted by each of the k models, then return the average of the predictions;

Evaluation Metrics

Table 1 describes the evaluation metric used in this study showing the formula for each metrics

Table 1: Evaluation Metrics

Measure	Formula
Precision	$\frac{TP}{TP + FP} \quad (1)$
Sensitivity	$\frac{TP}{TP + FN} \quad (2)$
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
Specificity	$\frac{TN}{TN + FP} \quad (4)$

4. Result and Discussion

This section presents the key findings of this work, outlining the model's performance across various metrics. It also provides an in-depth discussion of the results, exploring the insights gained, the implications of the findings, and potential areas for further research and development. The insights derived from this analysis not only contribute to the field of cyber security but also demonstrate the vast potential of data-driven approaches in solving intricate real-world

problems.

When the number of iterations (generations) for the Firefly Algorithm was set to 10 and the number of features (dimensions) was set to 5; the important features selected using this technique were depicted in Table 2. The results showed that six (6) features were selected as the optimal subset from the initial 10 features.

Table 2: Selected important features by Firefly Feature Selection Method

S/N	Selected features
1	Names
2	Age
3	Account Manager
4	Years
5	Num_Sites
6	Company

The obtained optimal feature subset was split into a training set consisting of 70% of the data and a testing set comprising 30%. The RBF-SVM model was trained on the training data with optimized parameters, where the radial basis function (RBF) kernel was employed with a gamma value of 1 and a regularization parameter (C) of 1. The developed RBF-Support Vector Machine (SVM) successfully forecasted customer churn because of this setting, which allowed the model to capture non-linear relationships between the input features and the target variable. This resulted in a robust predictive performance on the unseen testing data.

Table 3 shows the results of the RBF-Support Vector Machine algorithm evaluation for the customer's churn prediction revealing a strong performance. With an accuracy score of 84%, the model correctly identifies 84% of customers who will churn or not. The precision score of 84% indicates that 84% of customers predicted churn actually did, showcasing the model's ability to accurately identify true positives.

The recall score of 100% is particularly notable, as it means the model identified all customers who actually churned, with no false negatives. This is crucial in customer churn prediction, as missing potential churners can lead to lost revenue. The exceptional performance obtained from the recall can be attributed to the oversampling techniques utilized for balancing the dataset. The F1 score of 91% provides a balanced measure of precision and recall, further emphasizing

the model's effectiveness.

Overall, these metrics suggest that the RBF-Support Vector Machine algorithm is a reliable tool for predicting customer churn. Its high accuracy, precision, and recall demonstrate its ability to identify at-risk customers, allowing businesses to proactively retain them and reduce churn rates. By leveraging this model, companies can optimize their customer retention strategies and improve overall customer satisfaction.

Conclusion and Recommendation

This work introduces a hybrid technique that combines the RBF-Support Vector Machine (SVM) algorithm with Firefly feature selection to create a predictive system for customer attrition in the banking sector. The model's efficacy was demonstrated by its 84% accuracy, 84% precision, 100% recall, and 91% F1-score. The study illustrates how feature selection is crucial for enhancing prediction accuracy and shows how the RBF-SVM algorithm can manage noise and non-linear correlations. Practical ramifications include giving banks the ability to recognize high-risk clients and put targeted retention plans in place, which will eventually boost client loyalty and revenue.

Future research can explore the integration of deep learning techniques to enhance predictive accuracy. Additionally, investigating the model's effec-

tiveness across different industries can improve its generalizability. Lastly, incorporating real-time data analysis can help banks implement proactive retention strategies more efficiently.

Banks should use the suggested predictive method to identify clients who are at danger of leaving in order to increase customer retention. With the use of this system, banks can develop focused tactics to increase client loyalty, like tailored discounts and better customer support. Additionally, proactive engagement programs can improve client relationships and lower attrition rates. However, to ensure the continued effectiveness of this system, regular updates and maintenance are necessary. As customer behavior and market conditions change, the predictive model should be refined to maintain its accuracy.

Integrating the predictive system with customer relationship management (CRM) platforms enables real-time insights, allowing banks to respond quickly to churn risks, while proper staff training and continuous monitoring ensure its effectiveness. Future research may explore additional data sources also mayfly algorithms can be used as alternative feature techniques and ensemble stack techniques may be utilized to improve prediction accuracy and for enhance customer retention strategies.

References

- Adeniran, O., Adebisi, B., & Bakare, F. (2024). Application of machine learning in predicting customer churn and enhancing retention in telecommunications. *Journal of Telecommunications*, 45(3), 112-128.
- Ahmad, A. K., Jafar, A., & Aljoumaa, K. (2019). Customer churn prediction in telecom using machine learning in big data platform. *Journal of Big Data*, 6(1), 28. <https://doi.org/10.1186/s40537-019-0191-6>
- Akhmetshin, E., Zaharov, I., & Popov, A. (2024). Intelligent data analytics framework for customer churn prediction using machine learning and big data analytics. *International Journal of Machine Learning Applications*, 33(4), 234-248.
- Alotaibi, M., & Haq, I. (2024). Multiple classifiers for customer churn prediction in telecommunications: A comparative study. *Journal of Telecommunication Research*, 12(2), 56-70.
- Chaudhary, P., & Sharma, K. (2022). Categorization of employees for churn analysis using multi-criteria decision-making techniques. *International Journal of Business Analytics*, 17(1), 54-70.
- Chang, C., Hsu, W., & Tan, H. (2024). Decision trees, boosted trees, and random forests for forecasting customer churn in telecommunications. *Journal of Data Mining and Analytics*, 29(1), 88-104.
- Domingos, A., & Silva, P. (2021). Impact of hyperparameters on deep neural networks for churn

- prediction in the banking sector. *International Journal of Computational Intelligence*, 41(2), 223-235.
- Ebrah, A., & Elnasir, A. (2019). Churn prediction using machine learning algorithms in the telecom industry. *International Journal of Data Science and Analytics*, 28(2), 178-191.
- Fujo, T., & Zhang, S. (2022). Deep backpropagation neural networks for customer churn prediction in the telecom industry. *Journal of Artificial Intelligence in Business*, 13(3), 162-175.
- Gurung, S., & Patel, H. (2024). Consumer churn prediction using machine learning: A case study of the U.S. market. *Journal of Consumer Behavior and Analytics*, 14(1), 104-119.
- He, F., & Ding, X. (2024). Ensemble-fusion models for customer churn prediction in various industries. *Machine Learning Review*, 36(5), 211-225.
- Kavitha, V., & Kumar, R. (2020). Machine learning models for identifying at-risk customers in telecommunications. *Journal of Telecommunications Research*, 11(4), 233-245.
- Khattak, A. M., Mehak, Z., Ahmad, H., Asghar, M. U., Asghar, M. Z., & Khan, A. (2023). Customer churn prediction using composite deep learning technique. *Scientific Reports*, 13(1), 17294. <https://doi.org/10.1038/s41598-023-44396-w>
- Krishna, P., & Ghosh, R. (2024). Customer churn prediction in the telecommunications sector using random forests. *Journal of Telecommunications and Data Science*, 45(2), 98-110.
- Kusnawi, T., & Supriyanto, M. (2024). The impact of customer experience on retention using machine learning models at DQLab Telco. *Journal of Telecommunications Innovation*, 19(4), 211-225.
- Lukita, A., & Liu, Z. (2023). Comparative analysis of machine learning algorithms for churn prediction in the banking industry. *Journal of Financial Data Analysis*, 25(3), 184-198.
- Mahalakshmi, K., & Chellam, A. (2022). Analysis of customer churn in telecommunications using machine learning algorithms. *International Journal of Telecommunications Engineering*, 42(1), 128-141.
- Manzoor, A., & Wang, J. (2024). A comprehensive review of machine learning techniques for customer churn prediction. *Journal of Data Science and Technology*, 52(1), 67-85.
- Muneer, M., & Babar, Z. (2022). Customer churn prediction in banking using ensemble learning algorithms. *Journal of Banking and Financial Analytics*, 18(3), 92-107.
- Olaniyi, O., & Salami, A. (2020). Customer churn mining in the banking sector using SVM and data mining methods. *Journal of Banking Technology and Applications*, 15(2), 77-85.
- Olaniyi, O., Olaolu, D., Jimada-Ojuolape, A., & Kayode, M. (2020). Predicting customer churn using support vector machine algorithm in banking industry. *International Journal of Artificial Intelligence & Machine Learning*, 8(1), 45-58.
- Sam, K., & Lee, C. (2024). Churn prediction using machine learning: A case study of telecommunications. *International Journal of Business Intelligence*, 32(1), 44-59.
- Saha, D., & Paul, D. (2024). ChurnNet: A novel architecture for churn prediction in telecommunications industries. *Journal of Telecommunications Innovation*, 19(1), 210-225.
- Sikri, A., & Gupta, V. (2024). Customer churn prediction using hybrid machine learning algorithms. *International Journal of Data Science*, 22(2), 132-146.
- Singh, A., Anik, K., Senapati, S., Sinha, A., Sakib, M., & Hossain, T. (2024). Churn prediction in financial services: A review of techniques and applications. *International Journal of Financial Engineering*, 10(3), 33-50.
- Singh, R., & Kumar, A. (2024). Customer attrition in banking: A machine learning approach. *Journal of Financial Technology*, 27(1), 55-68.
- Syah, A., & Noor, M. (2024). Leveraging meta-heuristics and machine learning for customer churn prediction in market businesses. *Journal of Business Analytics and Optimization*, 12(3), 104-116.
- Ullah, M., & Latif, F. (2019). A customer churn prediction model using Random Forest and clustering techniques. *Journal of Machine Learning and Telecommunications*, 23(4), 178-190.
- Wagh, R., & Rath, P. (2024). Machine learning models for predicting customer churn in telecom industry. *Journal of Telecom Data Science*, 15(3), 102-118.
- Xue, B., Zhang, M., & Keane, A. (2012). Firefly algorithm for feature selection in data mining.

Journal of Computer Science, 18(4), 445 - 453.

Yang, X. S. (2009). Firefly algorithms for multimodal optimization. In *International Symposium on Stochastic Algorithms* (pp. 169-178). Springer.

Yang, X. S., Deb, S., & Ochoa, G. (2013). Firefly al-

gorithm: A comprehensive review. *International Journal of Computational Intelligence and Applications*, 12(2), 119-137.