Journal of Engineering Managment and Soft Computing (JEMSC)

Preventing Customer Churn in E-commerce Based on Feature Selection Using Metaheuristic Methods and Deep Learning

Vahideh Rahmani khorasani[1], Mohammad Reza Yamaghani[2]✉ and Mohammad Hossein Darvish Motevalli[3]

1. Assistance Prof. Department of Information Technology Management, IK.C., Islamic Azad University, kish, Iran. Email: vahide.rahmani@iau.ac.ir
2. Corresponding Author. Assistance Prof. Department of Computer Engineering and Information Technology, La.C., Islamic Azad University, Lahijan, Iran. Email: O_yamaghani@iau.ac.ir
3. Assistance Prof. Department of Industrial Management, WT.C., Islamic Azad University, Tehran, Iran. Email: Mhd.darvish@iau.ac.ir

Article Info

Article type:
Research Article

Article history:
Received 29 October 2025
Received in revised form
29 April 2026
Accepted 17 June 2026
Published online 1 July 2026

Keywords:
customer churn, e-commerce, feature selection, deep learning.

ABSTRACT

Given the extensive changes in today's business landscape, the continuous development of technology, and the need to pay attention to customers and respond to their demands in order to preserve organizational survival, preventing customer churn is necessary and unavoidable. Therefore, retaining existing customers, attracting new customers, and controlling customer churn for business growth and development is essential. The aim of this research is to present an algorithm to prevent customer churn in e-commerce based on feature selection using metaheuristic methods and deep learning. This research, in terms of its objective, results, and algorithm development, is based on operational research methods. The research variables are composite and the study is cross-sectional in time. The results indicate that the proposed algorithm is capable of finding the best approach for predicting customer churn. In this research, first, the features of customer churn are selected using genetic optimization algorithm based on binary human learning, and then, selected features as input of a deep learning model are used to predict customer churn in e-commerce. The research findings indicate that the optimization and deep learning algorithm, with an accuracy of 92.32, performed better than other customer churn prediction methods and can, as an efficient algorithm, assist organizations in improving performance.

Cite this article: Rahmani khorasani, V. & et al, (2026)., Preventing Customer Churn in E-commerce Based on Feature Selection Using Metaheuristic Methods and Deep Learning. *Journal of Engineering Management and Soft Computing,* 12 (3). 62-82.

DOI: https://doi.org/10.22091/jemsc.2026.14256.1316

DOI: https://doi.org/10.22091/jemsc.2026.14256.1316

Publisher: University of Qom

1) Introduction

In today's rapidly changing business world, where customer demands change almost as quickly as trends themselves, companies face an urgent need to place the customer at the core of their operations. It is no longer sufficient to merely react; businesses must actively align their value chains to start with customer needs and achieve genuine, tangible satisfaction (Guerola-Navarro et al., 2024; Oltra-Badenes et al., 2019).

These technologies are no longer just support systems. They have become strategic enablers. They harness the power of data analytics to help businesses predict what their customers might do next and allow them to tailor their services accordingly (Fakheri et al., 2024; Hosseini et al., 2024). This predictive capability gives customer relationship management a central role in digital transformation. This technology not only improves existing workflows but also creates space for the emergence of new ideas and the revelation of deeper insights.

At the same time, the rise of e-commerce tells its own compelling story. In 2018, global online sales reached the staggering figure of $2.86 trillion. The implication is clear: Understanding how people behave in the online space is not optional, it is essential. However, at this urgency, it is not easy to decode the online consumer behavior (Fakheri et al., 2024; Hosseini & Yaghmani, 2024). In the meantime, some variables are involved. Researchers try to understand all this by using a combination of statistical tools and machine learning models, but it is rarely simple. Different data sources, inconsistent methods, and constantly changing customer habits make it a continual challenge (Cirqueira et al., 2020; Turatti, 2025). Finally, the goal is to understand all those data—predict what customers will likely buy in the next phase and discover gaps where further research can have a real impact. It is a complicated puzzle, but it also has a value to solve (Guerola-Navarro et al., 2020).

Predicting customer churn in e-commerce is of vital importance, as it allows businesses to identify customers who are on the verge of leaving their services or products and retain them through actions such as offering discounts special offers, or improving the customer experience (Jamali & Navaei, 2016; Navaei & Afzali, 2014). In today's competitive market, the cost of acquiring new customers is far higher than retaining current ones. This approach not only helps reduce costs and strengthen loyalty but also ensures long-term profitability and sustainability (Khare & Arora, 2024).

From a technical perspective, churn prediction is a binary classification problem (churn or no churn) that requires careful selection of data features to build efficient models and understand the factors influencing customer behavior (Matuszelański & Kopczewska, 2022; Rad et al., 2023).

E-commerce data are typically large and diverse, and this complexity can make analysis difficult and time-consuming. To address this challenge dimensionality reduction and feature selection methods are employed to increase processing efficiency by removing unnecessary data (Babaee Khobdeh et al., 2023; Sam et al., 2024). Feature selection enables data simplification by preserving key information and removing redundant parameters, while feature extraction transforms data into a lower dimensional space that retains the core information. Although these methods can improve the accuracy and clarity of analysis, they may lead to the loss of some information if features are removed excessively or if designs are overly complex. Nevertheless, the intelligent use of these techniques can optimize customer behavior analysis and help prevent their churn (Sadr et al., 2025; Saranya & Pravin, 2021).

In this paper, a feature selection approach based on metaheuristic optimization algorithms is used to predict customer churn in e-commerce. The proposed method will use customer churn datasets from the standard Kaggle repository (Blastchar, 2018). The present data include information on customer behavior in e-commerce, where some of these customers have churned and others remain stable. Given that the dataset has many features, feature selection using novel metaheuristic optimization algorithms is used to select a quasi-optimal subset of features. The quasi optimal feature subset will be used as input to deep neural networks to learn customers' behavioral features for churn prediction.

The main innovation of this research is to provide a hybrid framework for predicting customer churn in e - commerce where a subset of the effective features is selected from the customer's behavioral data, and these selected features are then used as input of deep learning models. This approach, with focusing on optimal selection of features and taking advantage of learning ability of deep models, allows for more accurate identification of factors affecting the failure and improvement of prediction performance.

The main contributions of the paper are stated as follows.

- Extracting useful features from churn prediction-related data based on metaheuristic optimization algorithms
- Determining important factors in customer churn in e-commerce
- Applying deep learning to important factors
- Predicting new customers churn using patterns extracted from deep learning
- Comparing the accuracy of the proposed method based on different metaheuristic optimization algorithms and various classification methods

The remainder of this paper is organized as follows. In the second section, related work will be reviewed. In the third section, the details of the proposed method will be presented. In the fourth section, the implementation results of the proposed method will be presented. In the fifth section, the discussion and conclusion of the paper will be stated.

2) Literature Review

Given the importance of predicting customer churn in e-commerce, various studies have been conducted in this area; in this section we review some of these recent studies.

In Heshmati and Haq's (2023) study, using graph convolutional networks, a novel method for customer churn prediction was presented. This method models customer relationships as a graph structure, using dynamic network analysis over time and using each customer's features as graph nodes to improve churn prediction. In Sam Deli's (2023) research, the CRISP method was used to predict churn in electronic banking services. The aim of the present study is to identify the characteristics of churners from e-banking services. In Ghanbari Marvast et al. (2023), ensemble learning algorithms, which are advanced methods in data mining, were used to predict customer churn. First, the dataset under investigation was balanced using random oversampling; then, data preprocessing was performed, and finally, customer churn was predicted using voting and adaptive boosting ensemble learning algorithms.

In Subramanian et al. (2024), new customer churn prediction is developed using deep learning . In Gurung et al. (2024), the application of machine learning algorithms is discussed, such as random forests and decision trees to design forecasting models of churn and exploration in key factors that change probability. In He and Ding (2024), fusion model is based on machine learning and an intelligent system is introduced to help reduce the actual churn in the product. In Wagh et al.'s (2024) research, the classifier uses classification techniques to find customers churn and collect the reasons behind the absence of customers leave in the telecommunication industry. The main objective of this system is to analyze various machine learning algorithms required to develop customer churn prediction models and identify reasons of churn in order to provide strategies and their maintenance plans. In Abbasimehr and Bahrini (2024), a new feature selection method suggests that the imperialist competitive algorithm along with three methods of selecting the filter feature is used to select the best feature set.

Usman-Hamza et al. (2024) investigated the effectiveness of tree-based classifiers with various computational features in predicting customer churn. Specifically, the customer churn prediction performance of various tree-based classifiers, such as single, ensemble, boosted, and hybrid tree-based classifiers, is examined. In Murugesan and Trivedi's (2024) study, a Remora optimization algorithm proposed three support vector machine-based strategies for customer churn prediction.

Amin et al. (2023) proposed an adaptive learning approach to address the complex problem of customer churn prediction using a Bayesian classifier and a genetic algorithm-based feature weighting method (a subset of evolutionary algorithms). Durkaya Kurtcan and Ozcan (2023) presented a churn prediction model that incorporates feature selection and optimization into classification. This study performs principal component analysis for selecting the best features, support vector machine for predicting customer churn, and grey wolf optimization for optimizing the parameters of the support vector machine. In Quek et al. (2023), a customer churn prediction model based on feature selection analysis and support vector machine is proposed. The proposed model reduces the feature dimensionality while improving churn prediction performance by identifying the most important features of customer data. Liu et al. (2023) presented preprocessing of multidimensional data, feature extraction, and dataset processing provided by a telecom operator. Then, the k means algorithm is used to cluster different consumer groups, which in turn analyzes the factors of interest for different consumer groups and offers targeted suggestions.

In Khattak et al. (2023), deep neural networks were used to extract features without considering sequence information. Given these issues, the present study offers an effective method for churn prediction based on a hybrid deep learning model called bidirectional long short-term memory–convolutional neural networks. Suh (2023) conducted an analysis of customer behavior information at a real water purifier rental company, where customer churn often occurs, and a water churn prediction model was developed and validated. Jiang et al. (2023) proposed a model-based earnings forecast model that considers both efficiency and cost. Feature selection is modified based on a multi-objective nuclear orbital search and an extreme learning machine is used to obtain suitable variables for the prediction of failure with maximum efficiency and minimum cost. In Shir Ali Shahreza and Mahmoudzadeh (2026), prediction of customer churn in banking industry using behavioral characteristics of banking transactions and machine learning algorithms, ensemble learning methods, and deep learning models have been investigated on real data of a large bank. In Ghaedi Khani and Shahpasandi (2026), predicting customer churn by performing a multi-stage process involving data preprocessing, the feature extraction by the genetic algorithm, and the classification based on the combination of the basic classifiers is studied using genetic algorithm. In Delafrooz et al. (2026), presenting a model for reducing the churn in the insurance industry by using customer relationship management based on artificial intelligence and using the mixed research design including thematic analysis in the qualitative section and partial least squares modeling in the quantitative section have been studied. Namini (2025) evaluated customer loyalty detection using data mining methods and multiple machine learning algorithms, including support vector machine, multilayer neural network, decision tree, and logistic regression on real data of customers and a model based on neural network composition and teacher—student optimization algorithm is presented to improve performance. Jafari Dehkordi et al. (2025) investigated the prediction of customer churn in the field of e-commerce using customer behavioral data and an integrated forecasting model based on artificial intelligence methods. In Sarabadani and Cheragh (2024), identifying and predicting the failure of bank customers by using demographic and behavioral characteristics of customers and using a hybrid model based on random forest algorithm and jaya optimization method, has been studied. Golshan and Rahimi Khoshmakani (2025) identified the effective patterns in customer churn in online stores, using descriptive data mining approach and analyzing a set of customer data to discover the relationships and patterns of behavior. Ahmadi-pour (2025) identified customer churn patterns in the telecommunication industry using data mining in the context of customer relationship management and customer behavior analysis.

Previous studies indicate that although various methods based on machine learning, deep learning, and feature selection have been proposed to predict customer churn, challenges such as the selection of features in high-dimensional data, effective combination of feature selection methods, forecasting models, and improving the accuracy and consistency of models in the field of e-commerce has not been fully resolved. Therefore, the present study aims at selecting a feature based on meta-heuristic methods and taking advantage of deep learning to eliminate part of this research gap.

3) Methodology

As mentioned, the present research uses a hybrid feature selection approach based on a metaheuristic optimization algorithm to detect customer churn early. In this research, a subset of near-optimal features is selected using binary human learning optimization. Using this approach, important and useful features for customer churn are identified, and then used for deep learning and extracting customer churn patterns in e-commerce.

3.1 Feature Selection Based on Metaheuristic Optimization Algorithm

As noted, metaheuristic optimization algorithms are computational algorithms, inspired by optimization processes in nature. In the proposed method, metaheuristic optimization algorithms are used to select important features in the customer churn dataset. Considering the complexity of this problem and its nondeterministic polynomial time order, this problem is treated as an optimization problem. In this problem, two general objectives are considered in the context of feature selection. The first objective is to reduce the number of features present in the dataset, and the second objective is to increase the accuracy of customer churn prediction. Metaheuristic optimization algorithms attempt to reduce the number of features and increase churn prediction accuracy by using the important features present in the dataset. These algorithms can increase prediction accuracy using suitable features and improve system performance.

Metaheuristic search algorithms are used in a binary manner for feature selection, and the inputs to these algorithms are the features used in the original datasets. The initial population in these algorithms is chosen randomly, and by optimizing the initial population, the optimal solution is selected. In these algorithms, each solution is considered as a vector whose length equals the number of features in the dataset. The values within each element can be binary and accept only the values 1 and 0. The presence of 1 in an element indicates that the feature corresponding to that element's index in the original dataset is selected in the chosen solution. Likewise, the presence of 0 in an element indicates that the corresponding feature in the original dataset is not selected in the solution. Therefore, the initial population will be a matrix where each row of this matrix represents a solution and each column represents a feature. The matrix elements can only have values of 0 and 1. Table 1 shows an example of the initial population matrix.

Table 1. An Example of the Initial Population in the Proposed Method

Initial Population	Feature1 Index	Feature2 Index	...	Featuren Index
Solution1	$F_{1,1}$	$F_{1,2}$	...	$\boldsymbol{F_{1,n}}$
Solution2	$F_{2,1}$	$F_{2,2}$	...	$\boldsymbol{F_{2,n}}$
...	...	...	...	...
Solutionm	$F_{m,1}$	$F_{m,2}$	...	$\boldsymbol{F_{m,n}}$

According to Table 1, in this algorithm, a matrix of the initial population is used which includes m rows and n columns. The number of rows m is chosen by the parameters of the Human Learning Optimization algorithm as the number of solutions in the initial population or the number of individuals in the initial population. Additionally, the number of columns n is determined according to the number of features in the original dataset. Each of these rows represents a solution that has selected some of the features available in the original data set. This matrix is used as the input to the Human Learning Optimization algorithm.

In the selection phase of metaheuristic optimization algorithms, one or more optimal solutions are chosen from the initial population and the other solutions in the exploration phase tend toward them. These steps are repeated until the best solution is selected as the best subset of features. In general, metaheuristic search algorithms for binary feature selection, by using population-based optimization and selecting the best solution, attempt to choose the best subset of features. These algorithms can be effective in selecting important features and reducing the dimensionality of datasets, and can be used to improve system performance and prediction accuracy.

3.2 Proposed Fitness Function

The fitness function is one of the main stages in evolutionary algorithms, particularly in metaheuristic algorithms. In these algorithms, for each solution in the initial population, a fitness function value is calculated, which indicates the performance and quality of the solution. This fitness function value is calculated based on selected features from the dataset; additionally, by considering the prediction error of customer churn, the target value for the solution is calculated. In fact, this fitness function indicates the desirability and quality of the solutions. Solutions with a higher fitness function value are considered better and more suitable solutions and are selected in the subsequent stages of the evolutionary algorithm. Finally, solutions with the best fitness function value are selected as the final and optimal solution.

Next, each solution is evaluated based on the fitness function and defined objectives. The solutions are evaluated using the fitness function, which is a combined proposal of feature selection rate and customer churn prediction error, as defined in Relation 1.

$$Minimize\ F(x) = \begin{cases} f_1(x) = \frac{L}{A} & , L \in A, A \in \mathbb{R}^+ \\ f_2(x) = 1 - \frac{FP+FN}{P+N} & , (P+N) \in \mathbb{R}^+ \end{cases} \quad (1)$$

where A is the total number of features, L is the number of selected features, TP represents the total number of healthy individuals correctly identified as healthy, FP stands for the total number of healthy individuals incorrectly identified as sick, TN indicates the total number of sick individuals correctly identified as sick, FN is the total number of sick individuals incorrectly identified as healthy, P represents the sums of TP and TN, and finally N is the sums of FP and FN.

3.3 Deep Neural Network

Deep neural networks are composed of three types of layers. These are convolutional layers, pooling layers, and fully connected layers. When these layers are stacked, a deep neural network architecture is formed. A simplified deep neural network architecture for classification is illustrated in Figure 1.

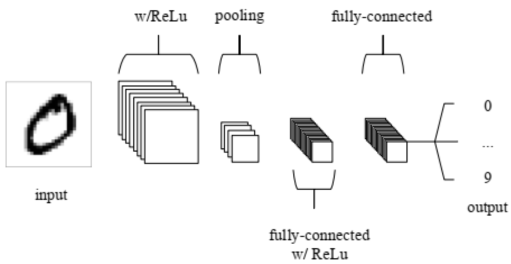

Figure 1. Simple Deep Neural Network Architecture Consisting of Five Layers (O'Shea & Nash, 2015)

According to Figure 1, the layers of deep neural networks are presented as follows.

- **Input Layer**

The input layer in neural networks plays an important role in processing and optimizing image data. This layer first centers the data by reducing the data mean to zero, creating favorable conditions for network training. Then, it normalizes the data to the range [0,1] to prevent large fluctuations and increase the convergence speed of learning. Additionally, by using data whitening techniques, unnecessary dependencies between features are reduced, and by analyzing principal components, data dimensions are reduced, while focus on key factors is increased. These steps ensure that data is available to learning algorithms in a compact and focused manner, improving the neural network's learning performance (Du, 2018).

- **Convolutional Layer**

The convolutional layer in deep neural networks, particularly for image processing, plays a key role in extracting local features. This layer uses convolutional kernels (filters) that slide over the inputs to extract specific features from each part of the image and produce convolutional results. The weights of the convolutional kernel are applied uniformly across all points of the image, a technique known as "weight sharing," which allows the network to identify similar features at different points in the image. Adjusting parameters such as kernel size, depth, stride, and other filter settings helps the convolutional layer effectively recognize patterns and perform better in data analysis. The algorithm for calculating the output size is defined by relation 2 (Wu, 2017):

$$H_{out} = 1 + \frac{H_{in} + (2*pad) - K_{height}}{S};\ W_{out} = 1 + \frac{W_{in} + (2*pad) - K_{width}}{S} \quad (2)$$

- **Activation Layer**

Non-linearization of the output of the convolutional layer in neural networks is performed through activation layers, which are essential for solving the vanishing gradient problem and improving model learning. Various activation functions, including sigmoid, hyperbolic tangent, rectified linear unit, leaky rectified linear unit, and others, are used for this purpose. In particular, the leaky rectified linear unit has become a popular choice in recent years due to its fast convergence speed and the absence of a dead region. The appropriate choice of activation function can significantly impact the performance and training speed of deep neural networks (Lalwani et al, 2022).

- **Pooling Layer**

The pooling layer in deep neural networks is used to reduce the dimensionality of extracted features and decrease computational complexity, helping to prevent overfitting. This layer includes three main types: global pooling, which involves max and average pooling with a stride equal to the step size; overlapping pooling, which preserves more information through overlapping regions; and spatial pyramid pooling, which allows the network to translate features from input images of different sizes into a uniform dimension, without damaging the information structure and preventing data loss. Spatial pyramid pooling has become a key component in the design of deep neural networks due to its ability to preserve information and improve efficiency when dealing with diverse images.

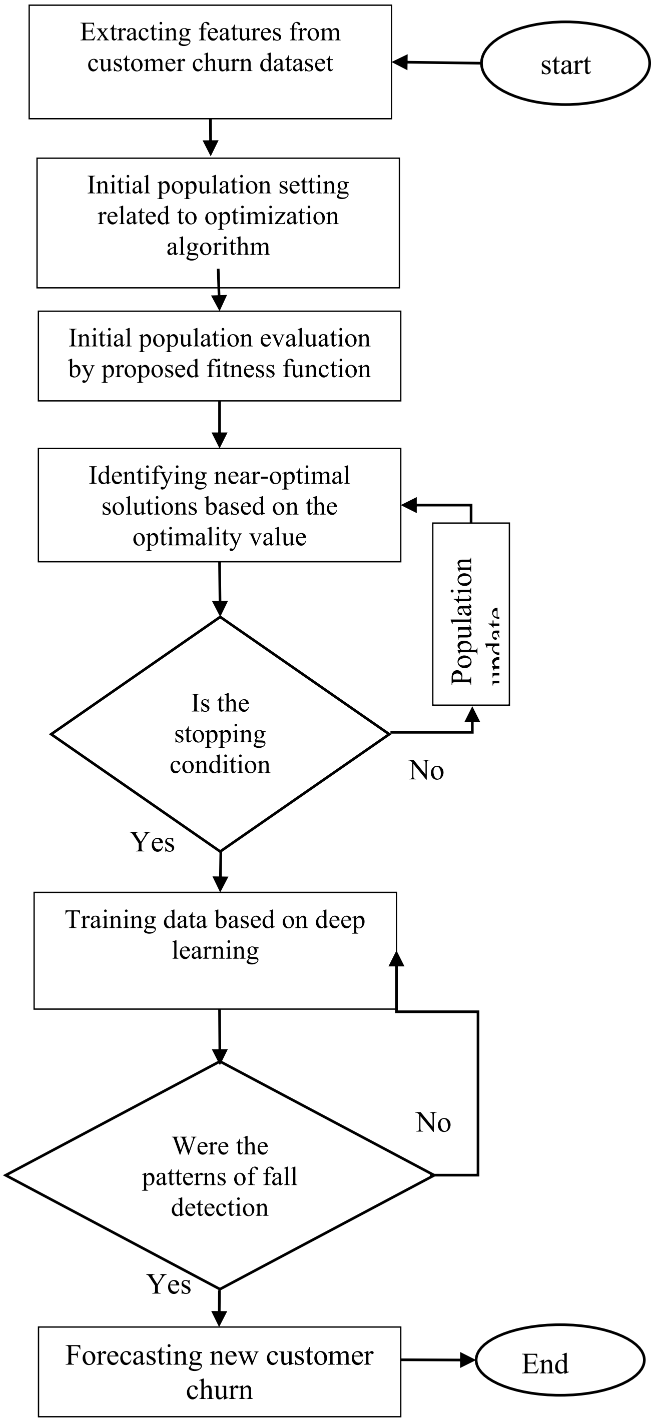

Figure 2. Flowchart of the Proposed Method

- **Fully connected layer**

Fully connected layers at the end of deep neural networks are used to transfer and analyze the information extracted from previous layers to produce the final output. In these layers, each neuron is connected to all neurons in the previous layer, which helps utilize all extracted features simultaneously and produces more accurate outputs. After several convolutional layers that extract different features from the image, the fully connected layers analyze this information comprehensively and help create a stronger representation of the data. These layers help simplify computations and increase processing speed; therefore, they are important in the final design and performance of deep neural networks (Lokesh & Vasantha, 2024).

The flowchart of the proposed method is depicted in Figure 2.

3.4 Evaluation of the Proposed Method

To evaluate the proposed method for analyzing new transactions and identifying new customer churn, various measurement metrics are used. These metrics are obtained based on the interplay between the predicted class label for customers and their actual class label, which is specified in the confusion matrix.

The main evaluation criteria that can be used to evaluate two-class classification methods are:

1. Accuracy: The ratio of the number of customers correctly classified to the total number of customers.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{TP+TN+FP+FN} \tag{3}$$

2. Sensitivity or Recall: The ratio of the number of positive customers correctly classified to the total number of actual positive customers.

$$Recall = \frac{\text{TP}}{TP+FN} \tag{4}$$

3. Specificity: The ratio of the number of negative customers correctly classified to the total number of actual negative customers.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{5}$$

4. F-measure: The harmonic mean of precision and recall, used as a combined metric of precision and the informational content of the model.

$$F - measure = \frac{2*Precision*Recall}{Precision+Recall} \tag{6}$$

These metrics help us evaluate the performance of the classification method on the test data and measure its accuracy and efficiency. The confusion matrix and the parameters true positive, false positive, true negative, and false negative can be used to calculate these metrics and to better understand the performance of the classification method.

4) Findings and Discussion

As mentioned, the proposed method attempts to predict customer churn based on feature selection using a metaheuristic method and deep learning. The proposed method uses the customer churn dataset from the Kaggle standard data repository. This dataset contains 20 categorical and numerical features and also a column for the class label of the samples.

This class label indicates whether customers have churned or not according to their characteristics. Table 2 shows the characteristics of this dataset.

Given the characteristics present in the dataset, it can be understood that some of these features may be irrelevant. Instead of increasing accuracy in customer churn detection, they can increase the required processing volume and reduce classification and prediction accuracy. Consequently, it is necessary to propose a model that can select important and relevant features for customer churn detection.

Table 2. Features of the Customer Churn Dataset

Feature Number	Feature Title	Feature Number	Feature Title
1	customerID	11	**OnlineBackup**
2	gender	12	**DeviceProtection**
3	SeniorCitizen	13	**TechSupport**
4	Partner	14	**StreamingTV**
5	Dependents	15	**StreamingMovies**
6	tenure	16	**Contract**
7	PhoneService	17	**PaperlessBilling**
8	MultipleLines	18	**PaymentMethod**
9	InternetService	19	**MonthlyCharges**
10	OnlineSecurity	20	**TotalCharges**

4.1 Selected Features

As stated, the proposed method uses the Human Learning Optimization (HLO) algorithm to find optimal features, which employs classification error to evaluate solutions. By examining various solutions, this algorithm finds an optimal combination of features for classifying and predicting customer churn. Given the available features, it provides the most optimal combination of features for classification and churn prediction.

Table 3 presents the best solution selected by the Human Learning Optimization algorithm.

According to Table 3, the HLO algorithm, by selecting 8 core features to use in classifying training samples and predicting test samples, improves performance in detecting customer churn. The evaluation criterion of this algorithm is the classification error, calculated by the fitness function, and the classification error is used as the output of the fitness function to evaluate the algorithm's performance. The convergence plot in the figure shows how the fitness function value and the classification error of customer churn decrease using the solutions of the HLO algorithm.

Table 3. Subset of Features Selected by the Human Learning Optimization Algorithm

Feature Number	Feature Title
4	Partner
5	Dependents
7	PhoneService
8	MultipleLines
11	OnlineBackup
12	DeviceProtection
14	StreamingTV
15	StreamingMovies

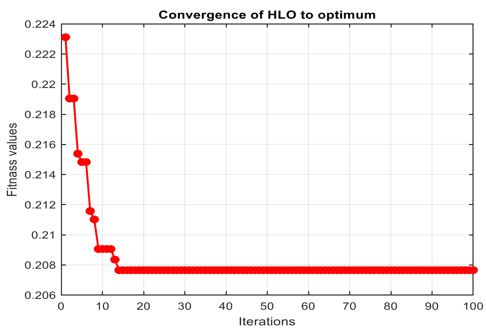

Figure 3. Convergence of the Fitness Function to the Optimum Point in HLO

According to Figure 3, the HLO algorithm in the feature subset selection problem moves toward the optimal point, which is zero percent error, as the number of iterations increases. This indicates that the optimization algorithm, as it continues the improvement process, reaches a state very close to the desired optimum. For example, after 100 iterations, the algorithm reached a fitness function value of 0.2076, indicating improved performance and high accuracy of the algorithm. These results show that the Human Learning Optimization algorithm performs well in finding the optimal solution for the feature subset selection problem and minimizes the classification error as much as possible.

4.2 Evaluation of the Proposed Method

As mentioned, in the proposed method, the predicted labels for new customers are compared with the actual customer labels to calculate the evaluation metrics. Well-known evaluation metrics used for assessing classification methods include accuracy, sensitivity, specificity, and the F-measure. These metrics are used as standard measures and allow us to comparatively evaluate the performance of different classification methods. In this paper, in addition to the feature selection combination based on the Human Learning Optimization algorithm with convolutional neural networks, other classification methods, such as K-nearest neighbors, random forest, decision tree, ensemble classification, and gradient boosted ensemble classification, were used. In Tables 4, 5, 6, and 7, the mean values of the accuracy sensitivity, specificity, and F-measure metrics for the classification methods over runs of the proposed method implementation are represented, respectively.

Table 4. Mean 20 Accuracy Metric for Classification Methods

	Accuracy					
	CNN	KNN	RF	DT	BAG	XGB
1	0.86	0.65	0.84	0.75	0.85	0.83
2	0.89	0.75	0.85	0.75	0.86	0.84
3	0.90	0.76	0.85	0.76	0.86	0.84
4	0.90	0.77	0.85	0.76	0.87	0.84
5	0.90	0.77	0.86	0.76	0.87	0.85
6	0.91	0.77	0.86	0.77	0.87	0.86
7	0.91	0.77	0.87	0.78	0.88	0.86
8	0.92	0.79	0.87	0.79	0.88	0.86
9	0.92	0.79	0.87	0.79	0.88	0.87
10	0.92	0.79	0.88	0.80	0.88	0.87
11	0.92	0.79	0.88	0.80	0.89	0.87
12	0.92	0.79	0.88	0.80	0.89	0.87
13	0.92	0.79	0.89	0.80	0.90	0.88
14	0.92	0.80	0.89	0.81	0.90	0.88
15	0.93	0.80	0.89	0.81	0.90	0.89
16	0.93	0.81	0.89	0.82	0.91	0.89
17	0.94	0.81	0.90	0.82	0.91	0.90
18	0.94	0.82	0.90	0.83	0.91	0.90
19	0.94	0.82	0.90	0.83	0.92	0.92
20	0.98	0.82	0.91	0.84	0.92	0.93

Table 5. Mean Sensitivity Metric for Classification Methods

Recall						
	CNN	KNN	RF	DT	BAG	XGB
1	0.86	0.78	0.85	0.85	0.86	0.86
2	0.87	0.79	0.86	0.86	0.86	0.86
3	0.88	0.84	0.87	0.86	0.87	0.87
4	0.89	0.84	0.87	0.87	0.87	0.87
5	0.89	0.85	0.88	0.87	0.88	0.87
6	0.90	0.85	0.88	0.88	0.88	0.87
7	0.90	0.85	0.88	0.88	0.88	0.87
8	0.90	0.86	0.88	0.88	0.88	0.87
9	0.91	0.87	0.89	0.88	0.89	0.88
10	0.91	0.87	0.89	0.88	0.89	0.88
11	0.91	0.87	0.89	0.88	0.89	0.88
12	0.92	0.88	0.90	0.88	0.89	0.88
13	0.92	0.88	0.90	0.89	0.89	0.89
14	0.92	0.88	0.90	0.90	0.90	0.89
15	0.92	0.88	0.90	0.91	0.90	0.90
16	0.93	0.88	0.90	0.91	0.90	0.90
17	0.93	0.89	0.91	0.91	0.91	0.90
18	0.94	0.90	0.91	0.92	0.92	0.90
19	0.95	0.90	0.92	0.92	0.92	0.92
20	0.95	0.90	0.92	0.93	0.93	0.92

Table 6. Mean Accuracy Metric for Classification Methods

Precision						
	CNN	KNN	RF	DT	BAG	XGB
1	0.90	0.79	0.89	0.85	0.89	0.87
2	0.91	0.82	0.90	0.86	0.92	0.89
3	0.92	0.82	0.91	0.86	0.92	0.89
4	0.92	0.83	0.92	0.87	0.92	0.90
5	0.93	0.83	0.92	0.87	0.93	0.90
6	0.93	0.83	0.93	0.87	0.93	0.90
7	0.94	0.83	0.93	0.88	0.93	0.91
8	0.94	0.83	0.93	0.88	0.93	0.91
9	0.95	0.84	0.93	0.89	0.93	0.91
10	0.95	0.84	0.94	0.89	0.93	0.91
11	0.95	0.85	0.94	0.89	0.93	0.91
12	0.95	0.85	0.94	0.89	0.94	0.91
13	0.95	0.85	0.94	0.90	0.94	0.91
14	0.95	0.86	0.94	0.90	0.94	0.92
15	0.95	0.86	0.94	0.90	0.94	0.93
16	0.95	0.87	0.95	0.91	0.94	0.93
17	0.96	0.87	0.95	0.91	0.95	0.93
18	0.97	0.87	0.96	0.91	0.96	0.94
19	0.99	0.88	0.97	0.92	0.96	0.94
20	0.99	0.98	0.97	0.94	0.97	0.96

Table 7. The Mean of F Criterion for Classification Accuracy

F-Measure						
	CNN	KNN	RF	DT	BAG	XGB
1	0.93	0.78	0.88	0.86	0.88	0.88
2	0.93	0.83	0.89	0.87	0.89	0.88
3	0.93	0.83	0.90	0.87	0.89	0.88
4	0.95	0.84	0.90	0.87	0.90	0.88
5	0.92	0.85	0.90	0.87	0.90	0.89
6	0.94	0.85	0.90	0.87	0.90	0.89
7	0.93	0.85	0.91	0.87	0.91	0.89
8	0.95	0.86	0.91	0.89	0.91	0.89
9	0.92	0.86	0.91	0.89	0.91	0.89
10	0.91	0.86	0.91	0.89	0.91	0.89
11	0.95	0.86	0.91	0.89	0.91	0.90
12	0.94	0.86	0.91	0.89	0.92	0.90
13	0.92	0.86	0.91	0.90	0.92	0.91
14	0.92	0.87	0.92	0.90	0.92	0.91
15	0.91	0.87	0.92	0.90	0.92	0.91
16	0.90	0.87	0.92	0.91	0.92	0.91
17	0.93	0.87	0.92	0.91	0.93	0.91
18	0.90	0.87	0.93	0.91	0.93	0.92
19	0.92	0.88	0.93	0.91	0.94	0.92
20	0.92	0.89	0.94	0.91	0.94	0.93

Considering Tables 4 to 7, it appears that the proposed method based on combining feature selection with the Human Learning Optimization algorithm achieves desirable results for evaluation metrics when running experiments on test data and new customers. The accuracy metric in evaluating classification methods means the rate of customers correctly identified in their respective category, i.e., churn or non-churn. This metric reveals the model's ability to recognize a new sample based on learning from previous samples. Figure 4 shows the chart of average accuracy of the classification methods across execution stages.

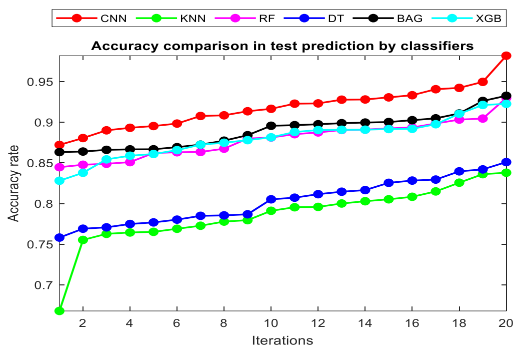

Figure 4. Comparison of Accuracy of Classification Methods

The sensitivity metric is actually an evaluation measure of a diagnostic system, indicating how well the customers who have churned are correctly identified by the diagnostic system. Figure 5 indicates the chart of average sensitivity of the classification methods across execution stages.

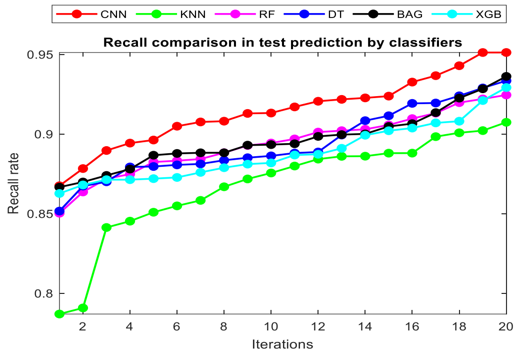

Figure 5. Comparison of Average Sensitivity of Classification Methods

The precision metric indicates the rate of churn detection among the test data, and the amount of churn samples that were not detectable is considered as precision error. Figure 6 shows the chart of average precision of the classification methods across execution stages.

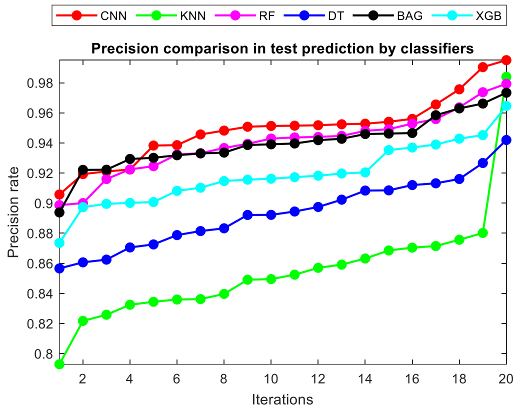

Figure 6. Comparison of Precision of Classification Methods

Finally, the F-measure is a comprehensive metric that takes the harmonic mean of precision and sensitivity. Figure 7 shows the chart of average F-measure of the classification methods across execution stages.

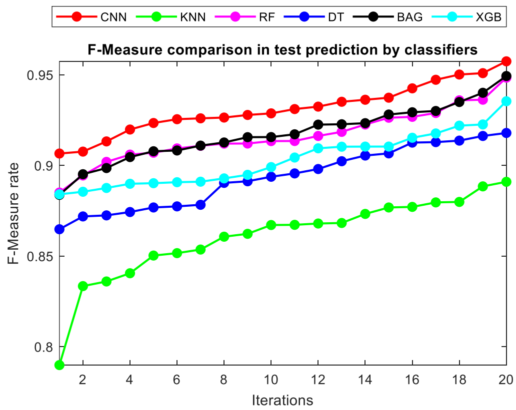

Figure 7. Comparison of F-Measure of Classification Methods

Considering the results in Tables 4 to 7 and Figures 4 to 7, it can be seen that the features selected in the proposed method based on the Human Learning Optimization algorithm and their combination with convolutional neural networks have high capability in terms of evaluation metrics for predicting churn of new customers. Additionally, to clearly compare the superiority of the proposed method based on combining the Human Learning Optimization algorithm and convolutional neural networks, boxplots of the overall mean values of the evaluation metrics can be used. In Figures 8, 9, 10, and 11, respectively, the boxplots related to the overall mean values of the accuracy, sensitivity, precision, and F-measure are shown.

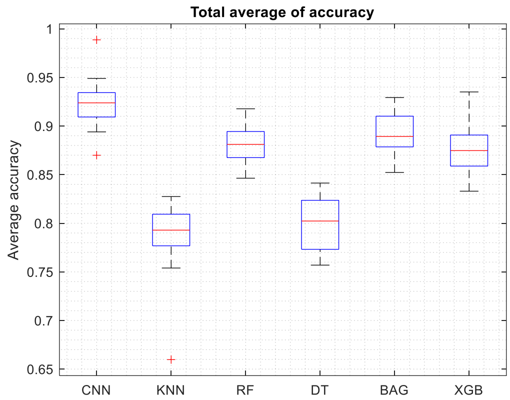

Figure 8. Boxplot of Overall Mean Accuracy

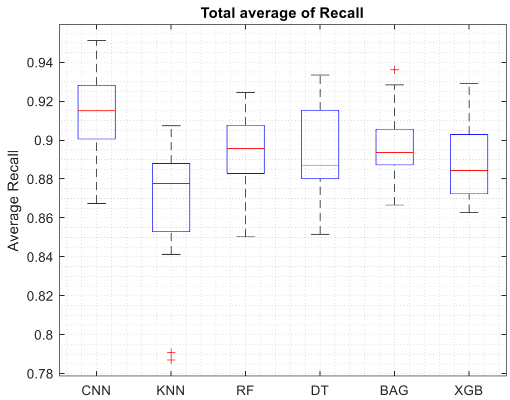

Figure 9. Box Plot of the Overall Average Sensitivity

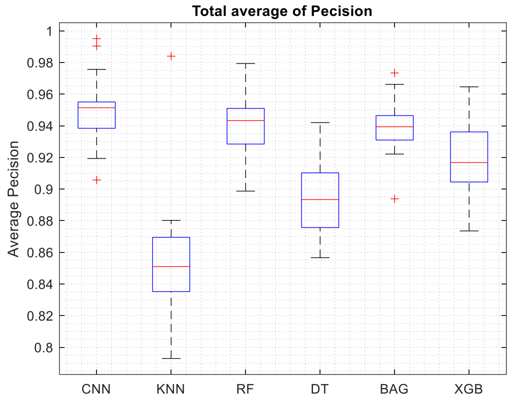

Figure 10. Box Plot of the Overall Average Accuracy

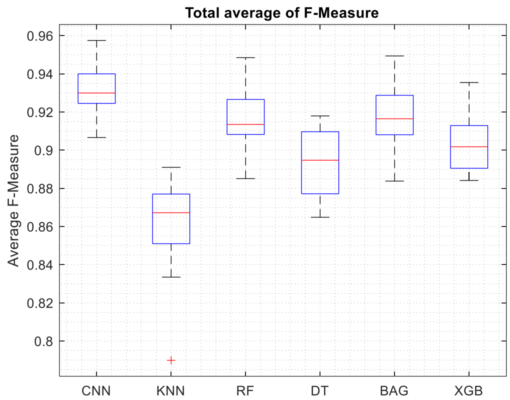

Figure 11. Box Plot of the Overall Mean of the F-Criterion

Based on the box plots in Figures 8 to 11, it can be seen that the proposed method, based on a combination of the human learning optimization algorithm and convolutional neural networks, has achieved higher overall average values compared to other classification methods.

4.3 Comparison of the Proposed Method with Previous Methods

In this section, after evaluating the performance of the proposed method against experimental datasets, the validity of the results obtained from this method is examined by comparing it with previous methods under the same conditions on a specific dataset in the field of customer churn detection (Turatti, 2025). Since the accuracy metric shows the relationship between the actual samples and the samples detected by machine learning models, it can be used as the best metric to evaluate the proposed method. Figure 12 shows a bar graph comparing the proposed method based on different classifications and previous methods. As shown in Figure 12, the proposed method, relying on feature selection based on the human learning optimization algorithm and using these features as input for deep learning, has been able to achieve favorable results for predicting the buying behavior of new customers. Therefore, comparing the results of the proposed method with previous methods, in addition to demonstrating the improvement made in the field of predicting customer buying behavior by the proposed method, also validates the method. The results obtained from the proposed method have shown a significant improvement compared to other methods.

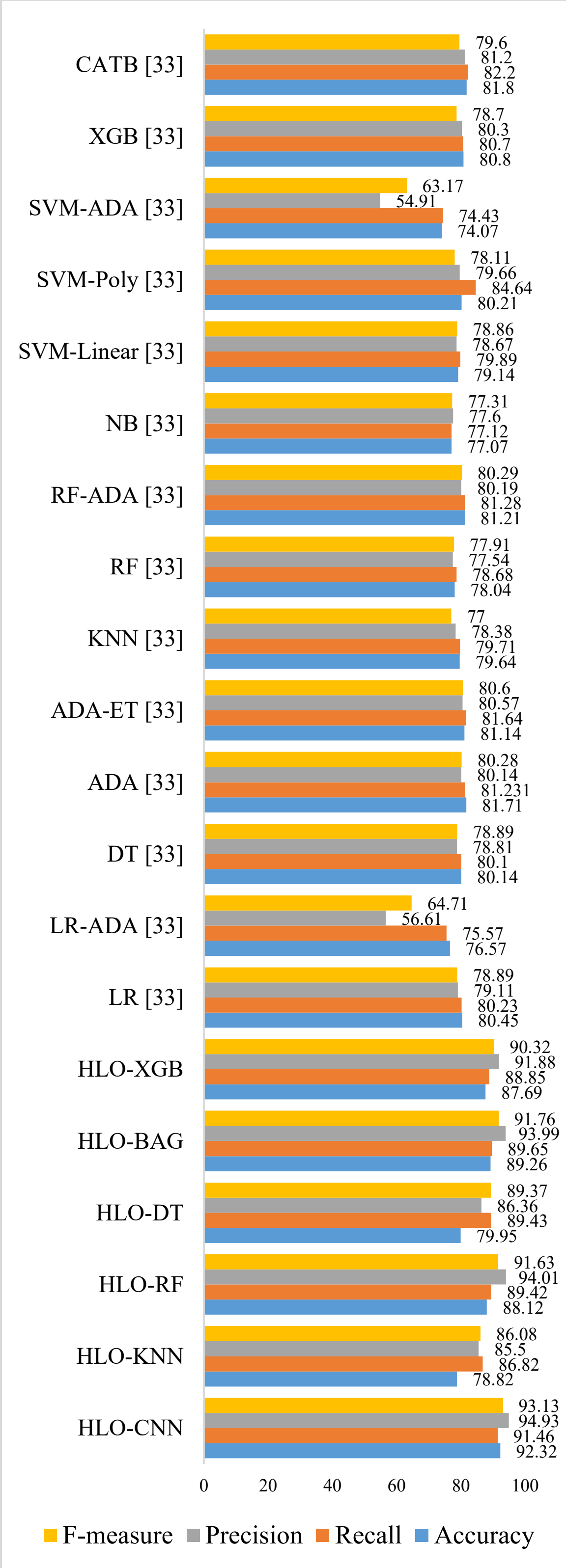

Figure 12. Comparison of the Proposed Method with Previous Methods

This research aims to develop a novel optimization algorithm based on feature selection, combining deep learning and optimization techniques to enhance the efficiency of customer churn prediction. Given the diversity of research variables and the volume of data, the use of a feature selection approach based on metaheuristic optimization in this article is correct to increase the accuracy of customer churn prediction in e-commerce. The proposed algorithm, which is introduced in this study by combining the human learning optimization algorithm and convolutional neural networks, has succeeded in identifying the best parameters for more accurate prediction through the integration of convolutional neural network structures and metaheuristic optimization algorithms, thereby significantly increasing the overall prediction accuracy.

Based on the results obtained and as illustrated in the figure, the combination of the human learning optimization algorithm and convolutional neural networks has outperformed previous methods—including categorized gradient boosting, additive gradient boosting, various support vector machine types, random forest, and other classification methods without using a feature selection phase—in all performance evaluation indicators. These include accuracy, sensitivity, prediction accuracy, and the harmonic mean of the F-measure, due to the selection of precise features as input for the classification method and optimal training on these features. Due to the deep learning capability of convolutional neural networks, the combination of the Human Learning Optimization algorithm and convolutional neural networks has outperformed classification methods that are also combined with the Human Learning Optimization algorithm and use the features it selects. For example, while conventional methods, such as categorized gradient boosting and additive gradient boosting, have achieved an accuracy of around 81%, the proposed method, by achieving an accuracy of approximately 92.32% and a harmonic mean of the F-measure of about 93.13%, demonstrates a significant advantage.

The findings of this research suggest that the combination of optimization and deep learning can serve as an efficient algorithm and a valuable tool for organizations operating in the e-commerce sector to have a more accurate prediction of their customers' behavior and improve customer retention strategies. The proposed algorithm not only has the ability to find the best parameters and model structure but, considering its stable performance across various criteria, can also be used as a general framework for similar problems in other areas of prediction and data classification. This advantage solidifies the position of the combination of the human learning optimization algorithm and convolutional neural networks as an innovative and effective approach, distinguishing it from other common methods.

comparing the results of the proposed method with the reported results in previous studies, it is observed that the HLO based model in combination with different classifications has better performance than most of the previous works. In particular, the HLO-CNN combination has demonstrated a remarkable performance in customer churn prediction with an accuracy of 92.32% and an F1 score of 93.13%, showing a significant improvement compared to most methods presented in Zahmatkesh Zakariaee et al. (2023), including LR, KNN, RF, and SVM.

The results also indicate that the use of feature selection based on the HLO algorithm has improved the performance of various classification models. For instance, HLO-RF and HLO-BAG models have reported better performance with accuracy values of 88.12% and 89.26%, respectively, and F1 scores above 91%, compared to many conventional machine learning models reported in previous studies. This suggests that proper feature selection plays a significant role in enhancing the accuracy of customer churn prediction.

In summary, comparing the obtained results with previous methods indicates that combining evolutionary optimization algorithms for feature selection with machine learning and deep learning models can effectively improve customer churn prediction performance. This shows that the proposed approach, in addition to reducing data dimensions, can identify customer behavioral patterns more accurately and provide acceptable results compared to many existing models.

5) Cunclusion

Online retailers have extensively invested in extracting important information from their customer and product databases. This data is used as a research branch called Market Basket Analysis. The term "market basket analysis" in the retail business refers to research that provides the retailer with the information needed to understand customer buying behavior. This information enables the retailer to understand buyer needs and modify the store layout according to customer needs and shopping patterns, cross-purchasing plans, design discount coupons, or even attract new buyers (very similar to the concept of cross-selling). Therefore, in recent years, machine learning has played an important role in predicting customer buying behavior. In this paper, an online customer purchase prediction system, using feature selection based on the human learning optimization algorithm and deep learning, is presented. The proposed method, by selecting features with the highest mutual information with the class label and customer purchase decision, provides favorable data for deep learning methods and other supervised classification methods. The experimental results show that the proposed method, relying on feature selection based on the human learning optimization algorithm and deep learning, has been able to achieve an accuracy of approximately 92.32% in predicting the buying behavior of new customers. On the other hand, the results obtained from the proposed method have been compared with other methods, and the improvement resulting from the proposed method compared to previous methods is clearly visible.

According to the results and the importance of customer churn prediction in e-commerce, we can suggest several paths for future research. In the competitive environment of e-commerce, companies need to maintain and increase their current customers loyalty, in addition to attracting new customers. This is because existing customer retention usually has less cost than attracting new customers and plays an important role in business sustainability. Therefore, developing more accurate models to identify customers at risk of churn can help organizations adopt more effective strategies to maintain customers.

One of the important paths for future research could be the use of more advanced deep learning models and combining multiple neural network architectures to better identify the complexity of customer behavioral patterns. For example, using deep stacked neural networks that combine multiple architectures, such as multi-layer perceptrons (MLP), convolutional neural networks (CNN), recurrent neural networks (RNN), and long-short-term memory (LSTM) networks, can enhance the model's ability to extract complex patterns from customer behavioral data. In such a framework, feature selection can be primarily done using optimization algorithms, such as human learning optimization algorithm, and then selected features are used as input of deep models.

Additionally, the development of hybrid models, such as deep stacked neural networks, can be considered as a potential direction for future research. In this approach, by combining feature selection based on HLO and deep stacked neural networks, the possibility of achieving very high prediction performance is provided.

The reported results for this method show that such structures can achieve very high levels of accuracy, sensitivity, precision, and F-measure in predicting customer churn. Therefore, investigating and developing these hybrid architecture designs on more diverse datasets in the e-commerce domain can lead to further improvements in the accuracy of customer churn prediction models in future studies.

References

Abbasimehr, H., & Bahrini, A. (2024). Improving churn prediction using imperialist competitive algorithm for feature selection in telecom. *International Journal of Business Information Systems, 45*(4), 490–502. https://doi.org/10.1504/IJBIS.2024.100342

Ahmadi-pour, B. (2025). Data mining in CRM for the telecommunications industry to identify customer churn patterns. In *Proceedings of the First International Conference on Information Technology, Management and Computer Scienc*. (In Persian) https://civilica.com/doc/2083373/

Amin, A., Adnan, A., & Anwar, S. (2023). An adaptive learning approach for customer churn prediction in the telecommunication industry using evolutionary computation and Naïve Bayes. *Applied Soft Computing, 137*, 110103. https://doi.org/10.1016/j.asoc.2023.110103

Babaee Khobdeh, S., Yamaghani, M. R., & Khodaparast, S. (2023). A novel method for clustering basketball players with data mining and hierarchical algorithm. *International Journal of Applied Operational Research, 11*(4), 25–38. https://doi.org/10.71885/ijorlu-2023-1-650

Blastchar. (2018). Telco Customer Churn. Kaggle. https://www.kaggle.com/datasets/blastchar/telco-customer-churn

Cirqueira, D., Hofer, M., Nedbal, D., Helfert, M., & Bezbradica, M. (2020). Customer purchase behavior prediction in e-commerce: A conceptual framework and research agenda. In M. Ceci, C. Loglisci, G. Manco, E. Masciari, & Z. Ras (Eds.), *New Frontiers in Mining Complex Patterns: 8th International Workshop, NFMCP 2019, Held in Conjunction* *with ECML-PKDD 2019, Würzburg, Germany, September 16, 2019, Revised Selected Papers* (pp. 119–136). Springer. https://doi.org/10.1007/978-3-030-48861-1_8

Delafrooz, N., Shadpour, M., & Shahroodi, K. (2026). A model to reduce customer churn using Artificial intelligence-based customer relationship management in the insurance industry. *ATU Press* (in press). (In Persian) https://civilica.com/doc/2265211/

Du, J. (2018). Understanding of object detection based on CNN family and YOLO. *Journal of Physics: Conference Series, 1087*, 062013. https://doi.org/10.1088/1742-6596/1087/6/062013

Durkaya Kurtcan, B., & Ozcan, T. (2023). Predicting customer churn using grey wolf optimization-based support vector machine with principal component analysis. *Journal of Forecasting, 42*(6), 1329–1340. https://doi.org/10.1002/for.2960

Fakheri, S., Nourbakhsh, A., & Yamaghani, M. (2024). Deepfake detection models and methods in artificial intelligence and insights from media and social culture perspective. *Innovation Management and Operational Strategies, 5*(3), 259-287. https://doi.org/10.22105/imos.2024.452298.1344

Fakheri, S., Nourbakhsh, A., & Yamaghani, M. (2024). Deepfake detection models and methods in artificial intelligence and insights from media and social culture perspective. *Innovation Management and Operational Strategies, 5*(3), 259-287. (In Persian). https://doi.org/10.22105/imos.2024.452298.1344

Ghaedi Khani, E., & Shahpasandi, B. (2026). Improving the efficiency of customer churn systems by combining whale-based feature selection and genetic-based ensemble classification. *National Conference on the Application of Artificial Intelligence in the Modern Era*. (In Persian) https//:civilica.com/doc/2294722/

Ghanbari Marvast, M., Sadeghieh, A., & Khademi Zare, H. (2023). Predicting customer churn in the insurance industry using crowd learning algorithms. In *Proceedings of the Sixth International Conference on Mechanical, Industrial and Aerospace Engineering* (pp. 1–10). https://en.civilica.com/l/101531/

Golshan, M., & Rahimi Khoshmakani, M. (2025). A review of customer churn prediction methods in online stores using data mining techniques. In *First International Conference on Information Technology, Management and Computer Science*. (In Persian) https://civilica.com/doc/2084042/

Guerola-Navarro, V., Gil-Gómez, H., Oltra-Badenes, R., & Soto-Acosta, P. (2024). Customer relationship management and its impact on entrepreneurial marketing: A literature review. *International Entrepreneurship and Management Journal, 20*(2), 507-547. https://doi.org/10.1007/s11365-022-00800-x

Guerola-Navarro, V., Oltra-Badenes, R., Gil-Gómez, H., & Gil-Gómez, J. A. (2020). Customer relationship management (CRM): A bibliometric analysis. *International Journal of Services Operations and Informatics, 10*(3), 242-268. https://doi.org/10.1504/IJSOI.2020.108988

Gurung, N., Hasan, M. R., Gazi, M. S., & Chowdhury, F. R. (2024). AI-based customer churn prediction model for business markets in the USA: Exploring the use of AI and machine learning technologies in preventing customer churn. *Journal of Computer Science and Technology Studies, 6*(2), 19–29. https://doi.org/10.32996/jcsts.2024.6.2.3

He, C., & Ding, C. H. (2024). A novel classification algorithm for customer churn prediction based on hybrid Ensemble-Fusion model. *Scientific Reports, 14*(1), 20179. https://doi.org/10.1038/s41598-024-71168-x

Heshmati, Z. A., & Haq, B. (2023). Predicting traffic crashes using modern convolutional graph networks to improve traffic experience. In *Proceedings of the Fifth International Conference on Soft Computing* (pp. 1–10). https://doi.org/10.1007/978-3-030-65045-0_1

Hosseini, S. S., Yamaghani, M., & Poorzaker Arabani, S. (2024). A review of methods for detecting multimodal emotions in sound, image and text. *Journal of Applied Research on Industrial Engineering, 11*(4), 669–705. https://doi.org/10.22105/jarie.2024.413341.1560

Hosseini, S., & Yaghmani, M. (2024). Identification of multi-modal human emotions by artificial intelligence and its application in psychology. *Arak University of Medical Sciences Journal, 27*(4), 194–204. https://doi.org/10.22100/jams.v27i4.7585

Jafari Dehkordi, M., Aali Pour, N. A., & Zangeneh, H. (2025). A study on identifying potential customers and predicting prevention of customer churn in e-commerce based on artificial intelligence technologies and generalizing its results to other countries. In *Sixth National Conference and Third International Conference on Novel Business Management Patterns in Unstable Conditions*. https://civilica.com/doc/2098736/

Jamali, S., & Navaei, Y. D. (2016). A two-level product recommender for E-commerce sites by using sequential pattern analysis. *International Journal of Integrated Engineering, 8*(1), 15–24. https://doi.org/10.30880/ijie.2016.08.01.003

Jiang, P., Liu, Z., Zhang, L., & Wang, J. (2023). Hybrid model for profit-driven churn prediction based on cost minimization and return maximization. *Expert Systems with Applications, 228*, 120354. https://doi.org/10.1016/j.eswa.2023.120354

Khare, P., & Arora, S. (2024). Predicting customer churn in SaaS products using machine learning. International *Research Journal of Engineering and Technology (IRJET), 11*(5), 754–765. https://doi.org/10.1007/978-981-99-8438-1_26

Khattak, A., 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

Lalwani, P., Mishra, M. K., Chadha, J. S., & Sethi, P. (2022). Customer churn prediction system: A machine learning approach. *Computing, 104*(2), 271–294. https://doi.org/10.1007/s00607-021-00908-y

Liu, Y., Fan, J., Zhang, Y., & Wang, H. (2023). Research on telecom customer churn prediction based on ensemble learning. *Journal of Intelligent Information Systems, 60*(3), 759–775. https://doi.org/10.1007/s10844-022-00739-z

Lokesh, S., & Vasantha, S. (2023, April). Customer relationship management a decision support system: Bibliometric analysis 1990–2023. In *International Conference on Artificial Intelligence and Digital Twin Technology* (pp. 55-68). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-77799-8_6

Matuszelański, K., & Kopczewska, K. (2022). Customer churn in retail e-commerce business: Spatial and machine learning approach. *Journal of Theoretical and Applied Electronic Commerce Research, 17*(1), 165–198. https://doi.org/10.3390/jtaer17010009

Murugesan, P., & Trivedi, P. (2024). Tri-strategy Remora optimization algorithm-based support vector machine for customer churn prediction. In *Proceedings of the 2024 International Conference on Integrated Circuits and Communication Systems (ICICACS)* (pp. 1–10). IEEE. https://doi.org/10.1109/ICICACS.2024.1234567

Namini, S. (2025). Extracting appropriate features for behavior analysis and predicting customer attrition in banks. *New Researches in the Smart City, 3*(1), 29-49. (in Persian) https://civilica.com/doc/2257044/

Navaei, Y. D., & Afzali, M. (2014). A survey on product recommendation system in E-commerce. *International Journal of Computer & Information Technologies (IJOCIT), 2*(4), 585–600. https://www.researchgate.net/publication/305755533_A_Survey_on_Product_Recommendation_System_in_E-Commerce

O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*.

Oltra-Badenes, R., Gil-Gomez, H., Guerola-Navarro, V., & Vicedo, P. (2019). Is it possible to manage the product recovery processes in an ERP? Analysis of functional needs. *Sustainability, 11*(16), 4380. https://doi.org/10.3390/su11164380

Quek, J. Y. V., Pang, Y. H., Lim, Z. Y., Ooi, S. Y., & Khoh, W. H. (2023). Customer churn prediction through attribute selection analysis and support vector machine. *Journal of Telecommunications and the Digital Economy, 11*(3), 180–194. https://doi.org/10.18080/jtde.v11n3.777

Rad, R. A., Yamaghani, M. R., & Nourbakhsh, A. (2023). A survey of sentiment analysis methods based on graph neural network. *International Journal of Data Science and Analytics*, *19*(3), 389-417. https://doi.org/10.21203/rs.3.rs-3173515/v1

Sadr, H., Nazari, M., Khodaverdian, Z., Farzan, R., Yousefzadeh-Chabok, S., Ashoobi, M. T., Hemmati, H., Hendi, A., Ashraf, A., Pedram, M. M., Hasannejad-Bibalan, M., & Yamaghani, M. R. (2025). Unveiling the potential of artificial intelligence in revolutionizing disease diagnosis and prediction: A comprehensive review of machine learning and deep learning approaches. *European Journal of Medical Research, 30*(1), 418. https://doi.org/10.1186/s40001-025-02680-7

Saha, S., Saha, C., Haque, M. M., Alam, M. G. R., & Talukder, A. (2024). ChurnNet: Deep learning enhanced customer churn prediction in telecommunication industry. *IEEE access*, *12*, 4471-4484. https://doi.org/10.1109/ACCESS.2024.1234567

Sam Deli, Z. (2023). Predicting mesh failure using machine learning techniques. In *Proceedings of the First National Conference on Innovation in Engineering: A Path to Development* (pp. 1–10). Islamic Azad University of Chalus. https://en.symposia.ir/NCIIE01

Sam, G., Asuquo, P., & Stephen, B. (2024). Customer churn prediction using machine learning models. *Journal of Engineering Research and Reports, 26*(2), 181–193. https://doi.org/10.9734/jerr/2024/v26i21081

Sarabadani, E., & Cheragh, S. (2024). A model based on the random forest algorithm and Jaya optimization to predict bank customer churn. *Engineering Management and Soft Computing, 9*(2), 132-148. https://civilica.com/doc/2094599/

Saranya, G., & Pravin, A. (2021). Feature selection techniques for disease diagnosis system: A survey. In *Artificial intelligence techniques for advanced computing applications: Proceedings of ICACT 2020* (pp. 1–10). Springer. https://doi.org/10.1007/978-3-030-65045-0_1

Shir Ali Shahreza, M., & Mahmoudzadeh, M. (2026). Predicting bank customer churn using machine learning. *Financial Research Journal, 27*(2), 218-245. https://civilica.com/doc/2311695/

Subramanian, R. S., Yamini, B., Sudha, K., & Sivakumar, S. (2025). Ensemble-based deep learning techniques for customer churn prediction model. *Kybernetes*, *54*(9), 4594-4625.. https://doi.org/10.1108/k-08-2023-1516

Suh, Y. (2023). Machine learning based customer churn prediction in home appliance rental business. *Journal of Big Data, 10*(1), 41. https://doi.org/10.1186/s40537-023-00721-8

Turatti, R. C. (2025). Application of artificial intelligence in forecasting consumer behavior and trends in E-commerce. *Brazilian Journal of Development, 11*(3), e78442. https://doi.org/10.34117/bjdv11n3-039

Turatti, R. C. (2025). Application of artificial intelligence in forecasting consumer behavior and trends in E-commerce. *Brazilian Journal of Development, 11*(3), e78442. https://doi.org/10.34117/bjdv11n3-039

Usman-Hamza, F. E., Balogun, A. O., Nasiru, S. K., Capretz, L. F., Mojeed, H. A., Salihu, S. A., Akintola, A. G., Mabayoje, M. A., & Awotunde, J. B. (2024). Empirical analysis of tree-based classification models for customer churn prediction. *Scientific African, 23*, e02054. https://doi.org/10.1016/j.sciaf.2023.e02054

Wagh, S. K., Andhale, A. A., Wagh, K. S., Pansare, J. R., Ambadekar, S. P., & Gawande, S. H. (2024). Customer churn prediction in telecom sector using machine learning techniques. *Results in Control and Optimization, 14*, 100342. https://doi.org/10.1016/j.rico.2023.100342

Wu, J. (2017). Introduction to convolutional neural networks. *National Key Lab for Novel Software Technology. Nanjing University. China*, *5*(23), 495. https://cs.nju.edu.cn/wujx/paper/CNN.pdf

Zahmatkesh Zakariaee, A., Sadr, H., & Yamaghani, M. R. (2023). A new hybrid method to detect risk of gastric cancer using machine learning techniques. *Journal of AI and Data Mining, 11*(4), 505–515. https://doi.org/10.22044/jadm.2023.13377.2464