Customer Churn Prediction using Sandpiper Optimization with Bidirectional Gated Recurrent Unit for Business Intelligence

Akbal Omran Salman [1], Mazin Abed Mohammed*[2]

[1] Department of Control & Automation Techniques Engineering, Electrical Engineering Technical College, Middle Technical University, Baghdad, Iraq

[2] College of Computer Science and Information Technology, University of Anbar, Ramadi, 31001, Anbar, Iraq

Emails: akbal.o.salman@mtu.edu.iq ; Mazinalshujeary@uoanbar.edu.iq

Abstract

Recent innovation in business intelligence (BI) assists companies to stay successful and competitive with the increasing business trend. Businesses have started to examine the succeeding level of data analytics and BI solution. At the same time, Customer Churn Prediction (CCP) is an essential procedure involved in business decision making that effectually determines the churn of clients and performs adequate processes to retain customers. With this motivation, this paper presents a sandpiper optimization with the bidirectional gated recurrent unit (SPO-BiGRU) for CCP on BI applications. The SPO-BiGRU model aims for determining the occurrence of customers into churners or non-churner. In addition, the SPO-BiGRU technique involves pre-processing, classification, and hyperparameter optimization. Followed by, the BiGRU model is applied to perform the predictive process. At last, the SPO algorithm is applied to optimally adjust the hyperparameters involved in the BiGRU model. For validating the enhanced performance of the SPO-BiGRU method, a wide range of simulations take place and the results are inspected under varying aspects. The experimental results portrayed the supremacy of the SPO-BiGRU technique over the recent state of art approaches.

Keywords; Business intelligence, Customer churn prediction, Deep learning, Sandpiper optimization, BiGRU model.

1. Introduction

The latest developments in business intelligence became crucial to be competitive and successful in the emerging business trend. Thus, businesses of each size (small and medium-scale businesses) have introduced the investment in the subsequent levels of business intelligence and data analysis outcome [1]. The efficient use of business intelligence techniques removes the analyses as well as visualizes the efficacy indications from a massive number of big data enterprises. It reduces the costs and increases the speed of resolving problems with useful information. However, it can be very critical for the business organization for generating proper decisions owing to numerous challenges [2]. Text analytics comprises wide-ranging applications. It can be employed to arrange product analyses and apply shopping pattern detection with categories, topics, annotations, tags, and keywords in textual information. The software is indispensable to running the advanced analytical method that concludes, prescribes and predicts information from the text can be very difficult for creating enterprises application.

Customer Relationship Management (CRM) is a wide-ranging approach for strengthening, building, and managing long-lasting and loyal customer relations [3]. It is extensively applied and largely acknowledged in dissimilar areas, for example, retail market, telecommunications, banking, insurance, and so on. The key objective is customer maintenance. The significance of this purpose is apparent,

DOI: https://doi.org/10.54216/AJBOR.060103
Received: May 15, 2021 Accepted: November 04, 2021

considering the fact that the costs for customer attainment are considerably higher when compared to the costs of customer maintenance (few instances it is twenty times costly). Therefore, methods to apply and develop a customer retention model (churning model) is needed and Business Intelligence (BI) application is essential [4]. In the dynamic market, churn can be the result of aggressive competitive strategies, lower-level customer satisfaction, regulations, new products, and so on. The churning model aims to recognize customers and earlier Churning signals with the highest probability.

In the telecommunication market, the customer could easily dismiss their subscription with the supplier and move to another business for better price rates and quality of services [5]. This challenging problem in marketing is called “customer churning”. Moreover, on the whole market, the costs of receiving novel customers are considerably higher. It was specified that the annual Churning rates in telecommunication range from 20-40% while the cost of gaining a novel customer might be 5-10 times superior to maintaining the existing client [6]. Therefore, the customers are considered a valuable resource for the business. For that reason, the telecommunication market has become very competitive and dynamic. As per the details, customer retention is considered the main aspect and great significance of CRM [7].

Accordingly, the Churning prediction is widely employed for identifying the customers about their subscription end/exit the businesses for another competitive supplier. A precise Churning prediction can effectively support economic marketing movement and customer retention plan, and, consequently, it can lead to the most important investments to suppliers. For enduring this aggressive competition, the telecommunication business has become very efficient through the investment of additional developments in machine learning (ML) based and data mining approaches for Churning analyses, prediction, and management [8]. Numerous ML systems were introduced in the work to predict the Churning. Learning from imbalanced class distribution is very difficult for each conventional ML method as they are inclined to classify the precise class and neglect the rarest one. The main problem is many factors affect customer churn and the relationship among them is very complex.

This paper presents a sandpiper optimization with the bidirectional gated recurrent unit (SPO-BiGRU) for CCP on BI applications. The SPO-BiGRU model aims for determining the occurrence of customers into churners or non-churner. In addition, the SPO-BiGRU technique involves preprocessing, classification, and hyperparameter optimization. Followed by, the BiGRU model is applied to perform the predictive process. At last, the SPO algorithm is applied to optimally adjust the hyperparameters involved in the BiGRU model. For validating the enhanced performance of the SPO-BiGRU model, a wide range of simulations take place and the results are inspected under varying aspects.

2. Recent State of Art CCP Approaches

Vafeiadis et al. [9] proposed a comparison analysis of the commonly used ML method employed to the stimulating challenge of client churn predictions in the telecommunication field. Initially, each model is evaluated and applied by cross-validation on a common, open-source database. Next, the efficiency improvements given through boosting have been investigated. Coussement et al. [10] carried out in a churning predictive modeling context, benchmark an enhanced logit method against 8 advanced data mining methods that utilize input data, including real-time cross-section information from huge European telecommunication suppliers. The result leads to succeeding decisions. (i) Analyst improved acknowledged that the data-preparation method selection essentially impacts churning predictive efficacy (ii) The enhanced logistic regression is competing with data mining methods.

Tsai and Chen [11] proposed the significant procedures of emerging MOD client churning predictive methods by data mining methods. They comprise the pre-processing phase to select significant parameters by relation rules, that haven’t been employed previously, the NN and DT that is broadly adopted in the survey, and 4 assessment measure including F-measure, predictive accuracy, precision, and recall, haven’t been considering to study the method efficiency. Amin et al. [12] present a smart rule-based decision-making method, based on RST, for extracting significant decision rules associated with the client churn and non-churn. The presented method efficiently performs classification of churning from non-churn clients, as well as predictions of these clients, will churn or might probably churn.

In De Bock and Van den Poel [13], 2 rotation-based ensemble classifier is presented as modeling methods for client churning predictions. In Rotation Forests, feature extraction is employed for feature

DOI: https://doi.org/10.54216/AJBOR.060103
Received: May 15, 2021 Accepted: November 04, 2021

subset to rotate the input data to train a base classifier, whereas RotBoost integrates Rotation Forest using AdaBoost.

Shirazi and Mohammadi [14] create a prediction churning method through big data containing the structured archival data, combined with unstructured data in sources like phone conversation logs, online web pages, and a number of website visits for the financial sector. Also, it studies the impact of distinct factors of customer behavior on churn decisions. The Datameer big data analytics on the Hadoop and prediction methods with the SAS business intelligence are employed for examining the customer retirement journey path. In De Caigny et al. [15], a novel hybrid system, the LLM, is projected to improve categorizing of information. The concept behindhand the LLM is distinct methods built on a segment of the information instead of on the whole data set lead to healthier prediction performances while retaining the simplicity of the model. The LLM contains 2 phases: a prediction stage and a segmentation stage.

Calzada-Infante et al. [16] present a new technique for extracting the dynamic significance of all the customers with social network analysis methods with a binary classification model named similarity forest. The dynamic significance of all the clients is defined by employing numerous centrality metrics on temporal graphs, to characterize the relations among clients and extracts behavioral pattern of non-churners and churners. Vo et al. [17] introduce a client churning predictive method using the unstructured information, i.e., the stated content in telecommunication. They gathered a larger-scale call center data set with 2 million calls from around a hundred thousand clients and performed wide-ranging research.

3. Material and Methods

This paper has presented an effective SPO-BiGRU technique for CCP on BI applications. The SPO-BiGRU technique comprises of different processes namely pre-processing, BiGRU based prediction, and SPO based hyperparameter tuning process. The detailed working of these processes is elaborated in the following sections.

3.1Design of BiGRU based Prediction Process

Lately, GRU, a class of RNN, was presented for handling gradient exploding or vanishing issues. GRU is a simple and powerful alternate to LSTM network. Like LSTM method, GRU is developed for adoptively reset/updating the memory contents with r^j reset gate and a z^j upgrade gate which is analogous to input and forgets gates of LSTM. In comparison with LSTM, GRU doesn't contain a memory cell and have 2 gates. The GRU activation h_t^j in time t is the linear interpolation of prior activation h_{t-1}^j and candidate activation $\tilde{h}_t^j$.

To calculate the state $\tilde{h}_t^j$ of jth GRU at time step t, they employ below formula:

$$h_t^j = \left(1 - z_t^i\right)h_{t-1}^j + z_t^j\tilde{h}_t^j \qquad (1)$$

In which $\tilde{h}_{t-1}^j$ and $\tilde{h}_{t-1}^j$ corresponding to the novel candidate and prior memory content, correspondingly. z_t^jcharacterizes the update gate which allows the method to determine the quantity of previous data (from prior time step) to be transported with the future and the quantity of novel memory contents to be included [18].

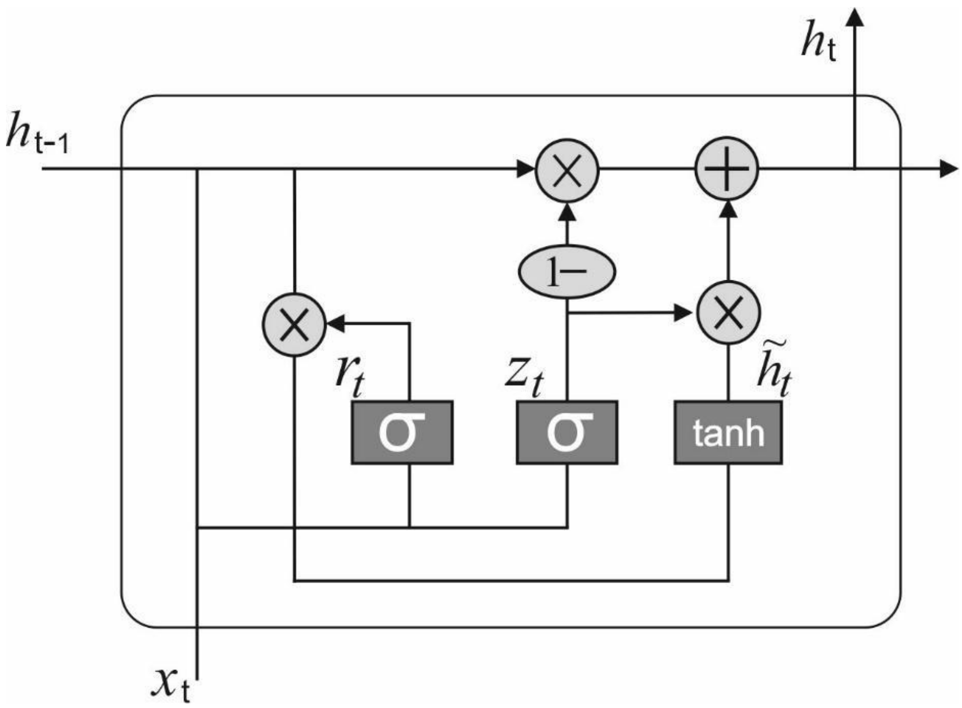
To compute the update gate z_t for time step t, they utilize the prior hidden state h_{t-1} and the present input x_t as:

$$z_t = \sigma(W_Z x_t + U_Z h_{t-1}) \qquad (2)$$

The novel memory content $\tilde{h}_{t-1}$ is estimated by:

$$\tilde{h}_t = \ tanh\ (Wx_t + r_t \odot Uh_{t-1}) \qquad (3)$$

DOI: https://doi.org/10.54216/AJBOR.060103
Received: May 15, 2021 Accepted: November 04, 2021



Figuer 1. GRU Structure

Which ⊙ represent the Hadamard product (called the element-wise product) and r_t denotes the reset gate employed for determining the number of data to forget from the historical as

$$r_t = \sigma(W_r x_t + U_r h_{t-1}) \quad (4)$$

A graphical depiction of GRU unit is demonstrated in Fig. 1.

GRU is quicker when compared to LSTM on training as GRU has basic framework with less parameters and thus employs lesser memory.

One downside of GRU network could utilize the prior context, without considering the future contexts, hence they handle sequence from forward to backward that leads to data loss. Hence several scientists have utilized bi-GRU i.e., able to process information in both directions, and data from single hidden layer is gathered in the output layer.

The basic framework of Bi-GRU network is essentially making 2 distinct GRUs. The input order is provided to one network in standard time sequence (from right to left, for Arabic) and for other, in reverse time sequence. At all the steps, the output of both networks is habitually integrated (there is another choice, i.e., summation). This architecture could offer comprehensive context data.

3.2Design of SPO based Hyperparameter Tuning Process

In this subsection, the stepwise process is considered to attain the optimum value of DL based FDCN through Sand piper optimization method. Initially, the Sandpiper optimization method generates the early uniformly distributed population of Sandpiper for optimizing the best variable value of DL based FDCN method. The optimum solution is updated by the Sand piper optimization method. Fig. 2 illustrates the flowchart of SPO. The complete steps are given in the following,

DOI: https://doi.org/10.54216/AJBOR.060103
Received: May 15, 2021 Accepted: November 04, 2021

Step 1: Initiation

Initiate the population of the sandpiper, that employed to reduce collision $q = 0,1,2, \dots .iter_{M\sigma x}$ (MRI image q with SPO and find the cardiac diseases).

Step 2: Arbitrary generation

Afterward, initiation, the input parameter DL based FDCN method are arbitrarily generated with sandpiper optimization method. In this phase, the sandpiper's higher fitness value is selected based on attacking behavior and sandpiper fitness migration [19].

Step 3: Fitness Function

The FF of the solutions is $1_{pk}^{(x)}$ measured and the objective function represents the optimization of function $\beta_{cc} and \lambda_{cc}$.

Step 4: Migration behavior

This method describes the life of sandpiper birds that move from one position to another positions. Now, collision might take place, it creates higher computation difficulty and cost function and decreases the accurateness.

Step 4.1: Collision prevention

In this phase, $\overrightarrow{K_{yl}}$ signified as optimal collision prevention agents in the searching method for detecting food (CVD detected image) as

$$\overrightarrow{K_{yl}} = K_B \times \overrightarrow{L_{yl}}(q) \tag{5}$$

whereas K_B represent the collision avoidance, $\overrightarrow{K_{yl}}$ denotes the position of the searching term which won't collide among further searching conditions, $\overrightarrow{K_{yl}}$ (q) symbolizes the current location of searching (detecting) conditions, q means the existing iteration in search space (diseases).

Step 4.2: Coverage in direction of optimal neighbor

Afterward, the collision avoidance method converges the search method to the direction of optimal neighbor. $\overrightarrow{L_{ayu}}$ indicates the optimal searching measures, that decreases the computation difficulty as

$$\overrightarrow{N_{yl}} = K_B \times (\overrightarrow{L_{ayu}}(q) - \overrightarrow{L_{yl}}(q) \tag{6}$$

Let $\overrightarrow{N_{yl}}$ be the position of probing measure $\overrightarrow{L_{yl}}$represents the optimal penetrating measure $\overrightarrow{L_{ayu}}$ signifies the values is lesser while the direction to the neighbor, K_B indicates the collision prevention.

Step 5: Updating the Sandpiper position

In this phase, update the location with the sandpiper optimization behavior that decreases the computation complexity as well as decreases the cost functions (Eq. (6)) and rises the performance.

A group of variables is utilized for reducing the computation difficulty, cost functions and increases the performance, $\lambda_{cc} = 3 \times 10^{-3}, \beta_{cc} = 4$. $1_{pk}^{(x)}$ denotes the binary mask utilized in the optimization method.

$$1_{pk}^{(x)} = K_A \times (\overrightarrow{L_{ayu}}(q) - \overrightarrow{L_{yl}}(q) \tag{7}$$

In which $\overrightarrow{N_{yl}}$ represent the position of probing measure (that measures the computation difficulty in CVD finding) $\overrightarrow{L_{yl}}$ represents the optimal search operation (detection of CVD disease) and decrease cost functions $\overrightarrow{L_{ayu}}$ means the values are lesser once the direction towards the neighbors. K_A represent the arbitrary parameter.

$$\overrightarrow{V_{yl}} = \overrightarrow{K_{yl}} + \overrightarrow{N_{yl}} \tag{8}$$

DOI: https://doi.org/10.54216/AJBOR.060103
Received: May 15, 2021 Accepted: November 04, 2021

$\overrightarrow{K_{yl}}$ denotes the position of finding the image with optimization method to maximize and minimize objective function in search conditions. Through the sandpiper optimization method, the abovementioned equation is enhanced. Now, $\beta = 10^{-4}$ and stated that automated recognition in the testing and training stage. The abovementioned equations are upgraded for optimizing or reducing the cost functions to distinguish the cardiovascular diseases with MRI image analyses

Step 6: Attacking behavior

During migration, the sandpiper might alter its angle and speed, later attacking the accurate location for examining the images at each point over the CVD recognition. The sandpipers could generate spiral behaviors to attack prey (CVD image) in the air. The 3D assessment of the attack (detection of CVD disease) behavior cardiac cycle as:

$$I' = W_{radius} \times sinc(x) \quad (9)$$

$$S' = W_{radius} \times cosc(x) \quad (10)$$

$$Q' = W_{radius} \times x \quad (11)$$

$$w = t \times \alpha^{cj} \quad (12)$$

whereas W_{radius} represents the radius of each spin, x indicates the parameter lies in the interval of $[0 \leq C \leq 2\pi]$, t and j signify the spiral shape constantly, and α represents base of natural method. Assume the t and j values are 1 and it performs as constant is upgraded. Now, the sandpiper easily attacks prey from Eqs. (9-12). This procedure could detect cardiovascular diseases and decrease the cost functions. Hence, the location is upgraded by,

$$Accuracy\ K_A = 0.5 \times W_{rand} \quad (13)$$

$$\overrightarrow{L_{yl}}(q) = \left(\overrightarrow{V_{yl}} \times (i' + j' + q')\right) \times \overrightarrow{L_{ayl}}(q) \quad (14)$$

whereas $\overrightarrow{L_{ayl}}(q)$ upgrade the locations of detection cardiac disease and give optimum solution saves the outcome.

Step 6: Termination

In this phase, the sandpiper optimization method to enhance objective functions like computation difficulty and cost functions are maximized and minimized the accurateness to distinguish cardiovascular disease in DL based FDCN method would repeat step 3 until the end conditions are encountered. Finally, the result of sandpiper optimization method is coming with skilled position of optimum objective function values for FDCN method.

DOI: https://doi.org/10.54216/AJBOR.060103
Received: May 15, 2021 Accepted: November 04, 2021

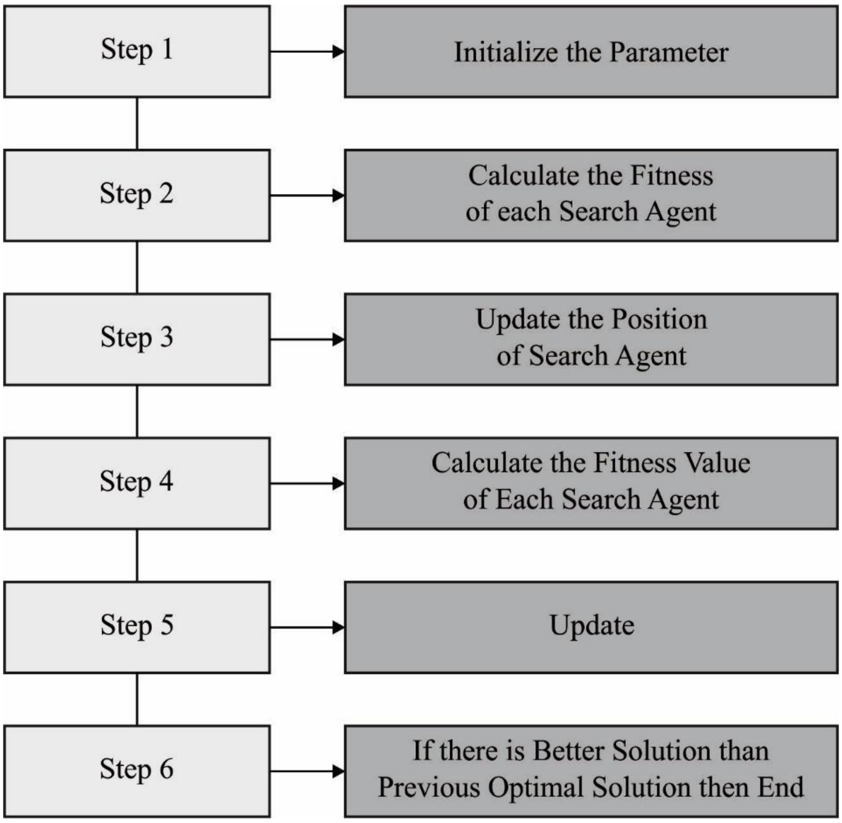


Figure 2. Flowchart Sandpiper Optimization Algorithm

4. Experimental Validation

In this section, the experimental outcomes analysis of the SPO-BiGRU approach takes place using 3 different datasets. An extensive comparative results analysis of the SPO-BiGRU model with other techniques is shown in Table 1 [20].

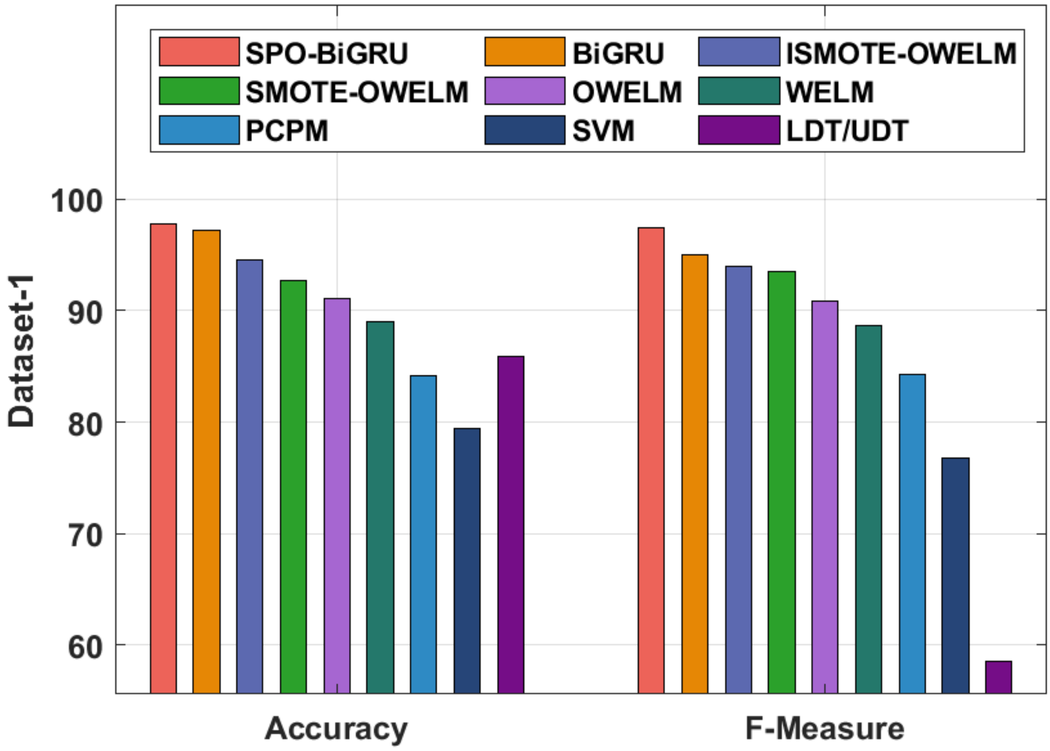
Fig. 3 investigates the accuracy and F-measure analysis of the SPO-BiGRU model on the test dataset-1. With respect to accuracy, the SVM, PCPM, LDT/UDT, and WELM techniques have reached lower accuracy values of 79.4%, 84.2%, 85.9%, and 89% respectively. In line with, the OWELM, SMOTE-OWELM, and ISMOTE-OWELM techniques have gained moderate accuracy of 91.10%, 92.7%, and 94.5% respectively. Though the BiGRU model has resulted in a near optimal accuracy of 97.2%, the proposed SPO-BiGRU technique has accomplished improved accuracy of 97.8%.

Table 1. Comparative analysis of SPO-BiGRU Method for Applied Dataset in terms of Accuracy and F-Score

Methods	**Dataset-1**		**Dataset-2**		**Dataset-3**	
	Accuracy	**F-Measure**	**Accuracy**	**F-Measure**	**Accuracy**	**F-Measure**
SPO-BiGRU	97.80	97.40	96.50	96.70	98.10	97.00
BiGRU	97.20	95.00	95.70	96.20	94.80	93.30
ISMOTE-OWELM	94.50	94.00	92.50	92.30	91.40	91.30
SMOTE-OWELM	92.70	93.50	92.30	92.20	90.40	90.40
OWELM	91.10	90.90	90.20	89.90	89.20	88.40

DOI: https://doi.org/10.54216/AJBOR.060103
Received: May 15, 2021 Accepted: November 04, 2021

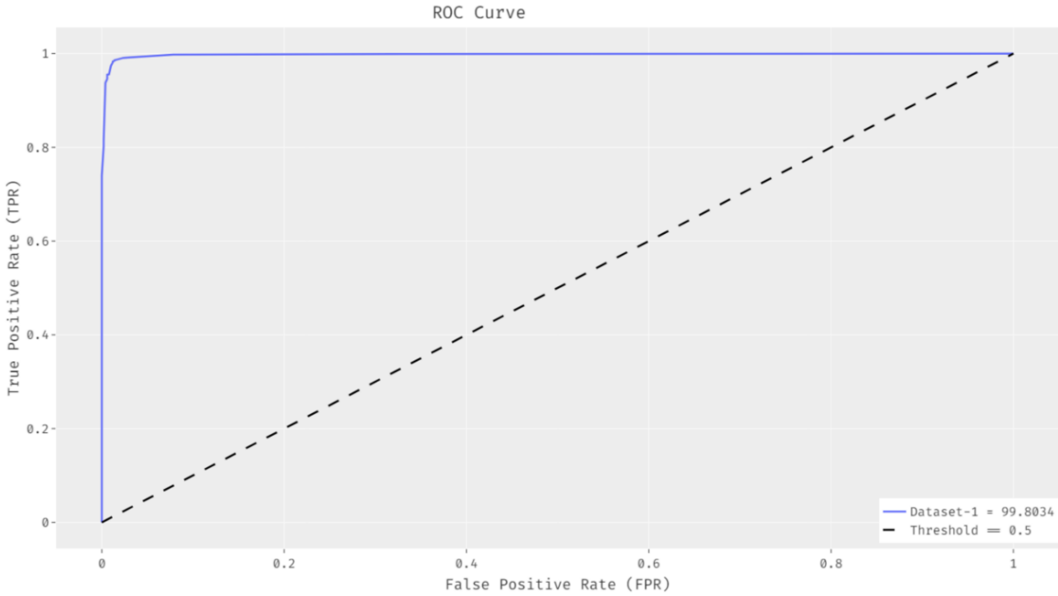
WELM	89.00	88.70	88.10	87.70	87.40	85.70
PCPM	84.20	84.30	83.30	83.60	82.30	81.30
SVM	79.40	76.80	73.00	73.60	68.40	68.70
LDT/UDT	85.90	58.50	77.50	67.50	58.50	59.70



Figuer 3. Result analysis of SPO-BiGRU model under dataset-1

Relating to F-measure, the SVM, PCPM, LDT/UDT, and WELM approaches have attained minimum F-measure values of 76.8%, 84.3%, 58.5%, and 88.7% correspondingly. Similarly, the OWELM, SMOTE-OWELM, and ISMOTE-OWELM techniques have obtained moderate F-measure of 90.90%, 93.5%, and 94% correspondingly. Then, the BiGRU method has resulted in a near optimal F-measure of 95%, the presented SPO-BiGRU methodology has accomplished higher F-measure of 97.4%.

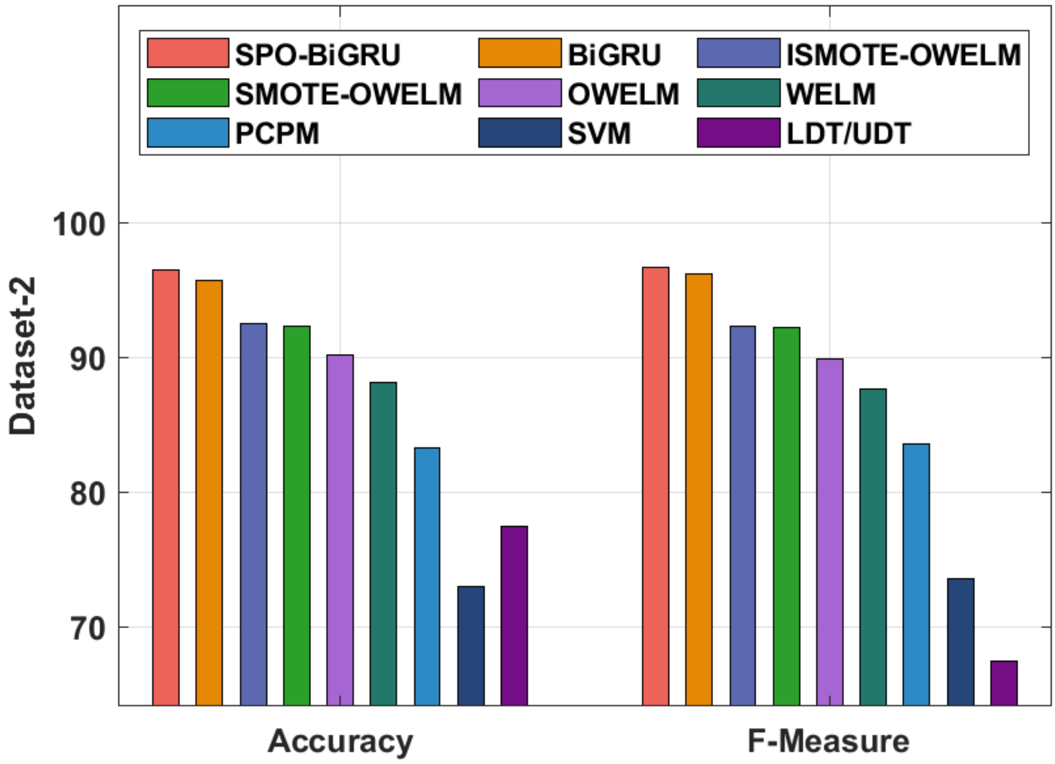
DOI: https://doi.org/10.54216/AJBOR.060103
Received: May 15, 2021 Accepted: November 04, 2021



Figuer 4. ROC analysis of SPO-BiGRU technique under dataset-1

Fig. 4 showcases the ROC analysis of the SPO-BiGRU technique under dataset-1. The figure depicted that the SPO-BiGRU approach has resulted in a higher ROC of 99.8034.

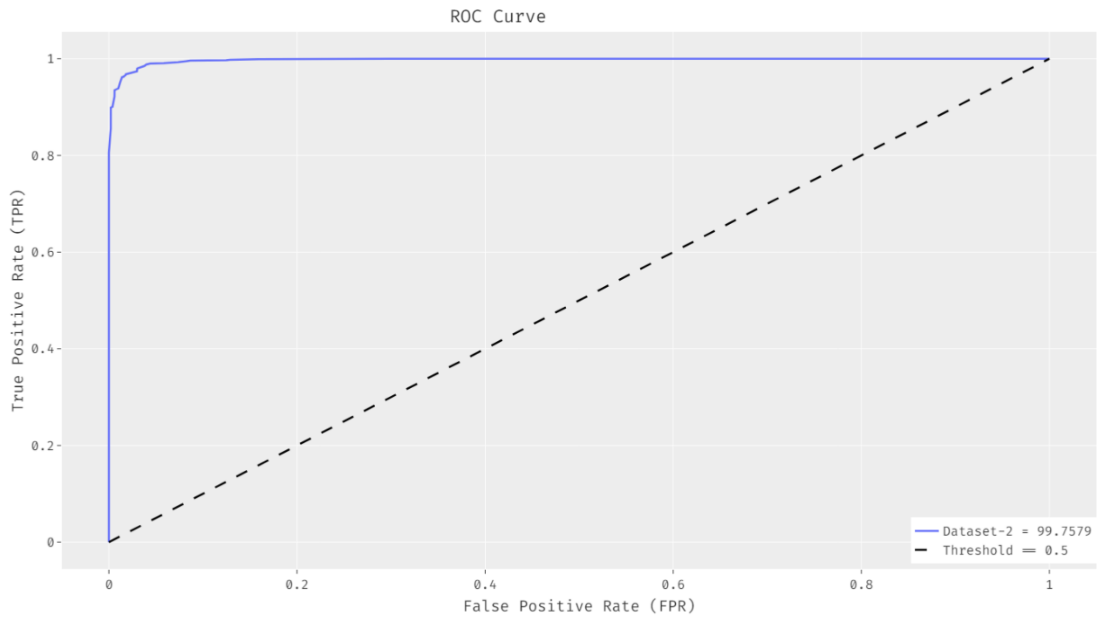
Fig. 5 examines the accuracy and F-measure analysis of the SPO-BiGRU system on the test dataset-2. With respect to accuracy, the SVM, PCPM, LDT/UDT, and WELM approaches have reached lower accuracy values of 73%, 83.3%, 77.5%, and 88.1% correspondingly. Followed by, the OWELM, SMOTE-OWELM, and ISMOTE-OWELM techniques have gained moderate accuracy of 90.20%, 92.3%, and 92.5% correspondingly. Next, the BiGRU model has resulted in a near optimal accuracy of 95.7%, the proposed SPO-BiGRU technique has accomplished increased accuracy of 96.5%.



Figuer 5. Result analysis of SPO-BiGRU model under dataset-2

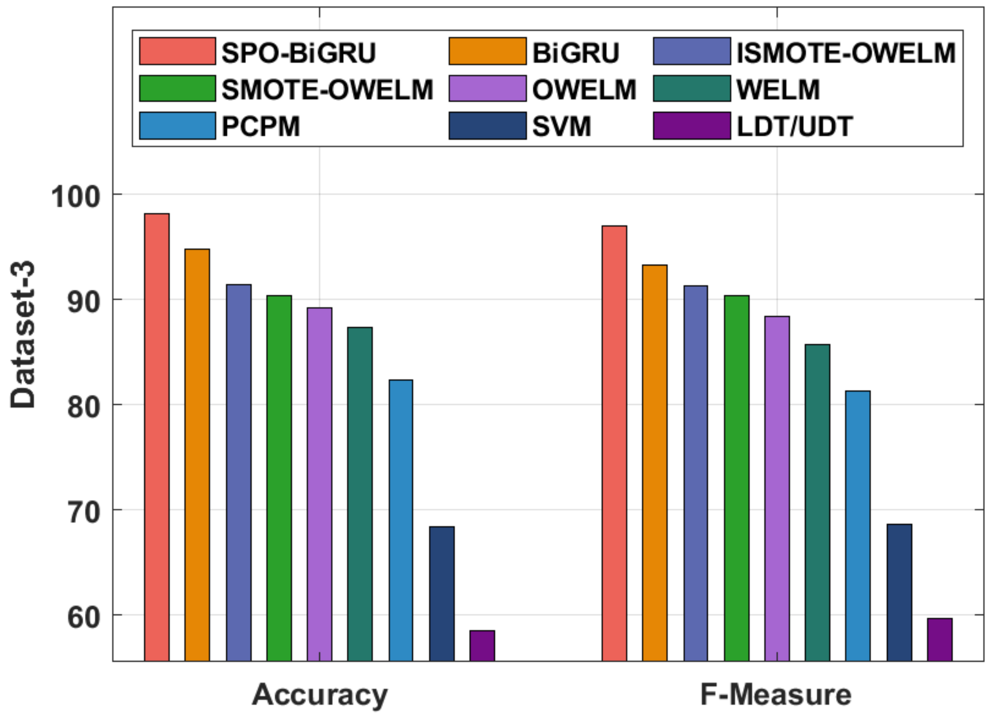
DOI: https://doi.org/10.54216/AJBOR.060103
Received: May 15, 2021 Accepted: November 04, 2021

In regard to F-measure, the SVM, PCPM, LDT/UDT, and WELM techniques have been obtained to lower F-measure values of 73%, 83.3%, 67.50%, and 87.70% correspondingly. Also, the OWELM, SMOTE-OWELM, and ISMOTE-OWELM techniques have gained moderate F-measure of 89.9%, 92.2%, and 92.3% respectively. Afterward, the BiGRU model has resulted in a near optimal F-measure of 96.2%, the projected SPO-BiGRU technique has accomplished improved F-measure of 96.7%.



Figuer 6. ROC analysis of SPO-BiGRU technique under dataset-2

Fig. 6 exhibits the ROC analysis of the SPO-BiGRU technique under dataset-2. The figure outperformed that the SPO-BiGRU technique has resulted in a superior ROC of 99.7579.

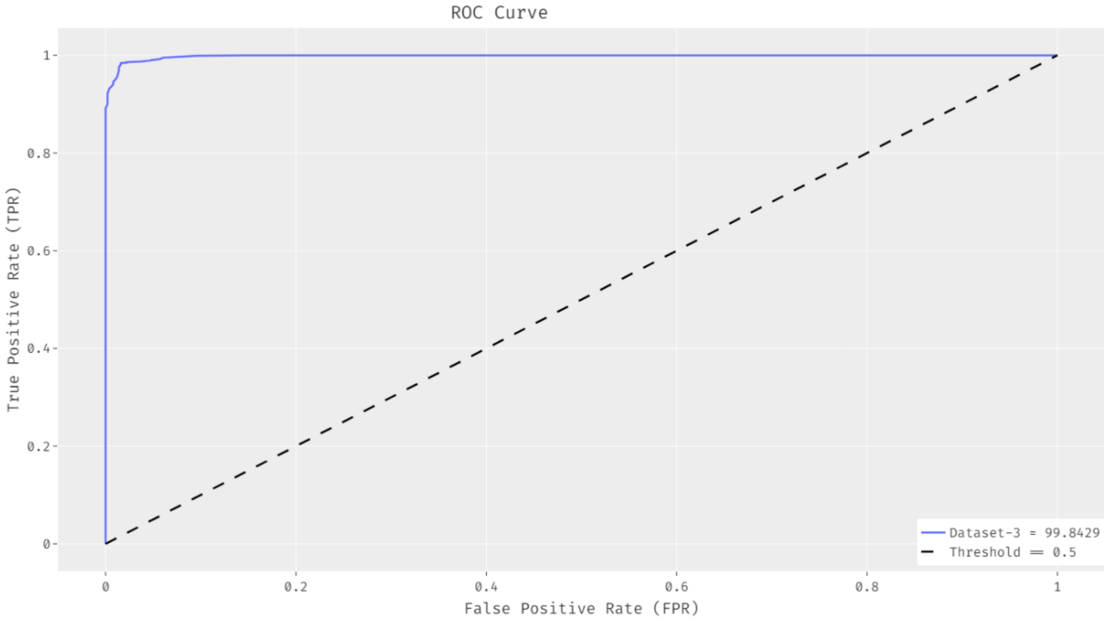


Figuer 7. Result analysis of SPO-BiGRU model under dataset-3

DOI: https://doi.org/10.54216/AJBOR.060103
Received: May 15, 2021 Accepted: November 04, 2021

Fig. 7 investigates the accuracy and F-measure analysis of the SPO-BiGRU model on the test dataset-3. With respect to accuracy, the SVM, PCPM, LDT/UDT, and WELM techniques have reached lesser accuracy values of 68.4%, 82.3%, 58.5%, and 87.4% respectively. In addition, the OWELM, SMOTE-OWELM, and ISMOTE-OWELM techniques have gained moderate accuracy of 89.2%, 90.4%, and 91.4% correspondingly. Though the BiGRU technique has resulted in a near optimal accuracy of 94.8%, the proposed SPO-BiGRU algorithm has accomplished maximum accuracy of 98.1%.

With respect to F-measure, the SVM, PCPM, LDT/UDT, and WELM methods have achieved lower F-measure values of 68.7%, 81.3%, 59.7%, and 85.7% correspondingly. Along with that, the OWELM, SMOTE-OWELM, and ISMOTE-OWELM manners have gained moderate F-measure of 88.4%, 90.4%, and 91.3% respectively. However, the BiGRU model has resulted in a near optimal F-measure of 93.3%, the projected SPO-BiGRU technique has accomplished improved F-measure of 97%.



Figuer 8. ROC analysis of SPO-BiGRU technique under dataset-3

Fig. 8 showcases the ROC analysis of the SPO-BiGRU technique under dataset-3. The figure outperformed that the SPO-BiGRU approach has resulted in a maximal ROC 99.8429. From the above mentioned results and discussion, it is evident that the SPO-BiGRU model has accomplished improved outcome over the other techniques on the applied three dataset. the BiGRU technique is applied to perform the predictive process. It is due to the fact that the SPO algorithm is applied to optimally adjust the hyperparameters involved in the BiGRU model. In order to validate the enhanced performance of the SPO-BiGRU model, a wide range of simulations take place and the results are inspected under varying aspects.

5. Conclusion

This paper has presented an effective SPO-BiGRU technique for CCP on BI applications. The SPO-BiGRU model aimed for determining the occurrence of customers into churner or non-churner. Moreover, the SPO-BiGRU technique involves preprocessing, classification, and hyperparameter optimization. Furthermore, the BiGRU technique is applied to perform the predictive process. Finally, the SPO algorithm is applied to optimally adjust the hyperparameters involved in the BiGRU model. In order to validate the enhanced performance of the SPO-BiGRU model, a wide range of simulations take place and the results are inspected under varying aspects. The experimental results portrayed the supremacy of the SPO-BiGRU technique over the recent state of art approaches. In future, the predictive performance can be further boosted by the use of clustering approaches.

DOI: https://doi.org/10.54216/AJBOR.060103
Received: May 15, 2021 Accepted: November 04, 2021

References

[1] Ahmad, A.K., Jafar, A. and Aljoumaa, K., 2019. Customer churn prediction in telecom using machine learning in big data platform. Journal of Big Data, 6(1), pp.1-24.
[2] Amin, A., Al-Obeidat, F., Shah, B., Adnan, A., Loo, J. and Anwar, S., 2019. Customer churn prediction in telecommunication industry using data certainty. Journal of Business Research, 94, pp.290-301.
[3] Verbeke, W., Martens, D., Mues, C. and Baesens, B., 2011. Building comprehensible customer churn prediction models with advanced rule induction techniques. Expert systems with applications, 38(3), pp.2354-2364.
[4] Burez, J. and Van den Poel, D., 2009. Handling class imbalance in customer churn prediction. Expert Systems with Applications, 36(3), pp.4626-4636.
[5] Huang, B., Kechadi, M.T. and Buckley, B., 2012. Customer churn prediction in telecommunications. Expert Systems with Applications, 39(1), pp.1414-1425.
[6] Lu, N., Lin, H., Lu, J. and Zhang, G., 2012. A customer churn prediction model in telecom industry using boosting. IEEE Transactions on Industrial Informatics, 10(2), pp.1659-1665.
[7] Stripling, E., vanden Broucke, S., Antonio, K., Baesens, B. and Snoeck, M., 2018. Profit maximizing logistic model for customer churn prediction using genetic algorithms. Swarm and Evolutionary Computation, 40, pp.116-130.
[8] De Caigny, A., Coussement, K., De Bock, K.W. and Lessmann, S., 2020. Incorporating textual information in customer churn prediction models based on a convolutional neural network. International Journal of Forecasting, 36(4), pp.1563-1578.
[9] Vafeiadis, T., Diamantaras, K.I., Sarigiannidis, G. and Chatzisavvas, K.C., 2015. A comparison of machine learning techniques for customer churn prediction. Simulation Modelling Practice and Theory, 55, pp.1-9.
[10] Coussement, K., Lessmann, S. and Verstraeten, G., 2017. A comparative analysis of data preparation algorithms for customer churn prediction: A case study in the telecommunication industry. Decision Support Systems, 95, pp.27-36.
[11] Tsai, C.F. and Chen, M.Y., 2010. Variable selection by association rules for customer churn prediction of multimedia on demand. Expert Systems with Applications, 37(3), pp.2006-2015.
[12] Amin, A., Anwar, S., Adnan, A., Nawaz, M., Alawfi, K., Hussain, A. and Huang, K., 2017. Customer churn prediction in the telecommunication sector using a rough set approach. Neurocomputing, 237, pp.242-254.
[13] De Bock, K.W. and Van den Poel, D., 2011. An empirical evaluation of rotation-based ensemble classifiers for customer churn prediction. Expert Systems with Applications, 38(10), pp.12293-12301.
[14] Shirazi, F. and Mohammadi, M., 2019. A big data analytics model for customer churn prediction in the retiree segment. International Journal of Information Management, 48, pp.238-253.
[15] De Caigny, A., Coussement, K. and De Bock, K.W., 2018. A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. European Journal of Operational Research, 269(2), pp.760-772.
[16] Calzada-Infante, L., Óskarsdóttir, M. and Baesens, B., 2020. Evaluation of customer behavior with temporal centrality metrics for churn prediction of prepaid contracts. Expert Systems with Applications, 160, p.113553.
[17] Vo, N.N., Liu, S., Li, X. and Xu, G., 2021. Leveraging unstructured call log data for customer churn prediction. Knowledge-Based Systems, 212, p.106586.
[18] Abdelgwad, M.M., Soliman, T.H.A., Taloba, A.I. and Farghaly, M.F., 2021. Arabic aspect based sentiment analysis using bidirectional GRU based models. arXiv preprint arXiv:2101.10539.
[19] Metan, J., Prasad, A.Y., Kumar, K.A., Mathapati, M. and Patil, K.K., 2021. Cardiovascular MRI image analysis by using the bio inspired (sand piper optimized) fully deep convolutional network (Bio-FDCN) architecture for an automated detection of cardiac disorders. Biomedical Signal Processing and Control, 70, p.103002.
[20] Pustokhina, I.V., Pustokhin, D.A., Aswathy, R.H., Jayasankar, T., Jeyalakshmi, C., Díaz, V.G. and Shankar, K., 2021. Dynamic customer churn prediction strategy for business intelligence using text analytics with evolutionary optimization algorithms. Information Processing & Management, 58(6), p.102706.

DOI: https://doi.org/10.54216/AJBOR.060103
Received: May 15, 2021 Accepted: November 04, 2021