2025 7th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (ICHORA) | 979-8-3315-1088-6/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICHORA65333.2025.11017185

Enhanced Churn Prediction in Telecom with PSO-Based Feature Selection

Merve GÜLLÜ
R&D Department
Türk Telekom
Ankara, Türkiye
merve.gullu@turktelekom.com.tr

Nurefşan AK
R&D Department
Türk Telekom
Ankara, Türkiye
nurefsan.ak@turktelekom.com.tr

Reyhan DEDE
R&D Department
Türk Telekom
Istanbul, Türkiye
reyhan.dede@turktelekom.com.tr

***Abstract*—Customer churn is a significant issue threatening sustainability in the telecommunications sector. This study employs the Particle Swarm Optimization (PSO) algorithm for feature selection to predict customer churn. PSO aims to enhance the performance of machine learning models by identifying the most relevant features in high-dimensional datasets. For this purpose, various data preprocessing techniques, including SMOTE, SMOTENN, undersampling, and oversampling, were applied to the widely used Cell2Cell dataset, and several classification algorithms (Naive Bayes, Logistic Regression, XGBoost, KNN, and Random Forest) were tested. Experimental results demonstrate that feature selection using PSO improves model accuracy and creates simpler, more interpretable models by eliminating redundant features. In particular, a high accuracy rate of 88.8% was achieved when used with the Random Forest algorithm. This study demonstrates that PSO is a powerful feature selection method for customer churn prediction in the telecommunications sector and shows promise for future research.**

***Index Terms*—customer churn, particle swarm optimization, feature selection, telecommunications, SMOTE, SMOTENN.**

I. Introduction

Customer churn has emerged as a significant issue for the sustainability of businesses in today's digital and competitive world. It poses a vital threat to companies, especially in highly competitive fields such as telecommunications [1]. Companies that fail to retain their customers are exposed to revenue losses, brand reputation damage, and market share reduction [2]. Therefore, the ability to accurately and timely predict customer churn holds strategic importance for companies to maintain their profitability [3].

The digital transformation accelerated by 5G technology is reshaping customer expectations. The arrival of 6G is set to raise these expectations even higher with promises of ultra-fast connectivity, low latency, and massive network capacity. The emergence of 6G-enabled services, including virtual and augmented reality applications and autonomous systems, will generate unprecedented volumes of data and bring new complexity to customer behavior. Timely and accurate analysis of this flood of data is critical to identify customers vulnerable to churn and implement proactive retention strategies. Advanced churn prediction models are essential to meet these challenges of the 6G era.

In models developed to predict customer churn in advance, the goal is to identify the right customer groups, enabling companies to develop preventive strategies for customers at high risk of churning [4]. Recent studies have introduced machine learning and deep learning algorithms in churn prediction. In particular, decision trees, support vector machines, artificial neural networks, ensemble methods, and recently advanced deep learning approaches have been employed for this purpose and have drawn attention due to their high accuracy rates [5].

Algorithms success mainly depends on accurately identifying and selecting the features in the dataset [6]. Customer data in the telecommunications sector typically contains thousands of customer records, hundreds of variables, and numerous hidden relationships. Therefore, to build an effective prediction model, it is necessary to select the most meaningful features from these massive datasets [7]. Feature selection increases the accuracy and precision of the model and reduces training time and complexity, generating more interpretable results [6]. However, traditional statistical methods and fixed filtering techniques do not always provide sufficient performance, especially in areas like telecommunications, where data structures are highly complex [8]. This necessitates the exploration of more advanced techniques for feature selection. Hence, there is a need for more flexible and powerful methods, such as heuristic and meta-heuristic algorithms [7].

In recent years, meta-heuristic algorithms like Particle Swarm Optimization (PSO) has emerged as an effective and powerful tool in feature selection and model improvement [9]. PSO is a population-based global optimization algorithm inspired by the collective movements of social beings in nature [10]. The fundamental strength of PSO comes from its ability to navigate effectively in large and complex datasets, dynamically find solutions, and reach the global optimum [11]. Especially in large-scale and multidimensional datasets of the telecommunications sector, PSO's multi-solution search capability helps overcome feature selection challenges that other methods cannot easily handle [12]. This makes PSO a promising approach for feature selection in churn prediction, particularly in increasingly complex telecommunications data.

TThis work is supported by The Scientific and Technological Research Council of Türkiye (TÜBİTAK) 1515 Frontier R&D Laboratories Support Program for Türk Telekom 6G R&D Lab under project number 5249902.

979-8-3315-1088-6/25/$31.00 © 2025 IEEE

Authorized licensed use limited to: ULAKBIM UASL - Pamukkale University. Downloaded on December 04,2025 at 20:07:30 UTC from IEEE Xplore. Restrictions apply.

The literature shows that PSO has been effectively used in churn prediction processes and has achieved significant success. Several studies have demonstrated its efficacy in selecting only the most meaningful features from large datasets, enabling models to reach accuracy levels exceeding 90% [13]. Other approaches have leveraged PSO for feature selection and hyperparameter optimization, leading to improvements in the overall performance of machine learning and deep learning-based churn prediction models [12]. In such integrated approaches, PSO determines which features to use and which parameters the models should operate with, facilitating the generation of more precise results [13].

On the other hand, several studies have demonstrated that hybrid models, where PSO is combined with other meta-heuristic methods such as Simulated Annealing, Genetic Algorithms, and Fuzzy Logic, can create more balanced and high-performing churn prediction models [14]. These hybrid models enhance accuracy and increase processing speed, offering practical solutions for large-scale telecom datasets [7].

The success of models used for churn prediction relies on identifying the right customer groups and developing policies specific to these groups [6]. In this regard, PSO-based feature selection methods contribute to companies' ability to develop more effective and personalized strategies by revealing complex and implicit factors affecting customer churn [15]. Moreover, thanks to PSO's ability to generate multiple solutions, unique feature sets can be identified for different customer segments, allowing more accurate measures to be taken for groups with high churn risk [5].

Compared to other methods in feature selection processes, PSO's key advantage is its ability to transition between global and local optima efficiently, thus detecting hidden features in the dataset that influence churn decisions [9]. Additionally, PSO requires few parameters, enabling the algorithm to produce fast and stable solutions even when working with large datasets [10].

In conclusion, a deeper examination of the contributions of PSO algorithms in churn prediction processes and a comparison with other methods to highlight its advantages emerge as an important necessity. Accordingly, the present study aims to systematically and comprehensively address the role of the PSO algorithm in the feature selection process for churn prediction. The study will also evaluate the performance, flexibility, and advantages of PSO in churn prediction processes in comparison with other methods in the existing literature. Thus, both theoretical and practical contributions of PSO in churn prediction processes will be revealed, and its superiority over traditional methods will be emphasized. In this way, the present work aims to contribute new insights to the academic literature and guide sectoral applications.

II. Material and Method

A. Dataset Used for Churn Prediction

The dataset used in this study is an open-source dataset known in the literature as the Cell2Cell dataset and published on the Kaggle platform under the title "Datasets for Churn - Telecom" [16]. This dataset, which was created to address the problem of predicting customer churn in the telecommunications sector, includes variables such as customer behavior, service usage habits, demographic information, and financial data.

The dataset consists of 71047 customer records and contains 58 attributes for each customer. These variables enable churn prediction models to produce more accurate results by providing a wide range of information about the customer. Factors such as customers' income level, call duration, device characteristics, credit history, interactions with customer service, and responses to marketing campaigns all play a critical role in determining the probability of churn.

B. Data Preprocessing Methods

Data preprocessing steps are critical to achieving better results and improving the effectiveness of machine learning models. There are several problems, such as unbalanced classes, variables with different scales, and distributional differences in the data, with the datasets used to predict customer churn in the telecommunications industry. For this reason, several data preprocessing techniques are used. These aim to improve the modeling process by cleaning and preprocessing data before analysis.

1) Data Balancing Methods:

a) Random Undersampling: Random undersampling is a method of randomly removing samples from the majority class to balance the class distribution in an unbalanced dataset. Although this technique allows the model to learn the minority class more effectively, it can lead to random data loss, so it is important to avoid the loss of important information for more sophisticated strategies such as clustering-based selection [17].

b) Random Oversampling: Random oversampling is a sampling method similar to random undersampling. In this method, samples from the minority class are randomly selected, copied, and added to the dataset to balance the unbalanced class distribution [18].

c) SMOTE: SMOTE is an oversampling method that balances class distributions by creating new instances of minority classes in unbalanced datasets [19]. In contrast to the random oversampling technique, it uses the K-Nearest Neighbour (KNN) algorithm to generate synthetic data points rather than simply copying the existing samples in the minority class. This process better reflects the distribution of minority classes in the dataset, reducing the risk of model over-fitting and clarifying decision boundaries. The SMOTE method is widely used to reduce the effects of class imbalance. It is instrumental in machine learning models such as SVM, decision trees, and neural networks.

d) SMOTENN: It is a hybrid sampling technique created by combining the SMOTE (Synthetic Minority Oversampling Technique) and ENN (Edited Nearest Neighbors) methods to balance the class distribution in unbalanced datasets [19]. In this method, the first synthetic examples of the minority class are generated using SMOTE. Then, the ENN algorithm is applied to remove noisy or erroneous examples from the dataset

Authorized licensed use limited to: ULAKBIM UASL - Pamukkale University. Downloaded on December 04,2025 at 20:07:30 UTC from IEEE Xplore. Restrictions apply.

in both the minority and majority classes. ENN removes misclassified instances or class boundaries and evaluates the class congruence of each data point with its K-nearest neighbors. This process improves the model's generalization ability by reducing the in-class noise that SMOTE can generate due to oversampling. It also allows the model to generate more stable decision boundaries [20].

2) PSO Algorithm for Feature Selection: Particle Swarm Optimization is a population-based heuristic optimization algorithm developed by Kennedy and Eberhart [21]. Inspired by swarm behavior in nature, it uses a population of particles to find the best solution. Compared to traditional optimization methods, it has a low risk of getting stuck in local minima. It provides fast convergence and has continuously improvable parameters. Each particle is a candidate solution representing different combinations of features for PSO in feature selection. By evaluating these candidates with a fitness function, it aims to identify the best subset of features. Its global optimization capability contributes to more accurate models by eliminating redundant features in high-dimensional datasets [22].

a) Working Mechanism of the PSO: PSO allows a group of particles to reach the optimal solution through coordinated movement in solution space, mimicking swarm behavior in nature. The algorithm treats each particle as a candidate solution, which is evaluated according to a fitness function and directed towards better solutions. The basic principle of PSO is that particles act by learning from both their best position and the best position of the swarm. Each particle updates its position concerning its personal best position $pbest$ and the best position in the swarm $gbest$ [23].

Mathematically, the optimization process in a d-dimensional solution space for a population of N particles can be expressed as follows:

- $x_i^t \in \mathbb{R}^d$: Position of the i-th particle at iteration t.
- $v_i^t \in \mathbb{R}^d$: Velocity vector of the i-th particle.
- $pbest_i$: The best position attained so far by the i-th particle.
- $gbest$: The global best position found by the swarm.

The velocity and position of each particle are updated using the following equations:

$$v_i^{t+1} = wv_i^t + c_1 r_1 (pbest_i - x_i^t) + c_2 r_2 (gbest - x_i^t) \quad (1)$$

$$x_i^{t+1} = x_i^t + v_i^{t+1} \quad (2)$$

where:

- w represents the *inertia weight*, controlling the extent to which the particle retains its previous velocity.
- c_1, c_2 are *learning coefficients*, which govern the particle's responsiveness to its own experience and that of the swarm.
- r_1, r_2 are randomly generated values within the range $[0, 1]$, introducing stochasticity to enhance exploration in the solution space.

The velocity update equation consists of three key components:

1) *Inertia component* (wv_i^t): Preserves a portion of the particle's previous velocity, enabling broader exploration of the solution space.
2) *Cognitive component* ($c_1 r_1 (pbest_i - x_i^t)$): Directs the particle towards its own best-known position, leveraging individual experience.
3) *Social component* ($c_2 r_2 (gbest - x_i^t)$): Guides the particle towards the global best solution found by the swarm, promoting collective intelligence.

These steps continue by evaluating the particles by the fitness function and stop when the maximum number of iterations is reached or when the fitness function falls below a certain threshold.

b) Fitness Function: The performance of Particle Swarm Optimization (PSO) is highly dependent on the fitness function, which evaluates the quality of different solutions within the search space. The fitness function serves as a criterion to assess how effectively a given solution satisfies the requirements of a specific problem. Typically, the optimization objective is either to minimize (*minimization*) or maximize (*maximization*) the function's value to achieve the best possible outcome.

In the context of feature selection, the primary goal of PSO is to identify the optimal subset of features by effectively optimizing the fitness function. A well-designed fitness function can significantly enhance the convergence speed of PSO, facilitating its ability to reach the global optimum efficiently.

c) Optimization Process: The optimization process in Particle Swarm Optimization (PSO) is an iterative procedure in which particles are evaluated based on a predefined fitness function and continuously updated to find the optimal solution. The process begins with a randomly initialized set of particles, where each particle is assessed according to the fitness function. Particles adjust their velocities and positions by learning from both their own best-known positions ($pbest$) and the best solution found by the swarm ($gbest$). This learning mechanism guides the particles toward the optimal solution.

At each iteration, the velocity of each particle is updated using the inertia weight (w), cognitive coefficient (c_1), and social coefficient (c_2). The updated velocity values determine the new positions of the particles. The optimization process continues until either a predefined maximum number of iterations is reached or no further improvement is observed in the fitness function. This iterative approach enables PSO to converge toward the global optimum while mitigating the risk of getting trapped in local optima, making it a crucial aspect of the algorithm's effectiveness.

3) Machine Learning Models:

This study tested various machine learning algorithms and their performances compared. Each model possesses distinct advantages and limitations; thus, multiple algorithms were explored to identify the most suitable method for the

Authorized licensed use limited to: ULAKBIM UASL - Pamukkale University. Downloaded on December 04,2025 at 20:07:30 UTC from IEEE Xplore. Restrictions apply.

given dataset. Below, the fundamental characteristics of each model and the rationale for their selection in this study are elucidated.

Naïve Bayes is a probabilistic classifier based on Bayes' theorem, operating under the feature independence assumption. This model is renowned for its computational efficiency and ability to provide stable predictions even with limited data, making it particularly advantageous for datasets with categorical variables.

Logistic Regression is a statistically robust model known for its high interpretability. It quantifies the influence of independent variables on the dependent variable, facilitating a straightforward analysis of variable importance. Additionally, logistic regression performs effectively on linearly separable datasets by establishing linear decision boundaries.

XGBoost is a highly efficient and widely utilized machine learning algorithm, particularly noted for its ability to handle missing data and incorporate regularization mechanisms to prevent overfitting. Its capacity to learn complex interactions among variables enables it to achieve high accuracy on large datasets.

k-Nearest Neighbors (KNN) is a straightforward yet effective classification model that assigns class labels based on the classes of the nearest neighbors. When the dataset is well-structured and exhibits clear patterns, KNN can serve as a robust classifier.

Random Forest is a robust ensemble learning method aggregating multiple decision trees to produce more stable and reliable predictions. This model captures complex relationships among variables and demonstrates resilience against overfitting.

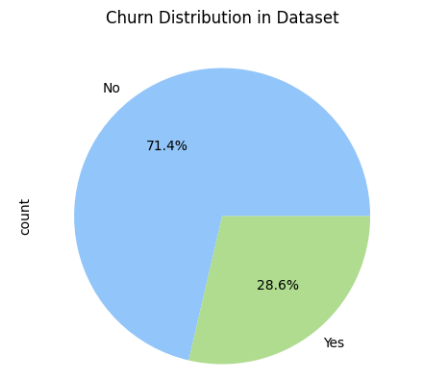

Fig. 1. Churn Distribution in Dataset

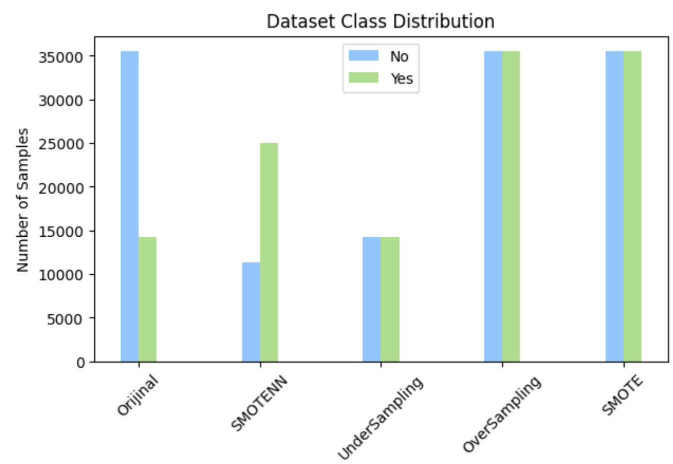

Fig. 2. Churn Distribution in Dataset

III. Methodology and Experimental Results

This study aims to predict churn and investigate the reasons for churn. The processes in line with these objectives are addressed in three stages. These stages include data preprocessing, feature selection, classification and making sense of the outcome.

A. Data Preprocessing

The raw dataset contains a total of 71,047 records and a total of 58 features, with churn information for each record. As a result of removing duplicate and nan values in the dataset, 49,752 samples were obtained. The high proportion of nan values was removed because they were included in the class information. As the dataset distribution is shown in Figure 1, the number of instances with no churn is high.

Data balancing procedures were performed on the dataset. These are random undersampling, random oversampling, smote and smoteen techniques. The four techniques and the original dataset are shown in Figure 2 with the number of samples for the class distributions.

B. Feature Selection

In this study, PSO is applied to determine the best feature subset. The basic mechanism of PSO is that each particle corresponds to a specific feature subset, and the fitness function evaluates its performance. The model performance of the dataset (original dataset without balanced class distribution) that passed the data preprocessing stage was monitored (see Table III). Since the highest success was obtained with the Random Forest algorithm, Random Forest was chosen as the fitness function used for PSO. Other parameters used for PSO and their values are presented in Table I. The model was trained using the feature subset selected by the PSO algorithm, and the model's performance was evaluated using the accuracy metric. In the optimization process of PSO, the subset of features with the best fitness value was determined.

The feature selection process was repeated five times with the PSO algorithm. Optimization processes with different initial conditions and randomness were applied in each run. The results were analyzed for each iteration, and the average success measures are reported in Table II. After five iterations, the average success rate was 71.87%, and the minimum success rate was 71.49%. The minimum success was obtained using only 26 out of 57 features, excluding the number of classes in the original dataset, with a slight difference of 0.003. The 25 features of the model with the highest success obtained

Authorized licensed use limited to: ULAKBIM UASL - Pamukkale University. Downloaded on December 04,2025 at 20:07:30 UTC from IEEE Xplore. Restrictions apply.

TABLE I
PSO ALGORITHM PARAMETERS

Params	Values
Population Size	10
Number of Iterations	30
Fitness Function	Random Forest Model Prediction
Fitness	(1-accuracy)
Normalization Range	[0,1]
PSO Inertia Weight (w)	Adaptive
PSO Cognitive Coefficient (c1)	2
PSO Social Coefficient (c2)	2
Number of experiment repetitions	5

in five iterations (Run1 stage) were selected as the dataset to solve the problem.

TABLE II
PSO AND RANDOM FOREST ALGORITHM RESULTS IN FIVE TEST RUNS

Runs	Accuracy Score	Number of Feature
Run1	0.7248	25
Run2	0.7208	35
Run3	0.7174	32
Run4	0.7160	26
Run5	0.7149	26

The obtained results are analyzed to evaluate the effectiveness of PSO in the feature selection process, and the features that provide the best performance in customer churn prediction are identified.

C. Classification

Classes are often unbalanced in real-world churn prediction problems, and the proportion of churning may be low compared to non-churning. This imbalance can make machine learning models insensitive to imbalanced classes and produce poor predictions for the low-frequency class. Therefore, in this study, different data balancing techniques such as Random Oversampling, Random Undersampling, SMOTE, and SMOTENN are applied, and their performances are compared.

Before (dataset with 57 features) and after (dataset with 25 features), the feature selection process and classification were performed using classical machine learning and ensemble learning methods for customer churn prediction. Among the methods, Naïve Bayes, Logistic Regression, XGBoost, KNN, and Random Forest algorithms were used. Testing was carried out on the balanced dataset versions using different methods. All results are presented in Table III.

When the results are analyzed, the most successful classification result is obtained after over sampling and feature selection with 88.8%. This success value shows that more accurate results can be obtained with fewer features in churn prediction. In models such as Logistic Regression and Random Forest, the effect of feature selection was limited. In general, feature selection did not cause a negative effect in all models, similar accuracy rates were maintained in the worst case scenario, and significant improvements were observed in some models.

TABLE III
ACCURACY VALUES OF MODELS BEFORE AND AFTER FEATURE SELECTION

Dataset	Model	BFS[a] Acc.	AFS[b] Acc.
original	Logistic Regression	0.705	0.705
	Random Forest	0.718	0.724
	XGBoost	0.709	0.716
	KNN	0.664	0.671
	Naive Bayes	0.637	0.652
under sampling	Logistic Regression	0.578	0.578
	Random Forest	0.608	0.612
	XGBoost	0.602	0.598
	KNN	0.536	0.540
	Naive Bayes	0.552	0.558
over sampling	Logistic Regression	0.577	0.577
	Random Forest	0.881	0.888
	XGBoost	0.718	0.707
	KNN	0.640	0.637
	Naive Bayes	0.545	0.552
smote	Logistic Regression	0.685	0.660
	Random Forest	0.794	0.789
	XGBoost	0.792	0.789
	KNN	0.686	0.671
	Naive Bayes	0.594	0.584
smotenn	Logistic Regression	0.755	0.738
	Random Forest	0.834	0.829
	XGBoost	0.834	0.846
	KNN	0.748	0.744
	Naive Bayes	0.689	0.675

[a]BFS means that Before Feature Selection.
[b]AFS means that After Feature Selection.

XGBoost, KNN and Naive Bayes showed the most improvement after feature selection. This may indicate that PSO is more effective especially on complex or distance-based models. In our study, we observed that SMOTENN gives the best results especially on complex models (such as XGBoost and Random Forest) and improves the overall model performance.

IV. CONCLUSION AND ASSESSMENT

Churn forecasting is crucial for customer relationship management in telecommunications, banking, and subscription-based services. An accurate and effective forecasting model can help companies increase customer loyalty, optimize marketing strategies, and prevent revenue losses. With the development of 6G technology, big data analytics, and AI-based forecasting models are becoming more critical in the telecommunications industry. The ultra-high speed, low latency, and broad connectivity offered by 6G networks will make customer behavior more dynamic and data flow more complex. In this context, scalable and optimizable churn prediction models become a significant requirement.

In this study, feature selection for the churn prediction problem is performed with the PSO algorithm, and classification processes are performed using different machine learning and deep learning methods. The performance of the models was improved, and the best feature subset was determined as a result of the optimization with PSO.

Experimental results show that feature selection by PSO is efficacious in improving model accuracy and creates more

Authorized licensed use limited to: ULAKBIM UASL - Pamukkale University. Downloaded on December 04,2025 at 20:07:30 UTC from IEEE Xplore. Restrictions apply.

straightforward and more understandable models by eliminating redundant features. After five replicated experiments, it is observed that the selected features contribute to the model performance. The optimization process using Random Forest based fitness function improved the model's accuracy and reduced the computational cost.

The results show that the PSO algorithm is a successful feature selection method for churn prediction and can produce more efficient results than traditional all-feature models. In future studies, the advantages and disadvantages of PSO can be evaluated in more detail by comparing it with different metaheuristic optimization algorithms. In addition, its performance on more complex datasets can be examined by using it together with deep learning models.

REFERENCES

[1] N. Mustafa, L. S. Ling, and S. F. A. Razak, "Customer churn prediction for telecommunication industry: A malaysian case study," *F1000Research*, vol. 10, p. 1274, 2021.

[2] S. K. Wagh, A. A. Andhale, K. S. Wagh, J. R. Pansare, S. P. Ambadekar, and S. Gawande, "Customer churn prediction in telecom sector using machine learning techniques," *Results in Control and Optimization*, vol. 14, p. 100342, 2024.

[3] A. Amin, A. Adnan, and S. Anwar, "An adaptive learning approach for customer churn prediction in the telecommunication industry using evolutionary computation and naïve bayes," *Applied Soft Computing*, vol. 137, p. 110103, 2023.

[4] S. Dhariya, "Customer churn prediction in telecommunication industry using machine learning and deep learning approach," in *2023 3rd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*. IEEE, 2023, pp. 804–810.

[5] R. Liu, S. Ali, S. F. Bilal, Z. Sakhawat, A. Imran, A. Almuhaimeed, A. Alzahrani, and G. Sun, "An intelligent hybrid scheme for customer churn prediction integrating clustering and classification algorithms," *Applied Sciences*, vol. 12, no. 18, p. 9355, 2022.

[6] M. Imani and H. R. Arabnia, "Hyperparameter optimization and combined data sampling techniques in machine learning for customer churn prediction: a comparative analysis," *Technologies*, vol. 11, no. 6, p. 167, 2023.

[7] C. Praseeda and B. Shivakumar, "Fuzzy particle swarm optimization (fpso) based feature selection and hybrid kernel distance based possibilistic fuzzy local information c-means (hkd-pflicm) clustering for churn prediction in telecom industry," *SN Applied Sciences*, vol. 3, pp. 1–18, 2021.

[8] H. Jain, A. Khunteta, and S. Srivastava, "Telecom churn prediction and used techniques, datasets and performance measures: a review," *Telecommunication Systems*, vol. 76, pp. 613–630, 2021.

[9] J. Nayak, H. Swapnarekha, B. Naik, G. Dhiman, and S. Vimal, "25 years of particle swarm optimization: Flourishing voyage of two decades," *Archives of Computational Methods in Engineering*, vol. 30, no. 3, pp. 1663–1725, 2023.

[10] T. M. Shami, A. A. El-Saleh, M. Alswaitti, Q. Al-Tashi, M. A. Summakieh, and S. Mirjalili, "Particle swarm optimization: A comprehensive survey," *Ieee Access*, vol. 10, pp. 10 031–10 061, 2022.

[11] A. G. Gad, "Particle swarm optimization algorithm and its applications: a systematic review," *Archives of computational methods in engineering*, vol. 29, no. 5, pp. 2531–2561, 2022.

[12] E. H. Houssein, A. G. Gad, K. Hussain, and P. N. Suganthan, "Major advances in particle swarm optimization: theory, analysis, and application," *Swarm and Evolutionary Computation*, vol. 63, p. 100868, 2021.

[13] M. Sedighimanesh, A. Sedighimanesh, and M. Gheisari, "Optimizing hyperparameters for customer churn prediction with pso-enhanced composite deep learning techniques," *Preprints*, p. 2024031048, 2024.

[14] M. Hao, "Research on customer churn prediction based on pso-sa feature selection algorithm," in *2024 5th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT)*. IEEE, 2024, pp. 1055–1059.

[15] M. Z. Alotaibi and M. A. Haq, "Customer churn prediction for telecommunication companies using machine learning and ensemble methods," *Engineering, Technology & Applied Science Research*, vol. 14, no. 3, pp. 14 572–14 578, 2024.

[16] Jpacse, "Datasets for churn - telecom," 2021, erişim tarihi: 17 Mart 2025. [Online]. Available: https://www.kaggle.com/datasets/jpacse/datasets-for-churn-telecom

[17] J. Brownlee, "Random oversampling and undersampling for imbalanced classification," *Machine learning mastery*, vol. 14, 2020.

[18] R. Van den Goorbergh, M. van Smeden, D. Timmerman, and B. Van Calster, "The harm of class imbalance corrections for risk prediction models: illustration and simulation using logistic regression," *Journal of the American Medical Informatics Association*, vol. 29, no. 9, pp. 1525–1534, 2022.

[19] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.

[20] G. E. Batista, R. C. Prati, and M. C. Monard, "A study of the behavior of several methods for balancing machine learning training data," *ACM SIGKDD explorations newsletter*, vol. 6, no. 1, pp. 20–29, 2004.

[21] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95-international conference on neural networks*, vol. 4. ieee, 1995, pp. 1942–1948.

[22] B. Xue, M. Zhang, W. N. Browne, and X. Yao, "A survey on evolutionary computation approaches to feature selection," *IEEE Transactions on evolutionary computation*, vol. 20, no. 4, pp. 606–626, 2015.

[23] M. Y. Özsağlam and M. Çunkaş, "Optimizasyon problemlerinin çözümü için parçaçık sürü optimizasyonu algoritması," *Politeknik Dergisi*, vol. 11, no. 4, pp. 299–305, 2008.

Authorized licensed use limited to: ULAKBIM UASL - Pamukkale University. Downloaded on December 04,2025 at 20:07:30 UTC from IEEE Xplore. Restrictions apply.