Computers, Materials & Continua
DOI: 10.32604/cmc.2022.030428
Article

Tech Science Press

Optimal Deep Canonically Correlated Autoencoder-Enabled Prediction Model for Customer Churn Prediction

Olfat M. Mirza[1], G. Jose Moses[2], R. Rajender[3], E. Laxmi Lydia[4], Seifedine Kadry[5], Cheadchai Me-Ead[6] and Orawit Thinnukool[7,*]

[1]Department of Computer Science, College of Computers and Information Systems, Umm Al-Qura University, Makkah, Saudi Arabia
[2]Department of Computer Science Engineering, Malla Reddy Engineering College, 500100, India
[3]Department of Computer Science Engineering, Lendi Institute of Engineering & Technology, Denkada, 535005, India
[4]Department of Computer Science and Engineering, Vignan's Institute of Information Technology, Visakhapatnam, 530049, India
[5]Department of Applied Data Science, Noroff University College, Norway
[6]Program in Statistics and Information Management, Faculty of Science, Maejo University, Chiang Mai, 50290, Thailand
[7]Department of Modern Management and Information Technology, College of Arts Media and Technology, Chiang Mai University, Chiang Mai, 50200, Thailand
*Corresponding Author: Orawit Thinnukool. Email: orawit.t@cmu.ac.th
Received: 25 March 2022; Accepted: 07 May 2022

Abstract: Presently, customer retention is essential for reducing customer churn in telecommunication industry. Customer churn prediction (CCP) is important to predict the possibility of customer retention in the quality of services. Since risks of customer churn also get essential, the rise of machine learning (ML) models can be employed to investigate the characteristics of customer behavior. Besides, deep learning (DL) models help in prediction of the customer behavior based characteristic data. Since the DL models necessitate hyperparameter modelling and effort, the process is difficult for research communities and business people. In this view, this study designs an optimal deep canonically correlated autoencoder based prediction (O-DCCAEP) model for competitive customer dependent application sector. In addition, the O-DCCAEP method purposes for determining the churning nature of the customers. The O-DCCAEP technique encompasses pre-processing, classification, and hyperparameter optimization. Additionally, the DCCAE model is employed to classify the churners or non-churner. Furthermore, the hyperparameter optimization of the DCCAE technique occurs utilizing the deer hunting optimization algorithm (DHOA). The experimental evaluation of the O-DCCAEP technique is carried out against an own dataset and the outcomes highlighted the betterment of the presented O-DCCAEP approach on existing approaches.

Keywords: Churn prediction; customer retention; deep learning; machine learning; archimedes optimization algorithm

1 Introduction

Customer retention plays a significant part in predicting churn, and telecommunication system is straightaway connected to the financial corporations. Henceforth, several businesses executed dissimilar activities for emerging a strong connection with user and to decrease user defection. Recently, the cost cutting stress and competing nature raised the companies to exploit the customer relationship management (CRM) system [1]. Customer Churn prediction (CCP) is assisted to established approaches for customer retention. In consort with rising competitiveness in market for offering facilities, the hazard of customer churn exponentially rises [2]. Then, establishing approaches to follow loyal customers (non-churner) becomes an obligation. The customer churn model aims at identifying earlier [3,4] churn signals and attempting to forecast the customer that leaves willingly. Therefore, several firms have understood that the current dataset is one of the valued advantages [5], and based on Abbasdimehr, churn prediction is a valuable mechanism for predicting customers at risk. For capturing the abovementioned issue, corporation must forecast the customer behavior properly. Customer churn management is performed by: (1) Reactive and (2) Proactive manners. In the reactive method, corporations wait for the termination request established from the user, then, business provides the effective strategies to the user for retention [6,7]. It is a binary classifier issue in which churner is detached from the non-churner. For resolving the challenge, machine learning method has shown an effectual method, for predicting data according to formerly taken information [8,9]. In machine learning (ML) model, afterward, pre-processing feature selection performs a substantial part to enhance the classifier performance. Several techniques are proposed by authors for feature selection that is effective for reducing the overfitting, dimension, and computational difficulty [10].

Pustokhina et al. [11] proposed a dynamic CCP approach to business intelligence utilizing text analytics with metaheuristic optimization (CCPBI-TAMO) technique. Also, the chaotic pigeon inspired optimization based feature selection (CPIO-FS) approach was utilized to FS procedure and decreased computation complexity. Also, the long short term memory (LSTM) with stacked autoencoder (SAE) technique was executed for classifying the feature decreased data. Lastly, the sunflower optimization (SFO) hyper-parameter tuning procedure occurs to enhance the CCP performance. In [12], a new algorithmic extension, spline-rule ensembles (SRE) with sparse group lasso regularization (SRE-SGL) was presented for enhancing interpretability with infrastructure regularized. The experiments on 14 real world customer churn datasets from distinct industries illustrate the higher prediction performance of SRE.

Bilal et al. [13] presented a hybrid method that is dependent upon integration of clustering and classification techniques utilizing an ensemble. Primary, distinct clustering techniques (for instance K-medoids, random clustering, K-means, and X-means) are calculated separately on 2 churn prediction data sets. Next, the hybrid methods are established by integrating the cluster with 7 various classifier approaches separately afterward evaluation was executed utilizing ensembles.

Wu et al. [14] presented an enhanced forecast method by combining data pre-processed and ensemble learning for solving these 2 problems. In specially, 2 novel features were initially combined for optimum capturing customer behavior. Second, the principal component analysis (PCA) was implemented for reducing data dimensions. Tertiary, adaptive boosting (AdaBoost) was utilized for cascading several DL for minimizing the influences in the unbalanced data. Ramesh et al. [15] examine the effectual solution for all these challenging problem from CCP. The analysis utilizes data sets from the telecommunication industry, the artificial neural network (ANNs) and random forest (RF) for determining the features that stimulate consumer churn. A hybrid ANN based work was presented to forecast CCP.

This study designs an optimal deep canonically correlated autoencoder based prediction (O-DCCAEP) model for competitive customer dependent application sector. In addition, the O-DCCAEP approach purposes for determining the sentiment of the customers. The O-DCCAEP technique encompasses pre-processing, classification, and hyperparameter optimization. Furthermore, the DCCAE model is employed to classify the churners or non-churner. Furthermore, the hyperparameter optimization of the DCCAE technique occurs using the deer hunting optimization algorithm (DHOA). The experimental evaluation of the O-DCCAEP technique is carried out against an own dataset and the outcomes highlighting the betterment of the presented O-DCCAEP approach on the existing approaches.

The rest of the paper is organized as follows. Section 2 offers the proposed model and Section 3 discusses performance validation. Lastly, Section 4 concludes the paper.

2 Process Involved in O-DCCAEP Model

In this study, a novel O-DCCAEP approach was presented to forecast customer churns in the telecommunication industry. The presented O-DCCAEP model encompasses three different processes. Initially, the pre-processing is carried out to transform the customer data into meaning format. Moreover, the DCCAE model is employed to classify the churners or non-churner. Furthermore, the hyperparameter optimization of the DCCAE technique takes place using the DHOA. Fig. 1 illustrates the overall workflow of O-DCCAEP technique.

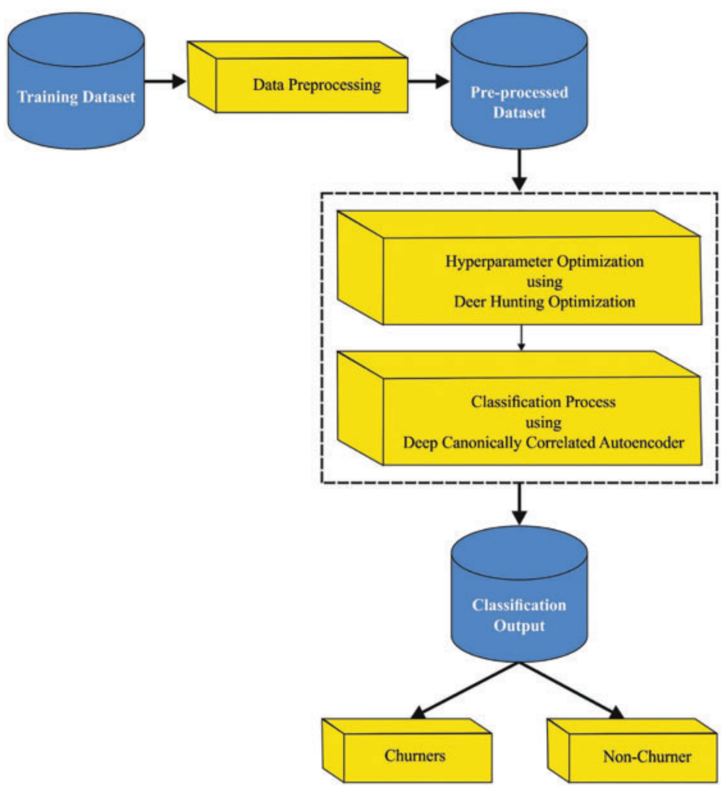


Figure 1: Workflow of O-DCCAEP model

2.1 DCCAE Based Classification

Once the customer data is pre-processed, they are fed into the DCCAE model to classify the churners or non-churner [16]. For tackling the limit of CCA technique, a deep neural network (DNN) is presented to canonical correlation analysis termed deep CCA (DCCA). The DCCA overcomes the constraint of CCA that couldn't recognize complexity nonlinear relation. In DCCA, 2 DNNs f and g are learned nonlinear representations to every data set. The DCCA is accomplished by exploited the canonical correlation of 2 DNNs consequences $f(X)$ and $g(Y)$:

$$\max_{W_f,W_g,U,V} \frac{1}{N} tr(U^T f(X) g(Y)^T V)$$

$$s.t. U^T(\frac{1}{N} f(X) f\left(X\right)^T + rI\big) U = I$$

$$V^T(\frac{1}{N} g(Y) g(Y)^T + sI) V = I$$

$$u_i^T f(Y) g(Y)^T v_j = 0, fori \neq j \tag{1}$$

Whereas N denotes the whole quantity of information, X and Y indicates the input matrices to 2 datasets, I denotes the identity matrix, f and g describes the nonlinear illustration of 2 DNNs with W_f and W_g parameters, $U = [u_{1,}, \ldots, u_L]$ and $V = [v_1, \ldots, v_L]$ indicates the CCA direction that existing DNN results to topmost layer with L unit, and $(r_x, r_y) > 0$ denotes the normalized parameter to covariance estimation. The DCCA is integrated with stacked autoencoder (SAE) for optimal representation of 2 datasets and a deep canonically correlated sparse autoencoder (DCSAE) which searches for deep network illustration of 2data sets for exploiting the canonical correlation among the 2 indifferent topmost feature, however, minimalizing the reform error of sparse AE. The DCSAE can be defined by:

$$\min_{W_f,W_g,U,V} -\frac{1}{N} tr(U^T f(X) g(Y)^T V) + \frac{\lambda}{N} \left\| \widehat{x^{(i)}} - x^{(i)} \right\|^2 + \| \overline{y^{(i)}} - y^{(i)} \|^2 + \alpha J_{KL}(\rho \| \widehat{\rho_j}) + \beta J_{KL}(\sigma \left\| \hat{\sigma}_{\rm k} \right) \tag{2}$$

where f and g imply the DNN employed to extract nonlinear features to each data set however, encode every input at the same time. $U = [u_1, \ldots, u_L]$ and $V = [v_1, \ldots, v_L]$ indicates the CCA direction that indicates the DNN results to topmost layer using L unit. $\widehat{x^{(i)}}$and $\widehat{y^{(i)}}$ suggests the reformation of input $x^{(i)}$ and input $y^{(i)}$ similarly $J_{KL}(\cdot)$ denotes as $JKL(\rho \| \widehat{\rho_j}) = \sum_{j=1}^{S} \rho log \frac{\rho}{\widehat{\rho_j}} + (1-\rho) log \frac{1-\rho}{1-\widehat{\rho_j}}$, whereas $\widehat{\rho_j} = \frac{1}{N} \sum_{i=1}^{N} h_j(x^{(i)})$ indicates the average activation of hidden layer j, correspondingly to $J_{KL}(\sigma \| \hat{\sigma}_{\rm k})$.

2.2 DHOA Based Parameter Optimization

At last, the hyperparameter optimization of the DCCAE technique occurs utilizng the DHOA [17]. The main purpose of the projected DHOA approach is for detecting an optimal place for an individual for hunting the deer, it can be important for exploring the deer's nature. It can contain special features which create difficult hunting for the predator. The distinct feature portrays visual power viz., 5 times superior related to humans. However, it required problem seeing green and red colors.

An important stage of method is the establishment of hunter population that is provided as:

$$Y = \{Y_1, Y_2, \cdots, Y_n\}; 1 < j \leq n \tag{3}$$

Assume n implies the total amount of hunters viz., the solution from the population Y.

Afterward, the beginning population, the deer place, and wind angles are essential parameters in defining an optimum hunter place were initializing. As the searching space is assumed as circles, the wind angle subsequently the circumference of circles.

$$\theta_i = 2\pi r \tag{4}$$

In which r refers the arbitrary number utilizing a value from the range of [0,1] and i signifies the existing iteration. In order to moment, the position angle of deer was signified as:

$$\phi_i = \theta + \pi \tag{5}$$

whereas θ signifies the wind angle.

While the place of optimal space was primarily unidentified, the method assumes the candidate solution nearby optimal viz., determined based on the fitness function (FF) as a better result. At this point, it regards the 2 outcomes, for instance, leadership place, Y^{lead}, refers the primary optimal place of the hunter and successor place, $Y^{successor}$ stands for the place of subsequent hunter.

(i) Propagation with leader place: then defining the optimal place all the individuals from the population tries to reach a better place and so, the process of upgrading the place begins. Accordingly, the surrounding nature was defined as:

$$Y_{i+1} = Y^{lead} - X \cdot p \cdot \left| L \times Y^{lead} - Y_i \right| \tag{6}$$

In which Y_i stands for the existing iteration place, Y_{i+1} represents the subsequent iteration place, X and L imply the co-efficient vector and p signifies the arbitrary number establish supposing the wind speed, whereas the value range of [0–2]. The co-efficient vectors are computed as:

$$X = \frac{1}{4} loglog \left(i + \frac{1}{i_{max}} \right) b \tag{7}$$

$$L = 2 \cdot c \tag{8}$$

Whereas i_{max} denotes the maximal iteration, b indicates the variables that contain a value amongst -1 and 1 and c defines the arbitrary numbers from the interval of [0,1]. Fig. 2 showcases the flowchart of DHOA.

In which (Y, Z) signifies the initial place of hunter that is upgraded dependent upon the prey place. The agent place was changed still an optimal place (Y^*, Z^*) has been obtained by adjusting L and X. All the hunters moved to leader place, once it can be effectual. However, the hunter remained from the existing place to ineffectual leader motion. The place upgraded subsequent in Eq. (7) if $p < 1$, refers the discrete is move arbitrarily from every direction regardless of angle places.

(ii) Propagation with angle location: to enhance the searching space, the idea was extended by considering the angles placed from the upgraded rule. The angle evaluation was essential to define the hunter place therefore prey could not be attentive to attack and hereafter, the hunting method is effectual. The visualization angle of prey/deer were computed as:

$$a_i = \frac{\pi}{8} \times r \tag{9}$$

Based on the difference amongst visual and wind angles of deer, the variable was computed that is support for upgrading the angling place.

$$d_i = \theta_i - a_i \tag{10}$$

Whereas θ signifies the wind angle. At that time, the angle places were upgraded to the subsequent iteration as:

$$\phi_{i+1} = \phi_i + d_i \tag{11}$$

With assuming the angling place, the place was upgraded for implementation as:

$$Y_{i+1} = Y^{lead} - p \cdot \left|\cos(v) \times Y^{lead} - Y_i\right| \tag{12}$$

In which $A = \phi_{i+1}$, Y_i^* signifies the optimal place and p signify the arbitrary number.

(iii) Propagation using successor location: During the exploration step, an identical model in surrounding nature was changed by implementing the vector L. Therefore, the upgrading place was based on the successor place rather than initial optimal solution reached. This allows a global search that is offered as:

$$Y_{i+1} = Y^{successor} - X \cdot p \cdot |L \times Y^{successor} - Y_i| \tag{13}$$

whereas, $Y^{successor}$ signifies the successor place of searching agent in the existing populations.

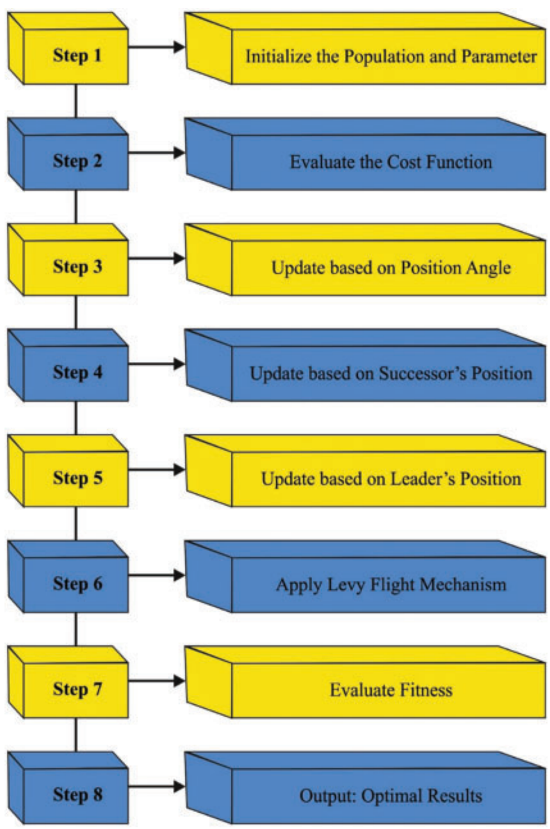


Figure 2: Flowchart of DHOA algorithm

3 Results and Discussion

In this section, a detailed experimental validation of the O-DCCAEP model is performed using benchmark churn dataset [2] from telecommunication industry. The dataset involves 3333 instances with 483 coming under churner class and 2850 coming under non-churner class.

Fig. 3 demonstrates five confusion matrices produced by the O-DCCAEP model on five distinct runs. On run-1, the O-DCCAEP model has recognized 322 samples into churners and 2649 samples into non-churners class. Besides, on run-2, the O-DCCAEP model has recognized 111 samples into churners and 2845 samples into non-churners class. In addition, on run-3, the O-DCCAEP model has recognized 150 samples into churners and 2847 samples into non-churners class. Also, on run-4, the O-DCCAEP model has recognized 303 samples into churners and 2643 samples into non-churners class. At last, on run-5, the O-DCCAEP model has recognized 166 samples into churners and 2843 samples into non-churners class.

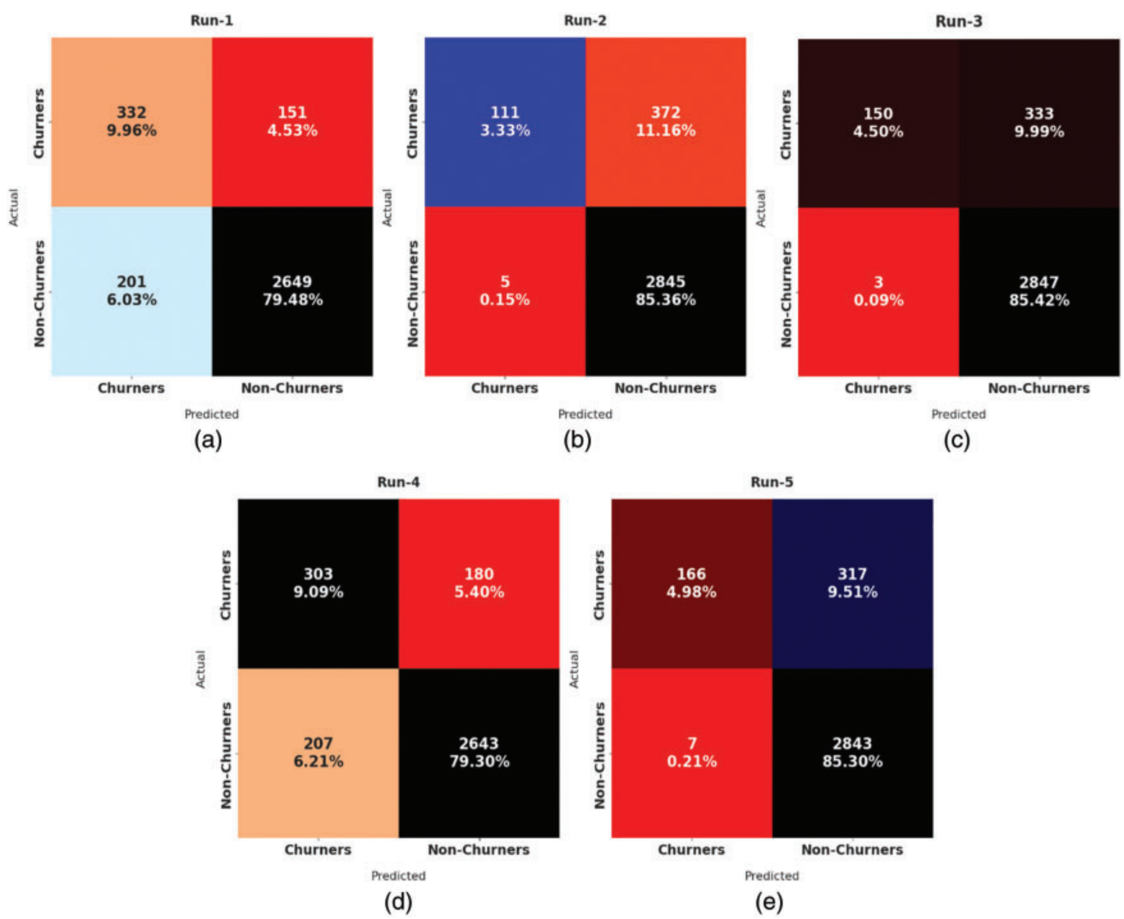
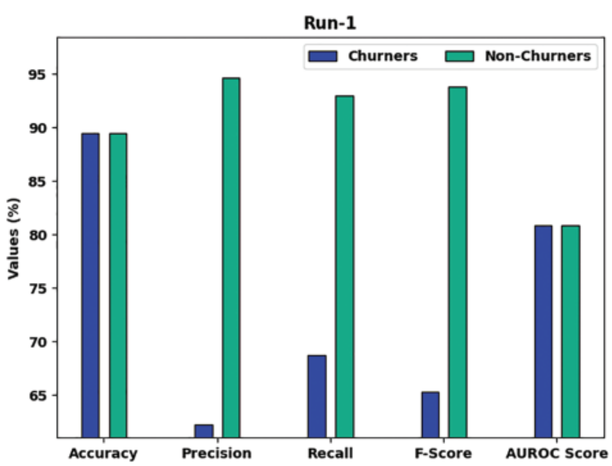


Figure 3: Confusion matrices of O-DCCAEP model

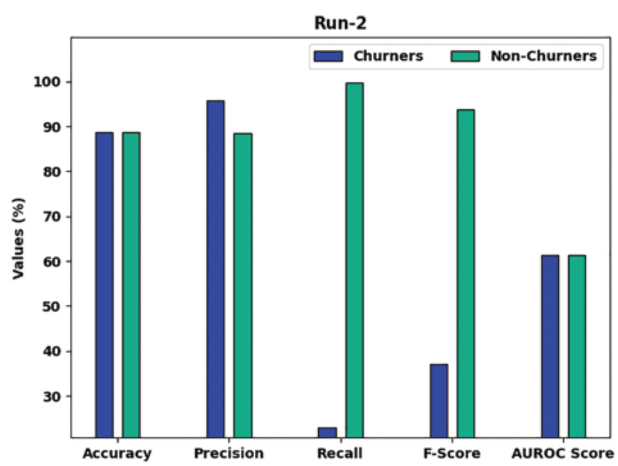
Tab. 1 and Fig. 4 report detailed CCP results of the O-DCCAEP model on the classification of churners and non-churners. The experimental results implied that the O-DCCAEP model has gained effectual outcomes with maximum CCP outcomes under distinct runs.

Table 1: CCP Results of O-DCCAEP model on five distinct runs

Class labels	Accuracy	Precision	Recall	F-score	AUC score
Run-1					
Churners	89.44	62.29	68.74	65.35	80.84
Non-churners	89.44	94.61	92.95	93.77	80.84
Average	89.44	78.45	80.84	79.56	80.84
Run-2					
Churners	88.69	95.69	22.98	37.06	61.40
Non-churners	88.69	88.44	99.82	93.79	61.40
Average	88.69	92.06	61.40	65.42	61.40
Run-3					
Churners	89.92	98.04	31.06	47.17	65.48
Non-churners	89.92	89.53	99.89	94.43	65.48
Average	89.92	93.78	65.48	70.80	65.48
Run-4					
Churners	88.39	59.41	62.73	61.03	77.73
Non-churners	88.39	93.62	92.74	93.18	77.73
Average	88.39	76.52	77.73	77.10	77.73
Run-5					
Churners	90.28	95.95	34.37	50.61	67.06
Non-churners	90.28	89.97	99.75	94.61	67.06
Average	90.28	92.96	67.06	72.61	67.06



(a)



(b)

Figure 4: (Continued)

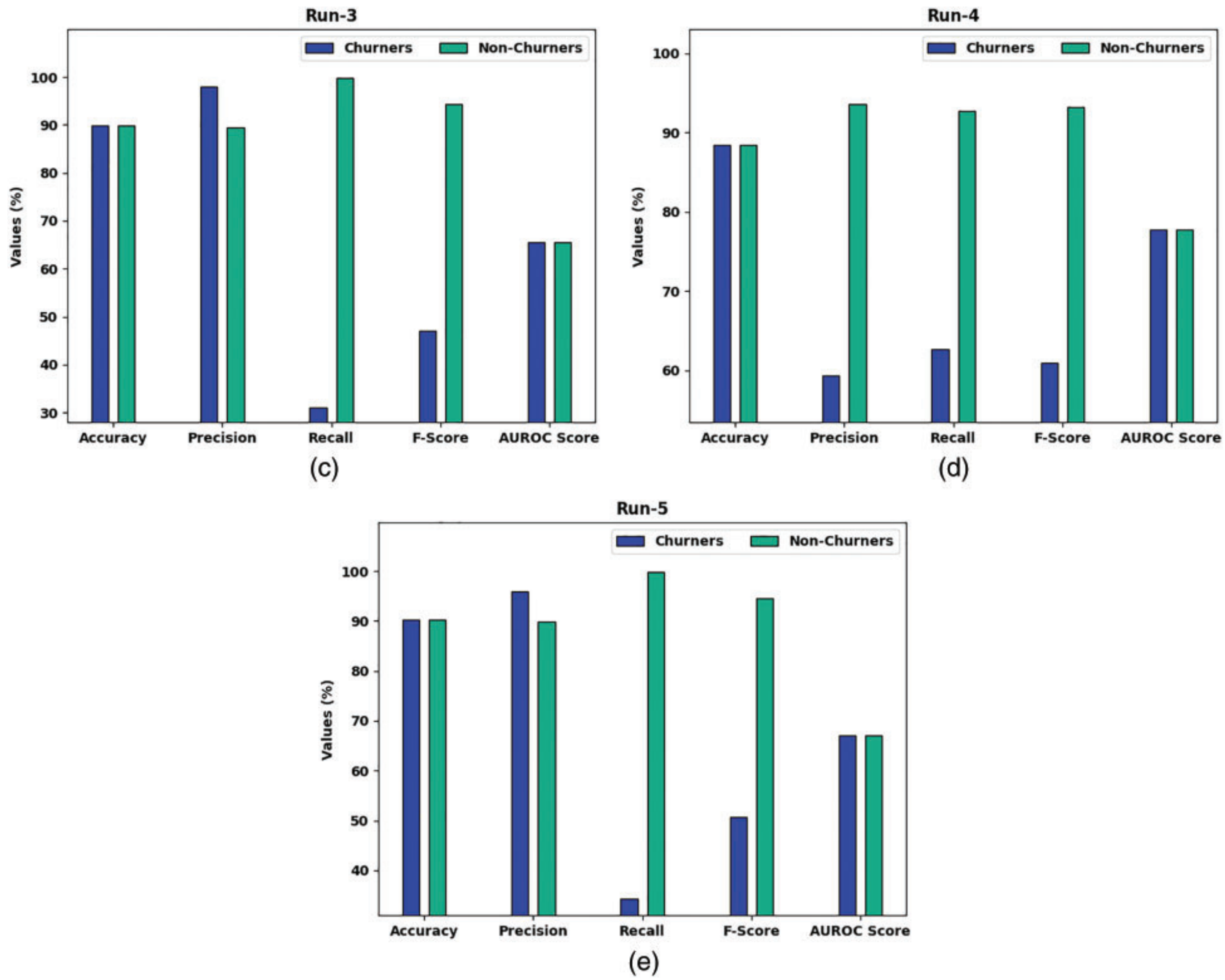


Figure 4: Overall CCP Results of O-DCCAEP model on five distinct runs

For sample, with run-1, the O-DCCAEP model has achieved average $accu_y$, $prec_n$, $reca_l$, F_{score},and AUC_{score} of 89.44%, 78.45%, 80.84%, 79.56%, and 80.48% respectively. In addition, with run-2, the O-DCCAEP model has accomplished average $accu_y$, $prec_n$, $reca_l$, F_{score},and AUC_{score} of 88.69%, 92.06%, 61.40%, 65.42%, and 61.40% respectively. Moreover, with run-3, the O-DCCAEP model has accomplished average $accu_y$, $prec_n$, $reca_l$, F_{score},and AUC_{score} of 89.92%, 93.78%, 65.48%, 70.80%, and 65.48% respectively. Furthermore, with run-5, the O-DCCAEP model has gained average $accu_y$, $prec_n$, $reca_l$, F_{score},and AUC_{score} of 90.28%, 92.96%, 67.06%, 72.61%, and 67.06% respectively.

Fig. 5 reports a detailed training accuracy (TA) and validation accuracy (VA) of the O-DCCAEP model on test data. The figure represented that the O-DCCAEP model has gained improved values of TA and VA.

Fig. 6 defines a comprehensive training loss (TL) and validation loss (VL) of the O-DCCAEP model on test data. The figure signified that the O-DCCAEP model has extended to least values of TL and VL.

Figure 5: TA and VA analysis of O-DCCAEP model on five distinct runs

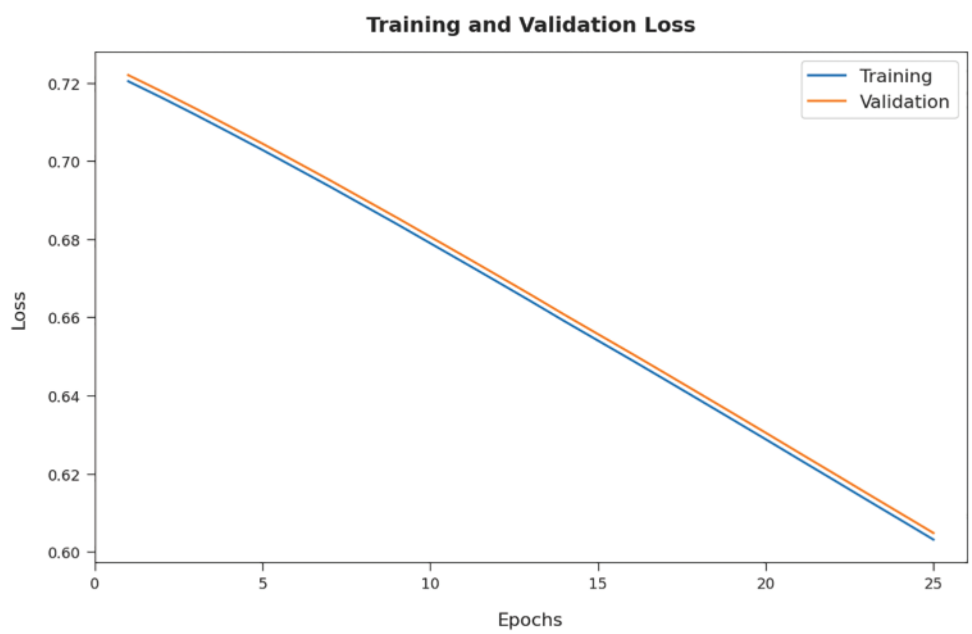


Figure 6: TL and VL analysis of O-DCCAEP model on five distinct runs

In order to exhibit the enhanced efficiency of the O-DCCAEP model, an extensive comparative examination is made in Tab. 2. Fig. 7 reports a detailed comparative study of the O-DCCAEP model with existing models in terms of $accu_y$, $reca_l$, and $prec_n$. The results implied that the decision tree (DT) and k-nearest neighbor (KNN) models have accomplished least values of $accu_y$, $reca_l$, and $prec_n$. Also, the logistic regression (LOR) method has reached somewhat higher values of $accu_y$, $reca_l$, and $prec_n$.

In line with, the Adaboost and CatBoost models have exhibited reasonable values of $accu_y$, $reca_l$, and $prec_n$. However, the O-DCCAEP model has shown effectual outcome with maximum $accu_y$, $reca_l$, and $prec_n$ values of 89.44%, 78.45%, and 80.84% respectively.

Table 2: Comparative CCP results of O-DCCAEP with recent models

Methods	Accuracy	Recall	Precision	F-score	AUC score
O-DCCAEP	89.44	78.45	80.84	79.56	80.84
LOR Model	81.64	75.71	80.11	78.48	76.75
DT Model	80.77	76.42	79.34	77.58	77.20
Adaboost	82.98	76.92	80.02	79.17	78.63
KNN Algorithm	80.58	77.24	79.41	77.54	72.26
CatBoost	82.92	76.97	78.11	78.92	74.61

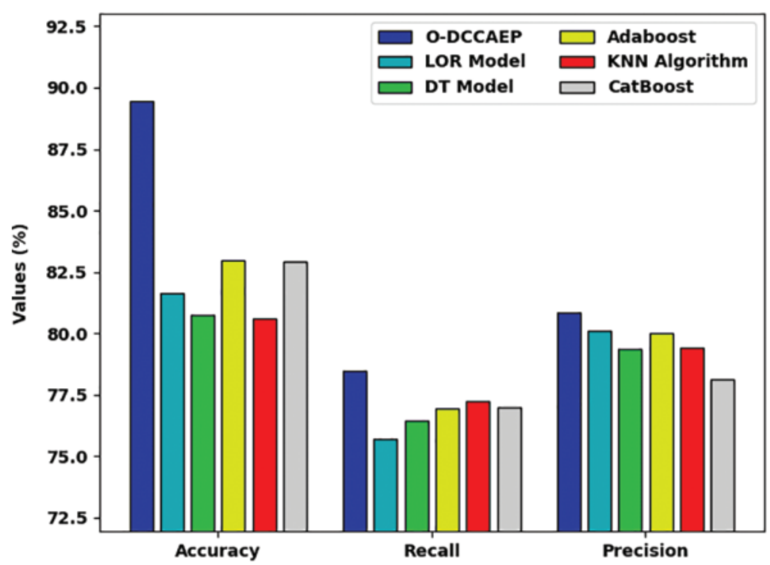


Figure 7: Comparative CCP results of O-DCCAEP model in terms of $accu_y$, $reca_l$, and $prec_n$

Fig. 8 portrays a comprehensive comparative study of the O-DCCAEP model with existing models in terms of F_{score} and. The experimental outcome demonstrated that the DT and KNN models have accomplished least values of F_{score} and. Eventually, the LOR model has gained slightly enhanced values of F_{score} and. Meanwhile, the Adaboost and CatBoost models have exhibited reasonable values of F_{score} and AUC. However, the O-DCCAEP model has shown effectual outcomes with maximum F_{score} and AUC values of 79.56% and 80.84% respectively. These results and discussion make sure the better outcomes of the O-DCCAEP model on CCP process.

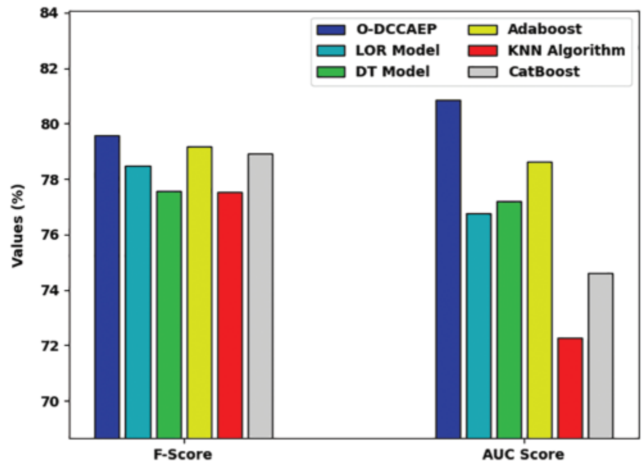


Figure 8: Comparative CCP results of O-DCCAEP model in terms of F_{score} and AUC

4 Conclusion

In this study, a novel O-DCCAEP model was introduced to forecast customer churns in the telecommunication industry. The presented O-DCCAEP model encompasses three different processes. Initially, the pre-processing is carried out to transform the customer data into meaning format. Moreover, the DCCAE model is employed to classify the churners or non-churner. Furthermore, the hyperparameter optimization of the DCCAE technique takes place using the DHOA. The experimental evaluation of the O-DCCAEP technique is carried out against an own dataset and the outcomes highlighted the betterment of the presented O-DCCAEP approach on the recent approaches with maximum $accu_y$, $reca_l$, and $prec_n$ values of 89.44%, 78.45%, and 80.84% respectively. Thus, the O-DCCAEP model can be utilized for improvising the customer churn prediction performance. In future, advanced DL models can be utilized for feature selection process.

Funding Statement: This research work was partially supported by Chiang Mai University and Umm Al-Qura University.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] N. N. Y. Vo, S. Liu, X. Li and G. Xu, “Leveraging unstructured call log data for customer churn prediction,” *Knowledge-Based Systems*, vol. 212, no. 4, pp. 106586, 2021.

[2] I. V. Pustokhina, D. A. Pustokhin, P. T. Nguyen, M. Elhoseny and K. Shankar, “Multi-objective rain optimization algorithm with WELM model for customer churn prediction in telecommunication sector,” *Complex & Intelligent Systems*, vol. 38, no. 12, pp. 15273, 2021.

[3] P. Lalwani, M. K. Mishra, J. S. Chadha and P. Sethi, “Customer churn prediction system: A machine learning approach,” *Computing*, vol. 104, no. 2, pp. 271–294, 2022.

[4] Y. Li, B. Hou, Y. Wu, D. Zhao, A. Xie *et al.,* “Giant fight: Customer churn prediction in traditional broadcast industry,” *Journal of Business Research*, vol. 131, no. 10, pp. 630–639, 2021.

[5] A. D. Caigny, K. Coussement, W. Verbeke, K. Idbenjra and M. Phan, "Uplift modeling and its implications for B2B customer churn prediction: A segmentation-based modeling approach," *Industrial Marketing Management*, vol. 99, pp. 28–39, 2021.

[6] M. Zhao, Q. Zeng, M. Chang, Q. Tong and J. Su, "A prediction model of customer churn considering customer value: An empirical research of telecom industry in china," *Discrete Dynamics in Nature and Society*, vol. 2021, no. 5, pp. 1–12, 2021.

[7] M. J. Shabankareh, M. A. Shabankareh, A. Nazarian, A. Ranjbaran and N. Seyyedamiri, "A stacking-based data mining solution to customer churn prediction," *Journal of Relationship Marketing*, vol. 21, no. 2, pp. 124–147, 2022. https://doi.org/10.1080/15332667.2021.1889743.

[8] A. Dalli, "Impact of hyperparameters on deep learning model for customer churn prediction in telecommunication sector," *Mathematical Problems in Engineering*, vol. 2022, no. 5, pp. 1–11, 2022.

[9] S. Kim and H. Lee, "Customer churn prediction in influencer commerce: An application of decision trees," *Procedia Computer Science*, vol. 199, no. 2, pp. 1332–1339, 2022.

[10] D. Melian, A. Dumitrache, S. Stancu and A. Nastu, "Customer churn prediction in telecommunication industry. A data analysis techniques approach," *Postmodern Openings*, vol. 13, no. 11, pp. 78–104, 2022.

[11] I. V. Pustokhina, D. A. Pustokhin, R. H. Aswathy, T. Jayasankar, C. Jeyalakshmi *et al.,* "Dynamic customer churn prediction strategy for business intelligence using text analytics with evolutionary optimization algorithms," *Information Processing & Management*, vol. 58, no. 6, pp. 102706, 2021.

[12] K. W. D. Bock and A. De Caigny, "Spline-rule ensemble classifiers with structured sparsity regularization for interpretable customer churn modeling," *Decision Support Systems*, vol. 150, pp. 113523, 2021.

[13] S. F. Bilal, A. A. Almazroi, S. Bashir, F. H. Khan and A. Ali Almazroi, "An ensemble based approach using a combination of clustering and classification algorithms to enhance customer churn prediction in telecom industry," *PeerJ Computer Science*, vol. 8, no. 8, pp. e854, 2022.

[14] Z. Wu, L. Jing, B. Wu and L. Jin, "A PCA-AdaBoost model for E-commerce customer churn prediction," *Annals of Operations Research*, vol. 66, no. 4, pp. 603, 2022.

[15] P. Ramesh, J. Jeba Emilyn and V. Vijayakumar, "Hybrid artificial neural networks using customer churn prediction," *Wireless Personal Communications*, vol. 124, pp. 1695–1709, 2022. https://doi.org/10.1007/s11277-021-09427-7.

[16] T. Yu, "AIME: Autoencoder-based integrative multi-omics data embedding that allows for confounder adjustments," *PLO0S Computational Biology*, vol. 18, no. 1, pp. e1009826, 2022.

[17] S. Raj, S. Tripathi and K. C. Tripathi, "ArDHO-deep RNN: Autoregressive deer hunting optimization based deep recurrent neural network in investigating atmospheric and oceanic parameters," *Multimedia Tools and Applications*, vol. 81, no. 6, pp. 7561–7588, 2022.