



Intelligent Model for Customer Churn Prediction using Deep Learning Optimization Algorithms

Abedallah Zaid Abualkishik ^{*1}, Rasha Almajed ², William Thompson³

^{1,2} American University in the Emirates, Dubai, UAE

³ Towson University, Towson University, Maryland's University, USA

Emails: abedallah.abualkishik@ae.ac; rasha.almajed@ae.ac; wvthompson@towson.edu

Abstract

Business intelligence (BI) mentions to the technical and procedural structure which gathers, supplies, and examines the data formed by company action. BI is a wide term that includes descriptive analytics, procedure analysis, data mining, and performance benchmarking. Customer churn is a general problem across businesses from several sectors. Companies are working always for improving their supposed quality by way of providing timely and quality service to its customer. Customer churn is developed most initial challenges which several firms were facing currently. Many churn prediction techniques and methods were presented before in literature for predicting customer churn from the domains like telecom, finance, banking, and so on. Researchers are also working on customer churn prediction (CCP) from e-commerce utilizing data mining and machine learning (ML) approaches. This manuscript focuses on the development of Stacked Deep Learning with Wind Driven Optimization based Business Intelligence for Customer Churn Prediction model. The proposed model is considered an intelligent system that applies golden sine algorithm (GSA) based feature selection approach to derive a set of features. In addition, the stacked gated recurrent unit (SGRU) model is applied for the prediction of customer churns.

Keywords: Business intelligence; Customer churn prediction; Deep learning; Feature selection; Machine learning; Intelligent Systems

1. Introduction

Customer churn (otherwise called customer attrition) means when a customer stops their relationship with a company [1]. Online businesses normally treat a customer as churned once a specific measure of time has slipped as the customer's last collaboration with administration or webpage. Decreasing churn is the main business goal of each and every web-based business. The capability for predicting that a particular customer is at a higher risk of churning, right now to take care of business, addresses a huge more potential income hotspot for each internet-based business [2]. Other than the immediate loss of income that outcomes from a customer abandoning the business, the expenses of at first gaining that customer might not have proactively been enclosed by the customer's spending to date [3]. For prevailing at hold customers who might somehow abandon the business, marketers and maintenance specialists should have the option to (a) foresee ahead of time which customers will churn through churn analysis and (b) know which marketing activities will affect every specific customer [4]. Outfitted with this knowledge, a huge extent of customer churn can be wiped out.

Churn was categorized in business terminologies as 'when a client drops a membership to help they were using [5]. A typical model is individuals dropping Spotify/Netflix memberships. In this way, Churn Prediction was usually foreseeing which clients were perhaps going to drop a membership that is 'leave

an organization based on their utilization of the help. Thus, the bits of knowledge acquired from Churn Prediction assist them with focusing more on the customers that are at a high risk of leaving [6]. Churn prediction displaying techniques endeavor for recognizing the exact customer ways of behaving and credits which signal the risk and timing of customer churn. All things considered, in the event that the marketer knows nothing about a customer going to churn, no move will be made for that customer [7]. Furthermore, extraordinary maintenance centered provides or impetuses may be coincidentally given to blissful, dynamic customers, bringing about diminished incomes inexplicably.

Tragically, the majority of the churn prediction demonstrating strategies depend on quantifying risk dependent upon static data and measurements. The most widely recognized churn prediction techniques are dependent upon more established factual and data-mining approaches, like logistic regression and other double demonstrating systems [8]. Such methods offer some worth and could distinguish a specific level of at-risk customers. An exact churn prescient could efficiently aid client holding plan and financial marketing developments, and, consequently, it could bring about significant reserve funds for the providers [9]. To withstand this forceful rivalry, telecommunication business turns out to be more successful by putting further improvement in data mining and ML based techniques for churn prediction, the executives, and examinations. A few ML techniques have been introduced in the review for anticipating churn. Learning from imbalanced class conveyance is exceptionally confounding for all conventional ML techniques while it can be slope to arrange the exact larger part class and dismiss the interesting one [10]. One of the key difficulties is that there are a few viewpoints impacting the client churn and the relations amongst them are profoundly muddled.

This manuscript focuses on the development of Stacked Deep Learning with Wind Driven Optimization based Business Intelligence for Customer Churn Prediction (SDLWDO-BICCP) model. The presented SDLWDO-BICCP model mainly concentrates on the identification of churn customers in the business environment. The SDLWDO-BICCP model applies golden sine algorithm (GSA) based feature selection (FS) approach to derive a set of features. In addition, the stacked gated recurrent unit (SGRU) model is applied for the prediction of customer churns. At last, the WDO algorithm is utilized for effectually modifying the hyperparameters involved in the SGRU model. To ensure the enhanced performance of the SDLWDO-BICCP model, a wide range of experiments are performed using benchmark dataset.

2. Literature Review

In [11], we propose an improved prediction method by coordinating ensemble learning and data to pre-processing to take care of these two issues. In particular, two novel features are first coordinated into the RFM analysis to more likely catch customers' ways of behaving. Second, the PCA is embraced to diminish data aspects. Third, adaptive boosting (AdaBoost) is utilized to flow various DTs to limit the effects of unequal data. Ramesh et al. [12] present successful answers for this large number of testing issues in Customer Churn Prediction (CCP). The review utilizes datasets in the telecommunication business, the ANN, and RF to determine the elements that impact purchaser churn.

Pustokhina et al. [13] investigation work plan a unique CCP methodology for business intelligence using text analytics with metaheuristic optimization (CCPBI-TAMO) process. Moreover, CPIO-FS method is utilized for the FS procedure and diminishes calculation intricacy. Additionally, LSTM with SAE method was implied to order the feature diminished data. In the LSTM-SAE method, the capacity of SAE in the identification of conservative features is coordinated with the classification ability of the LSTM method. At long last, the sunflower optimization (SFO) hyperparameter tuning procedure takes spot to additionally further develop the CCP execution. Rahmaty et al. [14] present a hybrid framework dependent upon fuzzy entropy criterion selection algorithm, and ANN to foresee the customer churn of those organizations which experience the ill effects of losing customers over time.

Xiahou and Harada [15] suggest a loss prediction method based on the blend of SVMtiu and k-means customer segmentation prediction. The technique separates customers into 3 classes and defines the center of customer gatherings. Perišić and Pahor [16] center around construction a churn prediction method for the versatile gaming market by using LR analysis in the drawn-out recency, frequency, and monetary (RFM) framework. Laid out features are utilized to foster a logistic system for churn prediction and to characterize likely churners in a populace of clients, no matter what their lifetime. Tariq et al. [17]

presented method purposes the 2D CNN (CNN; a technique of DL). The suggested method is layered engineering that involves 2 unique stages that are 2D CNN layer and pre-processing layer and data load. Likewise, the circulated framework and Apache Spark parallel were utilized to handle the data in parallel conditions.

3. The Proposed Model

This manuscript has developed a new SDLWDO-BICCP model for the identification of churn customers in the business environment. The SDLWDO-BICCP model applied GSA based FS approach to derive a set of features. Also, the SGRU model is applied for the prediction of customer churns. At last, the WDO algorithm is utilized for effectually modifying the hyperparameters involved in the SGRU model. Fig. 1 showcases the overall process of SDLWDO-BICCP algorithm.

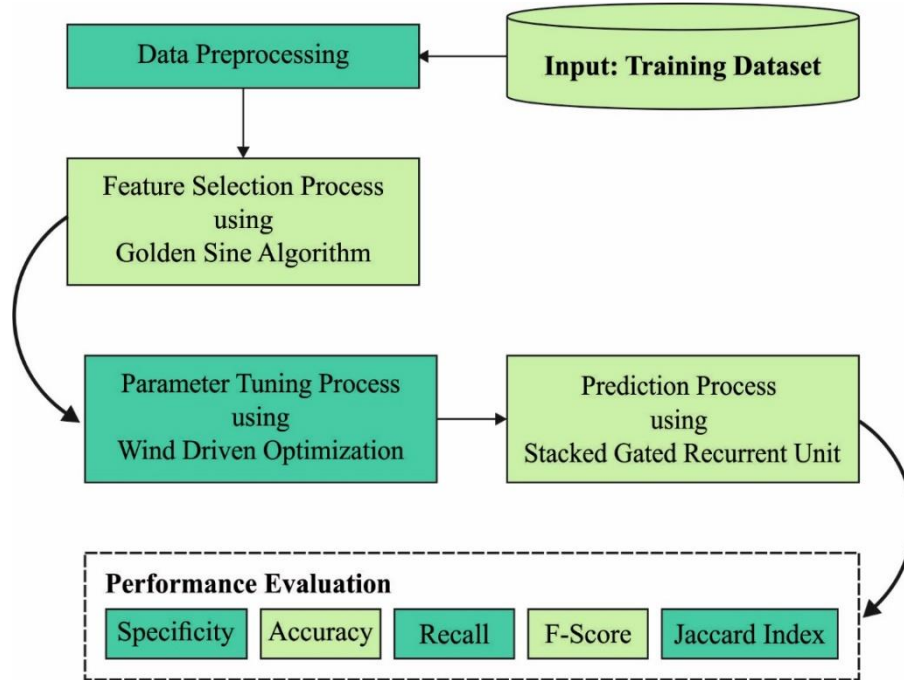


Figure 1: Overall process of SDLWDO-BICCP algorithm

3.1 GSA based Feature Selection Process

Primarily, the SDLWDO-BICCP model applied GSA based FS approach to derive a set of features. The motivation of the presented method is a sine function in arithmetic, and the individual explores the estimated optimum solution from the searching region based on the golden ratio [18]. The sine function ranges from $[-1, 1]$, with interval 2π . Once the x_1 value varies, the respective parameter y_1 varies. Merging the golden ratio and sine function assists in endlessly decreasing the searching region and searching in region whereby the optimum value is probable to be produced, thus enlightening the convergence rate. The computation equation is expressed by:

$$X_{i,j}(i+1) = X_{i,j}(t) \times |\sin(p_1)| - p_2 \times \sin(p_1) \times |d_1 \times X_{best,j}(t) - d_2 \times X_{i,j}(t)| \quad (1)$$

In Eq. (1), p_1 denotes the arbitrary integer lies within $[0, 2\pi]$, and p_2 indicates the arbitrary number lies within $[0, \pi]$, and d_1 and d_2 denotes the coefficient factor that is attained as follows:

$$d_1 = a \times \tau + b \times (1 - \tau) \quad (2)$$

$$d_2 = a \times (1 - \tau) + b \times \tau \quad (3)$$

From the expression, a and b denote the initial value that is $-\pi$ and π . τ represents the golden ratio, $(\sqrt{5} - 1)/2$. The pseudocode of GSA is demonstrated in Algorithm 1.

Algorithm 1: Pseudocode of GSA
Input: The variable of GSA namely the amount of searching agent (N) iteration (T).
Output: The optimal solution
Initialization of the searching agent in a random manner.
While ($t < T$) do
Check whether searching agent goes beyond the searching space and adjust it
Evaluate fitness for the provided searching agent.
For $i = 1$ to N do
Upgrade the arbitrary values p_1 and p_2 , correspondingly.
For $j = 1$ to D do
Upgrade location of searching agent using the Eq. (1).
End for
End for
$t = t + 1$.
End while

3.2 SGRU based Prediction Process

In this study, the SGRU model is applied for the prediction of customer churns. SGRU is composed of different GRU units [19]. For t time sequences, the input sequences $\{e_1, e_2, \dots, e_t\}$ firstly enter into hidden states $\{h_1^1, h_2^1, \dots, h_t^1\}$ for accomplishing every dataset from the preceding time steps. Following, the upper hidden states take the output from the low hidden layer at a comparable time step as the input to feature extraction. Especially, the upper layer of hidden layers are $\{h_1^2, h_2^2, \dots, h_t^2\}$. For every layer, h_t^i hidden state, as shown in Eq. (7), is demonstrated as Eqs. (4)–(6) to accomplish the reset gate, candidate value, and update gate. It noticed that in Eqs. (4)–(6), we added embedding vector e_t in the primary layer. Fig. 2 demonstrates the infrastructure of GRU technique. Beginning from the following layer upward, we employed the hidden layer from the existing time step in the preceding layer, h_t^{i-1} , before e_t in (4)–(6).

$$u_t^i = \sigma(W_u^i h_{t-1}^i + U_u^i e_t + b_u^i) \quad (4)$$

$$r_t^i = \sigma(W_r^i h_{t-1}^i + U_r^i e_t + b_r^i) \quad (5)$$

$$\tilde{c} = \tanh(W_c^i \cdot [r_t^i * h_{t-1}^i] + U_c^i e_t + b_c^i) \quad (6)$$

$$h_t^i = u_t^i * \tilde{c} + (1 - u_t^i) * h_{t-1}^i \quad (7)$$

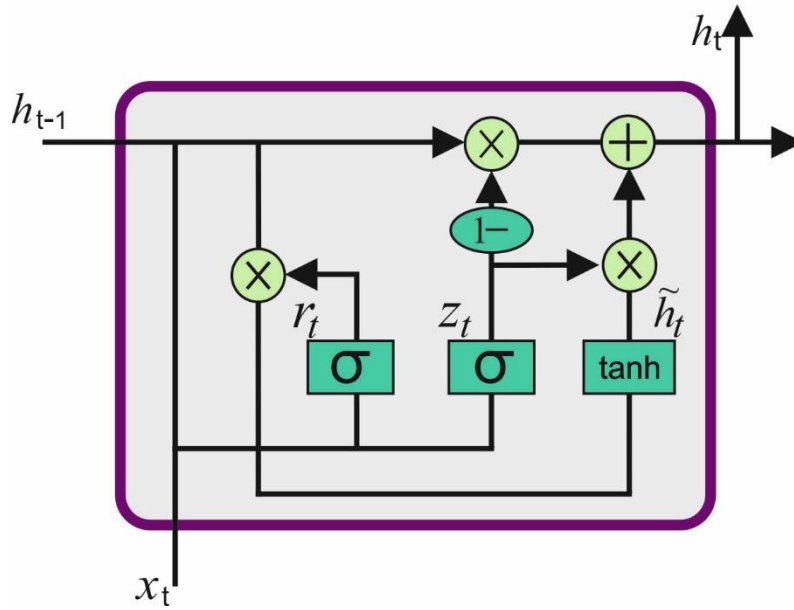


Figure 2: Framework of GRU

3.3 WDO based Parameter Optimization Process

At last, the WDO algorithm is utilized for effectually modifying the hyperparameters involved in the SGRU model. It is stimulated by the natural wind motion conception which acts as a stabilizer to equalize the air pressure inequality. Wind blow from a higher-to lower-pressure region, alongside a velocity that is directly proportionate to pressure gradient (the stronger the wind blows, the higher the pressure variance). The primary idea behindhand the presented technique is Newton's second law of motion:

$$\rho \vec{a} = \sum \vec{F}_i \quad (8)$$

Here, the acceleration vector is represented as a , the air density for a mechanism with small count is indicated by ρ , and the force acting on the masses is represented as F_j . Density, Temperature, and Air pressure are correlated with the equation as follows

$$P = \rho RT \quad (9)$$

Now, T , P , and R are indicated as temperature, pressure, and universal gas constant.

From the Eq. (8), the significant force that creates the wind blow in a direction for diverging the direction is categorized as gravitational force (F), frictional force (F_F), pressure gradient force (F_{PG}), and coriolis force (F_C).

$$\vec{F}_{PG} = -\nabla P \delta V \quad (10)$$

$$\vec{F}_C = -2\Omega \times \vec{u} \quad (11)$$

$$\vec{F}_G = \rho \delta \vec{g} \quad (12)$$

$$\vec{F}_F = -\rho \alpha \vec{u}, \quad (13)$$

Consider, g denotes the gravitational acceleration, the pressure gradient is characterized by ∇P , δV specifies a smaller volume of air, Ω indicates the revolution of earth, and u symbolizes the wind velocity vector. The combination of Eqs. (8) and (10)–(13) generate a formulation shown below:

$$\rho \vec{u} \Delta t = -\nabla P \delta V - 2\Omega \times \vec{u} + \rho \delta V \vec{g} - \rho \alpha \vec{u}. \quad (14)$$

The unit step of time, Δt is correspondent to 1. The combination of Eqs. (9) and (14) generate a formulation shown below:

$$\vec{u}_{new} = -g\vec{x}_{old} + (1 - \alpha)\vec{u}_{old} + \left|1 - \frac{P_{max}}{P_{old}}\right| RT(x_{max} - x_{old}) - \frac{cu_{old}^{other\ dim}}{P_{old}} \quad (15)$$

Now, u_{new} and u_{old} are the upgraded and existing velocity; x_{old} and x_{max} specifies the present place and the maximal pressure place of the air pack; P_{max} and P_{old} shows the maximal pressure and the pressure at the present location; T denotes the temperature; and R , α , and c indicate constant.

In Eq. (15) the pressure value is usually high. As a result, the velocity approximation is greater. This creates the efficacy level of the WDO to decrease.

$$\vec{u}_{new} = -g\vec{x}_{old} + (1 - \alpha)\vec{u}_{old} + \left|1 - \frac{1}{k}\right| RT(x_{max} - x_{old}) - \frac{cu_{old}^{other\ dim}}{k}. \quad (16)$$

From Eq. (16), k suggests the ranking amongst every air parcel ($k = 1, 2, \dots, 20$). The succeeding formulation is exploited to upgrade the place of air pack:

$$\vec{x}_{new} = \vec{x}_{old} + \vec{u}_{new} \times \Delta t. \quad (17)$$

The WDO grows a fitness function (FF) for accomplishing high classifier effectiveness. It determines a positive integer to signify the optimum efficacy of solution candidate. Here, the minimalized classifier error rate was considered as FF as prevailing.

$$\begin{aligned} fitness(x_i) &= ClassifierErrorRate(x_i) \\ &= \frac{\text{number of misclassified samples}}{\text{Total number of samples}} \times 100 \end{aligned} \quad (18)$$

4. Results and Discussion

The experimental assessment of the SDLWDO-BICCP model is tested using a churn dataset, comprising 3333 samples under two classes. A total of 2850 samples exist under churn class and 2850 samples exist under non-churn class as depicted in Table 1.

Table 1: Dataset details

Class	No. of Samples
Non Churn	2850
Churn	483
Total	3333

The confusion matrices created by the SDLWDO-BICCP model on the test data are demonstrated in Fig. 3. On 70% of TR data, the SDLWDO-BICCP model has categorized 1985 samples into non-churn and 330 samples into churn. Also, on 30% of TS data, the SDLWDO-BICCP method has categorized 850 samples into non-churn and 143 samples into churn. In addition, on 80% of TR data, the SDLWDO-BICCP technique has categorized 2258 samples into non-churn and 357 samples into churn. At last, on 20% of TS data, the SDLWDO-BICCP algorithm has categorized 560 samples into non-churn and 89 samples into churn.

Table 2 offers comprehensive classification results of the SDLWDO-BICCP model on 70% of TR data and 30% of TS data.

Fig. 4 indicates the classifier results of the SDLWDO-BICCP model on 70% of TR data. The SDLWDO-BICCP model has recognized the samples under non-churn class with $accu_y$ of 99.23%, $reca_l$ of

99.50%, $spec_y$ of 97.63%, F_{score} of 99.55%, and $Jaccard_{index}$ of 99.10%. Also, the SDLWDO-BICCP approach has recognized the samples under churn class with $accu_y$ of 99.23%, $reca_t$ of 97.63%, $spec_y$ of 99.50%, F_{score} of 97.35%, and $Jaccard_{index}$ of 94.83%.

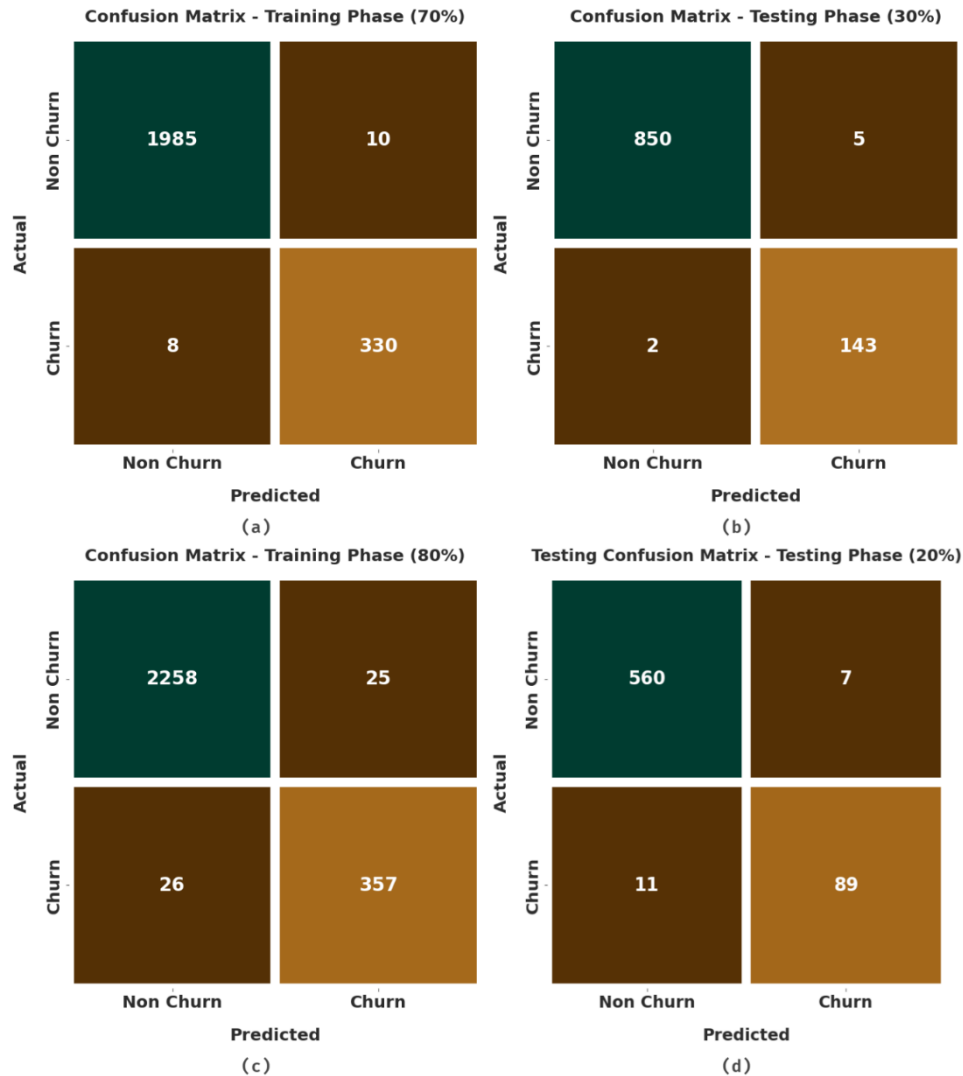


Figure 3: Confusion matrices of SDLWDO-BICCP approach (a) 70% of TR data, (b) 30% of TS data, (c) 80% of TR data, and (d) 20% of TS data

Table 2: Result analysis of SDLWDO-BICCP approach with various measures under 70% of TR and 30% of TS data

Labels	Accuracy	Recall	Specificity	F-Score	Jaccard Index
Training Phase (70%)					
Non Churn	99.23	99.50	97.63	99.55	99.10
Churn	99.23	97.63	99.50	97.35	94.83
Average	99.23	98.57	98.57	98.45	96.96
Testing Phase (30%)					
Non Churn	99.30	99.42	98.62	99.59	99.18
Churn	99.30	98.62	99.42	97.61	95.33
Average	99.30	99.02	99.02	98.60	97.26

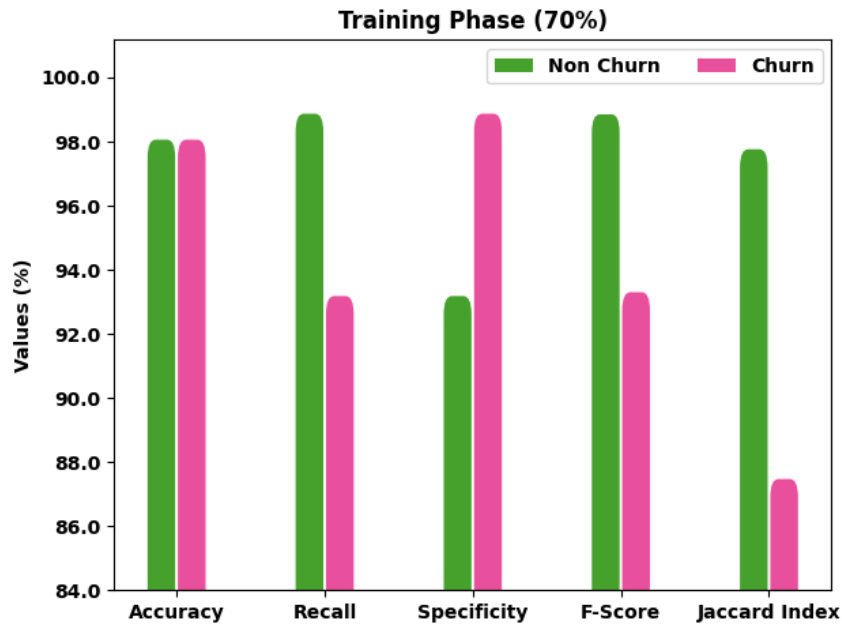


Figure 4: Result analysis of SDLWDO-BICCP approach under 70% of TR data

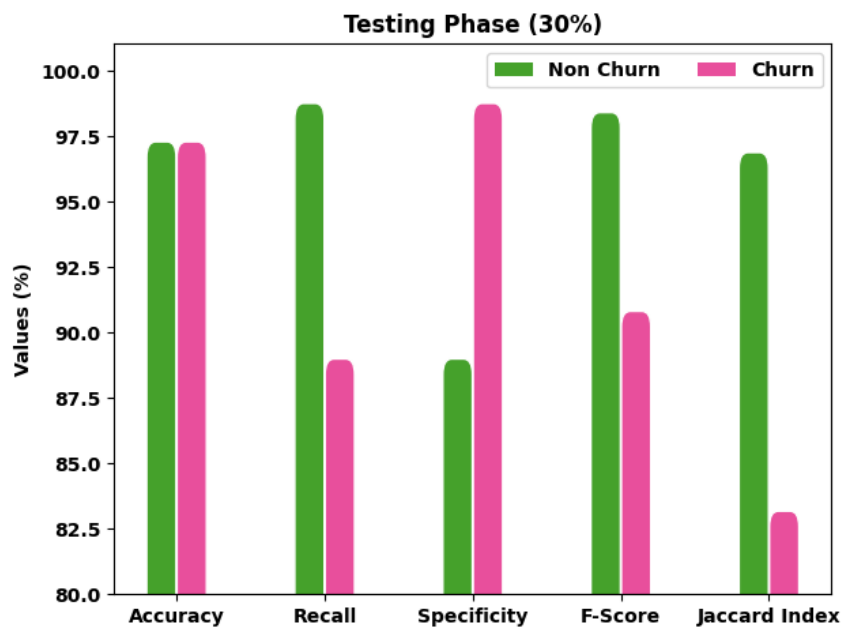


Figure 5: Result analysis of SDLWDO-BICCP approach under 30% of TS data

Fig. 5 represents the classifier results of the SDLWDO-BICCP method on 30% of TS data. The SDLWDO-BICCP technique has recognized the samples under non-churn class with $accu_y$ of 99.30%, $reca_l$ of 99.42%, $spec_y$ of 98.62%, F_{score} of 99.59%, and $Jaccard_{index}$ of 99.18%. Additionally, the SDLWDO-BICCP approach has recognized the samples under churn class with $accu_y$ of 99.30%, $reca_l$ of 98.62%, $spec_y$ of 99.42%, F_{score} of 97.61%, and $Jaccard_{index}$ of 95.33%.

Table 3 grants detailed classification results of the SDLWDO-BICCP approach on 80% of TR data and 20% of TS data. Fig. 6 denotes the classifier results of the SDLWDO-BICCP approach on 80% of TR data. The SDLWDO-BICCP technique has recognized the samples under non-churn class with $accu_y$ of

98.09%, $reca_l$ of 98.90%, $spec_y$ of 93.21%, F_{score} of 98.88%, and $Jaccard_{index}$ of 97.79%. Also, the SDLWDO-BICCP algorithm has recognized the samples under churn class with $accu_y$ of 98.09%, $reca_l$ of 93.21%, $spec_y$ of 98.90%, F_{score} of 93.33%, and $Jaccard_{index}$ of 87.50%.

Table 3: Result analysis of SDLWDO-BICCP approach with various measures under 80% of TR and 20% of TS data

Labels	Accuracy	Recall	Specificity	F-Score	Jaccard Index
Training Phase (80%)					
Non Churn	98.09	98.90	93.21	98.88	97.79
Churn	98.09	93.21	98.90	93.33	87.50
Average	98.09	96.06	96.06	96.11	92.65
Testing Phase (20%)					
Non Churn	97.30	98.77	89.00	98.42	96.89
Churn	97.30	89.00	98.77	90.82	83.18
Average	97.30	93.88	93.88	94.62	90.03

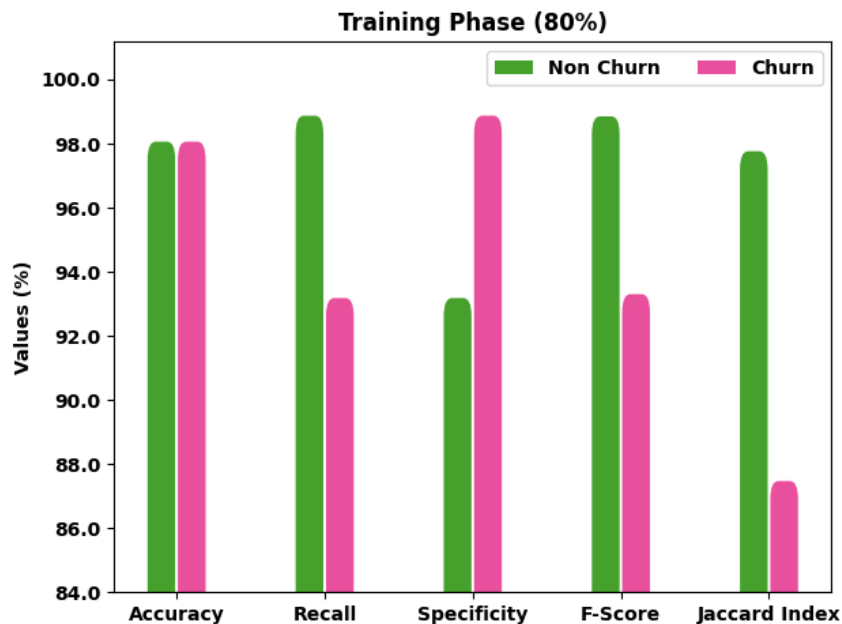


Figure 6: Result analysis of SDLWDO-BICCP approach under 80% of TR data

Fig. 7 demonstrates the classifier results of the SDLWDO-BICCP algorithm on 20% of TS data. The SDLWDO-BICCP technique has recognized the samples under non-churn class with $accu_y$ of 97.30%, $reca_l$ of 98.77%, $spec_y$ of 89%, F_{score} of 98.42%, and $Jaccard_{index}$ of 96.89%. along with that, the SDLWDO-BICCP algorithm has recognized the samples under churn class with $accu_y$ of 97.30%, $reca_l$ of 89%, $spec_y$ of 98.77%, F_{score} of 90.82%, and $Jaccard_{index}$ of 83.18%.

A comparative $accu_y$ inspection is made between the SDLWDO-BICCP model and existing models in Fig. 8 [13]. The figure indicated that the SVM model has shown lower $accu_y$ of 78.90% whereas the PCPM model has resulted to slightly enhanced $accu_y$ of 83.70%. Followed by, the OWELM model has shown moderate $accu_y$ value of 90.60%. Then, the CCPBI-TAMO and LSTM-SAE models have demonstrated reasonable $accu_y$ values of 95.56% and 94.11% respectively. However, the SDLWDO-BICCP model has accomplished maximum $accu_y$ of 99.30%.

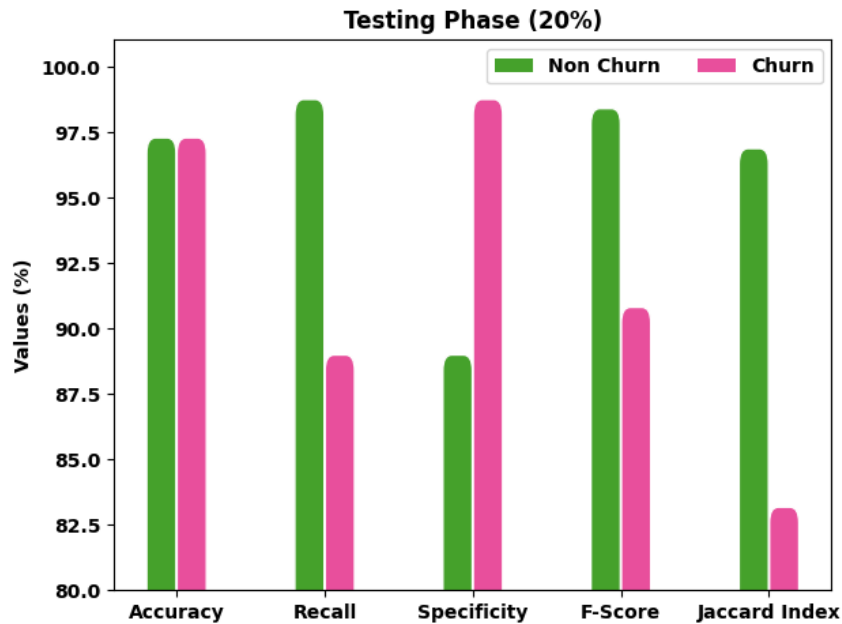


Figure 7: Result analysis of SDLWDO-BICCP approach under 20% of TS data

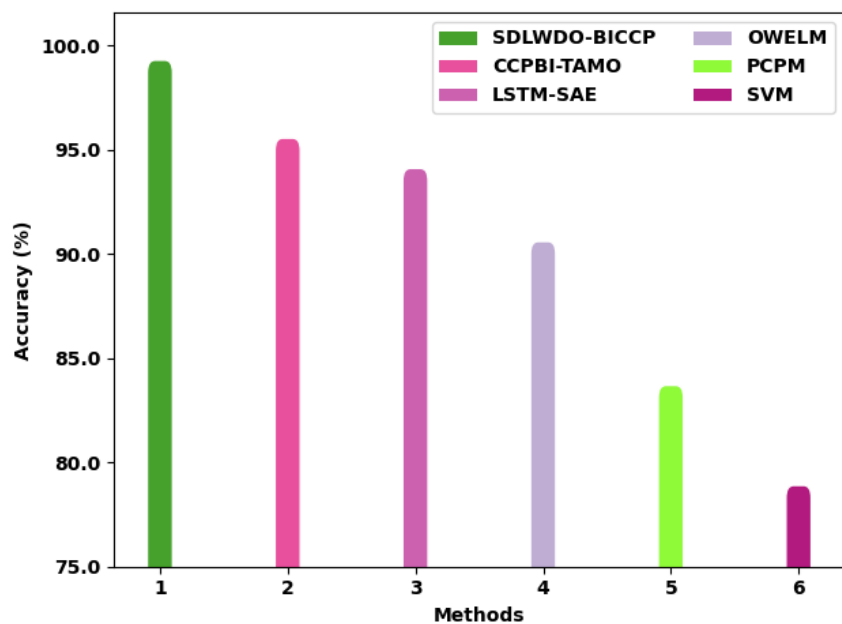


Figure 8: $Accu_y$ analysis of SDLWDO-BICCP approach with existing methodologies

A detailed F_{score} analysis is made among the SDLWDO-BICCP method and existing models in Fig. 9. The figure representing the SVM technique has shown lower F_{score} of 76.30% whereas the PCPM algorithm has resulted to slightly enhanced F_{score} of 83.80%. Subsequently, the OWELM methodology has shown moderate F_{score} value of 90.40%. And, the CCPBI-TAMO and LSTM-SAE approaches have demonstrated reasonable F_{score} values of 94.96% and 93.62% correspondingly. But, the SDLWDO-BICCP method has accomplished maximal F_{score} of 98.60%.

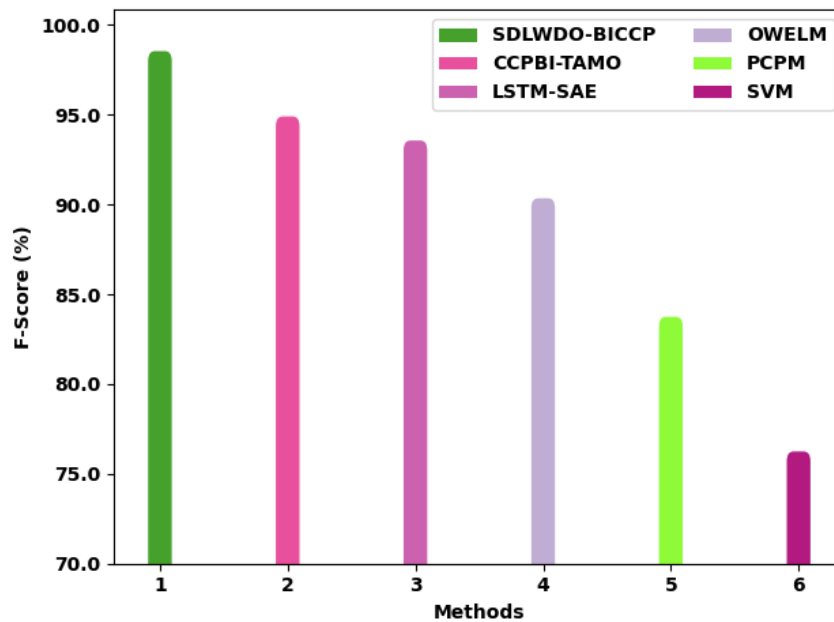


Figure 9: F_{score} analysis of SDLWDO-BICCP approach with existing methodologies

5. Conclusion

This paper proposed an intelligent system for the identification of churn customers in the business environment. The proposed system applied GSA based FS approach to derive a set of features. Additionally, the SGRU method was applied for the prediction of customer churns. At last, the WDO algorithm is utilized for effectually modifying the hyperparameters involved in the SGRU model. To ensure the enhanced performance of the proposed model, a wide range of experiments are performed using benchmark dataset.

References

- [1] Shakir, T.K. and Al Masri, A.N. 2021. Intelligent Optimal Deep Learning based Customer Churn Prediction Model in Telecom Industry. American Journal of Business and Operations Research, 4(2), pp. 57-64. doi: 10.54216/AJBOR.040202.
- [2] Rai, S., Khandelwal, N. and Boghey, R., 2020. Analysis of customer churn prediction in telecom sector using cart algorithm. In First International Conference on Sustainable Technologies for Computational Intelligence (pp. 457-466). Springer, Singapore.
- [3] Thakkar, H.K., Desai, A., Ghosh, S., Singh, P. and Sharma, G., 2022. Clairvoyant: AdaBoost with Cost-Enabled Cost-Sensitive Classifier for Customer Churn Prediction. Computational Intelligence and Neuroscience, 2022.
- [4] Nwaogu, V.C. and Dimililer, K., 2021, June. Customer Churn Prediction For Business Intelligence Using Machine Learning. In 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA) (pp. 1-7). IEEE.
- [5] 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.
- [6] Bandy, S.A. and Khan, S., 2021. Evaluation Study of Churn Prediction Models for Business Intelligence. In Big Data Analytics (pp. 201-213). Auerbach Publications.
- [7] Gursoy, U.F., Yildiz, E.M., Okay, M.E. and Aktas, M.S., 2021, September. Customer Churn Prediction and Promotion Models in the Telecom Sector: A Case Study. In Proceedings of SAI Intelligent Systems Conference (pp. 276-286). Springer, Cham.
- [8] Günesen, S.N., Şen, N., Yıldırım, N. and Kaya, T., 2021, January. Customer Churn Prediction in FMCG Sector Using Machine Learning Applications. In IFIP International Workshop on Artificial Intelligence for Knowledge Management (pp. 82-103). Springer, Cham.

- [9] Dai, M. and Wang, Y., 2020. Research on the application of business intelligence based on data mining technology in the new industry. *J. Simul*, 8(6).
- [10] Çelik, O. and Osmanoglu, U.O., 2019. Comparing to techniques used in customer churn analysis. *Journal of Multidisciplinary Developments*, 4(1), pp.30-38.
- [11] Wu, Z., Jing, L., Wu, B. and Jin, L., 2022. A PCA-AdaBoost model for E-commerce customer churn prediction. *Annals of Operations Research*, pp.1-18
- [12] Ramesh, P., Jeba Emilyn, J. and Vijayakumar, V., 2022. Hybrid Artificial Neural Networks Using Customer Churn Prediction. *Wireless Personal Communications*, 124(2), pp.1695-1709
- [13] 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
- [14] Rahmaty, M., Daneshvar, A., Salahi, F., Ebrahimi, M. and Chobar, A.P., 2022. Customer Churn Modeling via the Grey Wolf Optimizer and Ensemble Neural Networks. *Discrete Dynamics in Nature and Society*, 2022
- [15] Xiahou, X. and Harada, Y., 2022. B2C E-Commerce Customer Churn Prediction Based on K-Means and SVM. *Journal of Theoretical and Applied Electronic Commerce Research*, 17(2), pp.458-475
- [16] Perišić, A. and Pahor, M., 2020. Extended RFM logit model for churn prediction in the mobile gaming market. *Croatian Operational Research Review*, pp.249-261
- [17] Tariq, M.U., Babar, M., Poulin, M. and Khattak, A.S., 2021. Distributed model for customer churn prediction using convolutional neural network. *Journal of Modelling in Management*
- [18] Liu, Q., Li, N., Jia, H., Qi, Q., Abualigah, L. and Liu, Y., 2022. A Hybrid Arithmetic Optimization and Golden Sine Algorithm for Solving Industrial Engineering Design Problems. *Mathematics*, 10(9), p.1567.
- [19] Al Wazrah, A. and Alhumoud, S., 2021. Sentiment Analysis Using Stacked Gated Recurrent Unit for Arabic Tweets. *IEEE Access*, 9, pp.137176-137187.