SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V10I5P105

Volume 10 Issue 5, 48-59, May 2023

Original Article

Robust Adverse Drug Reaction Prediction and Classification by Employing Deer Hunting Optimization Driven Deep Learning Approach

S. Nithinsha[1], S. Anusuya[2]

[1]*Department of Computer Science, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences*
[2]*Department of Information Technology, Saveetha School of Engineering,*
Saveetha Institute of Medical and Technical Sciences

[1]*Corresponding Author : ayshaas2017@gmail.com*

Received: 04 March 2023 Revised: 12 April 2023 Accepted: 06 May 2023 Published: 29 May 2023

Abstract – *Drugs for medical purpose aims to save a person's life and improve the quality of life. They enhance one's mental or physical fitness and treat disease. However, drugs might cause adverse reactions or even unexpected effects on patients, known as Adverse Drug Reactions (ADRs). For preventing ADRs, drug trials were carried out to evaluate and analyze the potential danger in the procedure of drug development. Deep learning (DL) is a type of Machine Learning (ML) in Artificial intelligence (AI) that has developed as a highly effective and promising method that can interrogate and combine diverse biological data kinds to produce new hypotheses. DL is extensively used in drug repurposing and discovery, but its application in ADR prediction employing gene expression data is limited. Therefore, this study introduces a Deer Hunting Optimization Driven Deep Learning Model for Robust Adverse Drug Reaction Recognition and Classification (DHODL-ADRRC) technique. The DHODL-ADRRC technique involves diverse phases of data preprocessing to normalize the data. In addition, the Binary Fire Hawks Optimization (BFHO) based feature selection (FS) method is used to elect an optimum set of features. Moreover, the Attention-based Convolutional Bidirectional Long Short-Term Memory (ACBLSTM) algorithm is utilized for ADR identification. Furthermore, the DHO model is exploited for the tuning process adjusting of the ACBLSTM model, advancing the classification outputs. The simulation output of the DHODL-ADRRC algorithm was investigated on the ADR dataset, and the extensive outputs highlighted the advanced achievement of the DHODL-ADRRC method over other current methods in terms of different measures.*

Keywords - *Adversary drug reaction, Artificial intelligence, Feature selection, Deer hunting optimizer, Deep learning.*

1. Introduction

The World Health Organization (WHO) described Pharmacovigilance as the science and activity for preventing, detecting, assessing, and observing adverse impacts or any other drug-relevant issue [1]. Measuring the risk of medical products over medical trials is a significant aspect. An unintentional disease or symptom linked with a medical product is called an adverse event (AE) [2]. A linked medical product may not be concerned relevant to adverse events and is a suspect drug [3]. In addition, the linked medical product and adversarial events are considered to have a causal relation. Recognizing adversarial events from incoming reports is beneficial for constructing a knowledge base that could be leveraged for future detection of drug-adverse event pairs with the help of association rules mining [4]. The available wide-ranging biomedical data having concepts and entities offers data to present automation in Pharmacovigilance [5]. The issue of identifying novel drug-adverse event signals utilizing combined numbers of reports is broadly studied [6]. As it becomes easy for rule-related solutions to lack robustness, the existence of external attributes with high association with saved data provided robustness for identifying adversarial reactions [7].

Over the past, the growth of advanced and potential models for the target delivery of therapeutic agents with maximal efficacy [8] and minimal dangers has posed an excessive threat among biological and chemical researchers [9]. Moreover, emerging new therapeutic agents' time consumption and development cost were hindrances in the drug development and design procedure [10]. To minimalize such hurdles and challenges, authors moved toward computing approaches like molecular docking and virtual screening, which are classical approaches [11]. Still, such approaches impose challenges like inefficiency and inaccuracy. Hence, novel methods are increasingly applied,

which are self-sufficient for eradicating the difficulties in old calculation approaches. ML, AI, and containing DL algorithms have appeared as possible solutions that could solve drug discovery and design issues [12]. Drug discovery and devising also contain complicated and long steps like validation and target selection, clinical and pre-clinical trials, industrial practices, lead compound optimization and therapeutic screening [13]. These steps pose another massive challenge in detecting potential medication against diseases. Hence, AI has become an effective tool that helps pharmaceutical companies handle speed and procedure costs. Also, the rise in data computerization in the pharmacological firms and medical sector motivated the application of AI in resolving the problems of investigating complex datasets [14, 15]

This study introduces a Deer Hunting Optimization Driven Deep Learning Model for Robust Adverse Drug Reaction Recognition and Classification (DHODL-ADRRC) technique. The DHODL-ADRRC technique involves different stages of data preprocessing to normalize the data. In addition, Binary Fire Hawks Optimization (BFHO) based FS technique is used to elect an optimum set of features. Moreover, the Attention-based Convolutional Bidirectional Long Short-Term Memory (ACBLSTM) algorithm is utilized for ADR identification. Furthermore, the DHO model is exploited for the tuning process adjusting of the ACBLSTM model, advancing the classification outputs. The experimental investigation of the DHODL-ADRRC algorithm was investigated on the ADR dataset, and the outputs are assessed under various evaluation measures.

2. Related Works

Chi-Shiang Wang et al. [16] devised a DNN approach using drugs' biomedical, chemical, and biological data to detect ADRs. This method is intended to achieve two critical purposes: detecting the ADRs of medications and forecasting the likely ADRs of a novel medication. The author dispersed representations of targeted drugs in vector space for capturing drug relations to improve detection performance. [17], introduced a DL-related method for ADE detection and extraction utilizing BERT and compared outcomes to standard DL methods and recent extraction methods utilizing hybrid TL from sentence embedding and pre-trained BERT representations. Jiajing Zhu et al. [18] offer a method for ADDI prediction called Deep Attributed Embedding based Multi-task (DAEM) learning method. Specifically, two medication attributes, side effects and molecular structure, are implemented to devise the adverse contacts among drugs and a DNN was modelled to embed the hand-designed features.

Arnold K. Nyamabo et al. [19] present GMPNN (gated message passing NN), a data passing NN that learned chemical structures with various shapes and sizes from the molecular graph representation of medications for DDI forecast among the drug pair. In this presented method, edges are considered gates that regulate the data flow passing, thus de-restricting the structures in a learning manner. Christopher McMaster et al. [20] presented a DL-NLP method for finding ADRs in discharge summaries. This technique has been developed in two phases: first, a DeBERT; then, this model has been fine-tuned to identify ADR. This method was related to the version without the pretraining step. Chun Yen Lee, and Yi-Ping Phoebe Chen [21] introduced a method for classifying and detecting side effects utilizing DL methods. It is revealed that the potential incorporation of diverse, multidimensional medication data sources, with the state-of-the-art DL method deployment, aids in preventing or reducing the occurrence of ADRs. DL methods can be used to identify substitutes for drugs that have side effects in [22], devised an ensemble method for classification relation and extraction between medication-related and drug entities. The author incorporated named-entity recognition (NER) methods depending on BiLSTM networking and Conditional Random Fields (CRF) for end-wise extracting.

3. The Proposed Model

This article proposes a DHODL-ADRRC model for the ADR detection and classification process. The DHODL-ADRRC technique comprises four phases of operations, namely data pre-processing, BFHO-based FS, ACBLSTM-based ADR recognition, and DHO-based hyper parameter tuning. Figure 1 demonstrates the workflow of the DHODL-ADRRC model.

3.1. Data Preprocessing

Here, the DHODL-ADRRC method involves diverse phases of data pre-processing to normalize the information. Initially, the input dataset is pre-processed in three distinct phases. Initially, the missing value is eliminated from the data. Then, the data conversion process transforms the categorical value into a mathematical value. Next, the Z-score normalizing procedure is used for normalizing the input dataset into a uniform format. Now, the raw dataset gets adapted by using.

$$Z_{score} = \frac{X_i - X_{mean}}{S} \quad (1)$$

In Eq. (1), X_i represents the raw value that each variable continues, X_{mean} Denotes the average of the parameter value, and S characterizes the Standard Deviation (SD). Thus, the raw ratio was normalized, with 0 and SD as the unit average across the sample.

3.2. Feature Selection by Utilizing BFHO Algorithm

This section uses the BFHO-based FS method to choose an optimum set of features. The FHO method motivates Fire Hawks (FH) foraging behaviours that consider spreading fires, catching prey, and setting [23].

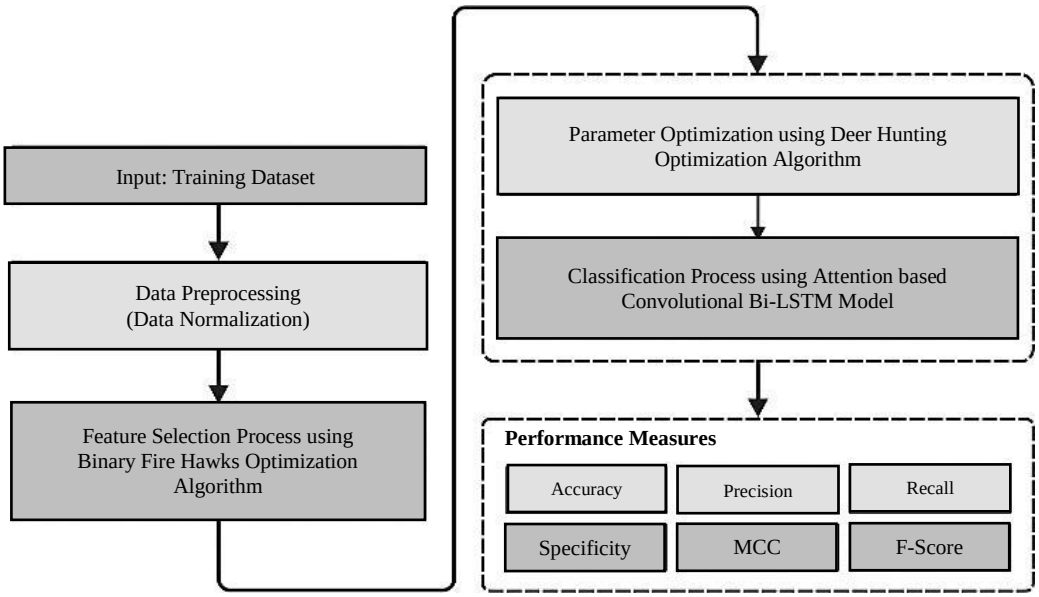

Fig. 1 Workflow of DHODL-ADRRC algorithm

Initially, a solution candidate (X) is described using the position vector of the FH and prey. The initial procedure is employed to recognize the initial position of these vectors in the search area.

$$x = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix}$$

$$\begin{bmatrix} x_1^1 & x_1^2 & \dots & x_1^j & \dots & x_1^d \\ x_2^1 & x_2^2 & \dots & x_2^j & \dots & x_2^d \\ \vdots & \vdots & \dots & \ddots & \dots & \vdots \\ x_i^1 & x_i^2 & \dots & x_i^j & \dots & x_i^d \\ \vdots & \vdots & \dots & \ddots & \dots & \vdots \\ x_N^1 & x_N^2 & \dots & x_N^j & \dots & x_N^d \end{bmatrix}, \begin{cases} i = \{1,2, \dots N\}. \\ j = \{1,2, \dots d\}. \end{cases} \quad (2)$$

$$x_i^j(0) = x_{i,\min}^j + rand.(x_{i,\max}^j - x_{i,\min}^j), \begin{cases} i = \{1,2, \dots \quad N\}. \\ j = \{1,2, \dots \quad d\}. \end{cases} \quad (3)$$

Now, $x_{i,\min}^j$ and $x_{i,\max}^j$ Signify the minimal and maximal limits of the $j\text{-}th$ search space for the $i\text{-}th$ candidate solution; N characterizes the total amount of candidate solution in the search area; X_i Symbolizes the $i\text{-}th$ candidate solution in the search region; d specifies the dimension vector; $x_i^j(0)$ symbolizes the initial location of the solution candidate; x_i^j Characterize the $j\text{-}th$ search space of the $i\text{-}th$ candidate solution; $rand$ denotes a random value distributed uniformly within [0,1]. Now, the global optimum solution is regarded as the most critical fire that is initially employed by the FH to b:

$$PR = \begin{bmatrix} PR_1 \\ PR_2 \\ \vdots \\ PR_k \\ \vdots \\ PR_m \end{bmatrix} k = 1,2, \dots m, \quad (4)$$

$$FH = \begin{bmatrix} FH_1 \\ FH_2 \\ \vdots \\ FH_l \\ \vdots \\ FH_n \end{bmatrix} L = 1,2, \dots n, \quad (5)$$

Consider PR_k as a $k\text{-}th$ prey in the search range concerning the total amount of m prey and FH_1 represents the lth FH regarded as an overall amount of n FH in the searching region. Then, the FH and the overall distance between the prey are assessed.

$$D_k^1 = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}, \begin{cases} l = 1,2, \dots \dots n. \\ k = 1,2, \dots \dots m. \end{cases} \quad (6)$$

Where D_k^1 indicates the overall distance between the $k\text{-}th$ prey and lth FH; m shows the overall amount of prey in the search field; n shows the overall amount of FH in the search domain, and (x_1, y_1) and (x_2, y_2) characterize the coordination of FH and prey in the search domain. Meanwhile, birds are enthusiastic about applying the burning stick from the FH terrain; for such reasons, this behaviour was used as a location updating method in the significant search looping of FHO:

$$FH_1^{new} = FH_1 + (r_1 \times GB - r_2 \times FH_{Near}), l = 1,2,, n, \quad (7)$$

In Eq. (7), FH_1^{new} indicates the novel location vector of lth FH (FH_1); GB denotes the optimal global solution in the searching domain assumed as the primary fire; FH_{Near} denotes the other FHs in the searching domain; and r_1 and r_2 are random numbers distributed uniformly within $(0,1)$ to define the movement of FH toward the significant fire and the other FHs' territory. Then, the motion of prey inside the territory of FH assumed a significant aspect of animal activities for the location updating method.

$$PR_q^{new} = PR_q + (r_3 \times FH_1 - r_4 \times SP_1), \begin{cases} l = \{1,2, \dots n\}. \\ q = \{1,2, \dots r\}. \end{cases} \quad (8)$$

In Eq. (8), PR_q^{new} indicates the novel location vector of qth prey (PR_q) enclosed by lth FH_1; GB shows the optimal global solution in the searching domain assumed as major fire; SP_1 indicate a safer position under lth FH territory; and r_3 and r_4 are random numbers distributed uniformly with $(0,1)$ to define the motion of prey towards the FH and the safer position?

In addition, the prey motion toward the other FH territories while there is a chance the prey might be approaching the FH in the nearby traps or hide in a safe position outside the FH territories where they are trapped:

$$PR_q^{new} = PR_q + (r_5 \times FH_{Alter} - r_6 \times SP), \begin{cases} l = \{1,2, \dots r\}. \\ q = \{1,2, \dots n\}. \end{cases} \quad (9)$$

In Eq. (9), PR_q^{new} denotes the novel location vector of qth prey (PR_q) enclosed by the lth FH_1; FH_{Alter} denotes the FH in the searching domain; SP indicates a safer potion outside the lth FH territories; r_5 and r_6 are random numbers distributed uniformly within $(0,1)$ to define the movement of prey towards the FH and the safer location outside the territory, and it is mathematically expressed in the following:

$$SP_1 = \frac{\sum_{q=1}^{r} PR_q}{r}, \begin{cases} q = 1,2, \dots r. \\ l = 1,2, \dots n. \end{cases} \quad (10)$$

$$SP = \frac{\sum_{k=1}^{m} PR_k}{m}, k = 1,2, \dots, m. \quad (11)$$

Now, PR_q indicates the qth prey bounded by the lth FH_1; PR_k denotes the k-th prey in the searching domain.

The fitness function (FF) projected for having balancing among the classifier accuracy (maximal) and the amount of FS in every outcome (minimal) accomplished using the FS, Eq. (10) signifies the FF for estimating the solution.

$$Fitness = \alpha \gamma_R(D) + \beta \frac{|R|}{|C|} \quad (12)$$

Where, α and β indicate the two parameters equivalent to the classifier quality's significance and the subset length. $\alpha \in [1,0]$ and $\beta = 1 - \alpha$. $\gamma_R(D)$ characterizes the classifier error rate of a provided classifier. $|R|$ denotes the selected subset's cardinality, and $|C|$ shows the overall amount of factors in the data.

3.3. ADR Recognition by Utilizing ACBLSTM Technique

At this stage, the ACBLSTM technique is employed for ADR detection. The CLSTM model is a practical kind of RNN, with LSTM and convolution layers that operate within the LSTM cell [24]. At the same time, the travelling time predictive bus travelling is processed as a time series predictive problem. Recently, LSTM can be a superior solution to time series prediction by using spatiotemporal datasets. In addition, the CLSTM employed the convolutional operator for capturing the spatial and temporal dependency in the data such that it is usually implemented well than FC-LSTM. The computation steps are given.

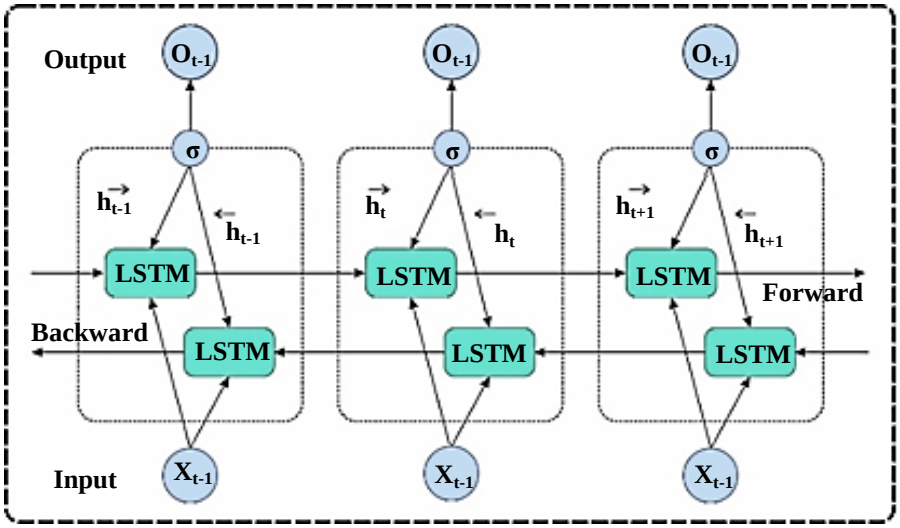

Fig. 2 Architecture of BLSTM

Firstly, evaluate the input gate:

$$i_t = \sigma(W_{xi} * x_\tau + W_{hi} * h_{t-1} + W_{ci} \circ c_{t-1} + b_i), \quad (13)$$

Forget gate:

$$f_t = \sigma\big(W_{xf} * x_t + W_{hf} * h_{t-1} + W_{cf} \circ c_{t-1} + b_f\big), \quad (14)$$

Cell state:

$$c_\tau = f_t \circ c_{t-1} + i_t \circ tanh(W_{xc} * x_t + W_{hc} * h_{t-1} + b_c), \quad (15)$$

Output gate:

$$o_\tau = \sigma(W_{xo} * x_\tau + W_{ho} * h_{t-1} + W_{co} \circ c_t + b_o), \quad (16)$$

Hidden state:

$$h_t = \mathrm{O}_t \circ tanh(c_t), \quad (17)$$

Where it shows the convolution operators, σ represents the sigmoid function and $\circ$ denotes the Hadamard product. W_{hi}, W_{hf}, W_{hc} and W_{ho} shows the weighted matrix linking the hidden state h_{t-1} to three gates and the cell input; W_{cf} and W_{co} represent the weight matrix connecting the c_t to three gates; W_{xi}, W_{xf}, W_{xc} and W_{xo} Indicate the weight matrix which connects the input. x_t to 3 gates and the cell input; and b_i, b_f, b_c and b_o indicates bias term of three gates and cell state. Unlike typical LSTM, the input flow is in both directions and can utilize data from both sides. Also, it is an effective method to model the sequential dependencies between phrases and words. In summary, BLSTM adds further LSTM layers that reverse the direction of data flow.

The attention mechanism concentrates on the major problem with the CBLSTM-based mechanism for predicting the bus travel time that tends to choose near-term dataset that is extremely related to upcoming travelling time. In this work, the encoded is the fundamental ACBLSTM method which generates the hidden state depiction. A self-attention mechanism was leveraged to the input after the operation of Eqs. (13) - (17):

$$m_{t,t'} = tanh(W_m h_t + W_{m'} h_{t'} + b_m), \quad (18)$$

$$e_{t,t'} = \sigma\big(W_a m_{t,t'} + b_a\big), \quad (19)$$

$$a_t = softmax(e_t), \quad (20)$$

$$l_t = \Sigma_{t'=1}^{n} a_{t,t'} \cdot h_{t'}, \quad (21)$$

Now $a_{t,t'}$ denotes an attention matrix; b_m and b_a indicate the biased terms; W_m and $W_{m'}$ shows weight matrices equivalent to the hidden state h_t, h_t and l_t characterizes a weighted amount of $h_{t'}$. Figure 2 represents the architecture of the BLSTM technique.

3.4. Hyper Parameter Tuning using DHO Algorithm

Finally, the DHO method is exploited for the hyper parameter adjusting the ACBLSTM model, enhancing the classification results. The DHO algorithm is proposed based on the inspiration or idea derived from the hunting model of humans near the deer [25]. Specific characteristics of deer are given in the following: awareness of ultra-higher frequency sound, outstanding sensing of visual and superior smelling ability. The mathematical model of the DHO model is provided as follows:

Step 1: Initialized populations.
In this work, the hunters' population is initialized and shown below.

$$T = \{T_1, T_2, \dots, T_n\}; 1 < p \le q \quad (22)$$

In Eq. (22), T denotes the population of the hunters, and q specifies the total hunter amount.

Step 2: Initialize the angle of wind and location.

$$\phi_t = 2\pi h \quad (23)$$

The notation ϕ characterizes the wind angle. θ denote the position angle. The random integer is denoted by$h$. The existing iteration is shown by t.

$$\theta_r = \phi + \pi \quad (24)$$

Step3: Propagation of location

Assume two locations. It can be successor locations ($T^{successor}$) and leader location (T^{lead}).

Propagation process performed by the location of the leader:

The location updating procedure of the hunter is initiated and performed as the optimum location is considered for finding the optimum solution.

$$T_{t+1} = T^{lead} - I.b.|Q \times T^{lead} - T_t| \quad (25)$$

In Eq. (25), the hunter location in the existing iteration is represented by T_t, the upgraded location of the hunters in the next iteration is represented as T_{t+1}, b denotes the random parameter that ranges from zero to two is present owing to the wind speed and the co-efficient vector represented with the representation I and Q, which is calculated using the mathematical equation as follows:

$$I = {}^{1/4} \log\left(t + \frac{1}{t_{\max}}\right) c \quad (26)$$

$$Q = 2.a \quad (27)$$

In Eq. (25) $t_{\max}$ denotes the maximal amount of iteration and the parameter value c to compute the coefficient vector I ranges between [1,1]. The random integer lies in the interval.

The location angle performs the propagation model:

The location angle is upgraded by presenting the newest variable d_r that is taken from the variation between the angle visualized and the wind angle. The angle of visualized k_t the prey is defined as follows.

$$d_r = \phi_t - k_t \quad (28)$$

$$k_t = \frac{\pi}{8} \times h \quad (29)$$

The successor location was provided as $^{the\ T-lead}$ of the searching agent. The algebraic formula can perform the location angle update, and the hunter position is upgraded considering the location angle as follows:

$$\theta_{r+1} = \theta_r + d_r \quad (30)$$

$$T_{t+1} = T^{lead} - b.|\cos(w) \times T^{lead} - T_t| \quad (31)$$

The propagation procedure is done with the successor position:

The concept of the encircling strategy has been implemented by adapting the vector Q existing in the exploration phase. Even though a random search was considered, the vector Q is regarded as lesser than 1. The hunter's location based on the successor's location is represented as follows.

$$T_{t+1} = T^{successor} - I.b.|Q \times T^{successor} - T_t| \quad (32)$$

Once the Q value is lesser than 1, the search agent is stochastically selected. Once the Q value is higher than or equivalent to 1, the better solution was chosen for updating the search agent location.

Step4: Termination of the procedure

The last phase in the DHO algorithm is to end the procedure. The better location is now accomplished, and the location updating is performed for every iteration.

The DHO model also derives Fitness Functions (FFs) to accomplish maximum $accu_y$ of the classification and defines positive integers to specify the improved $accu_y$ of the candidate outcomes. The decline of classifier error rate is assumed that FF as:

$$fitness(x_i) = ClassifierErrorRate(x_i)$$

$$= \frac{number\ of\ misclassified\ samples}{Total\ number\ of\ samples} * 100 \quad (33)$$

4. Performance Validation

The ADR detection and classification results of the DHODL-ADRRC technique are examined on the Metformin ADR dataset, comprising 16 features and two classes.

Among the available 16 features, the BFHO algorithm has chosen six features: adverse event, RRR, ROR_lower_bound, De, de, and N.

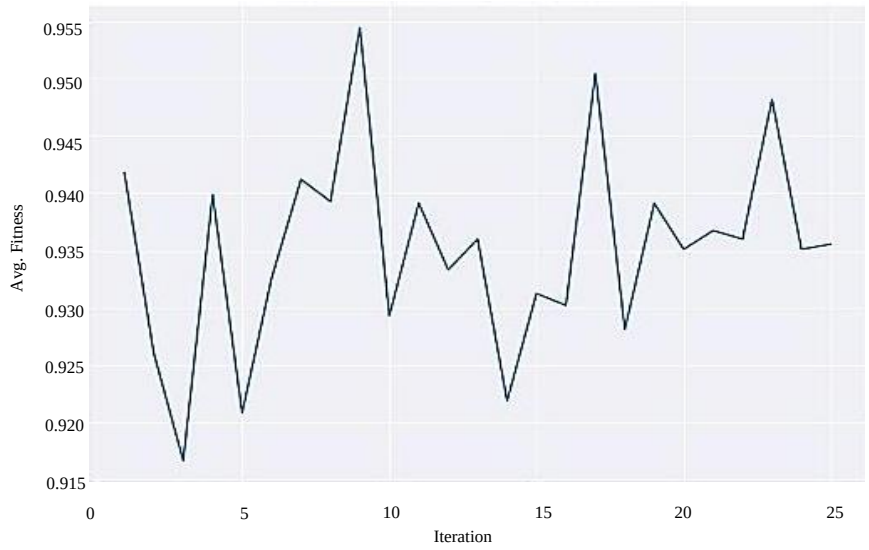

Fig. 3 Convergence analysis of binary FH 's optimization algorithm

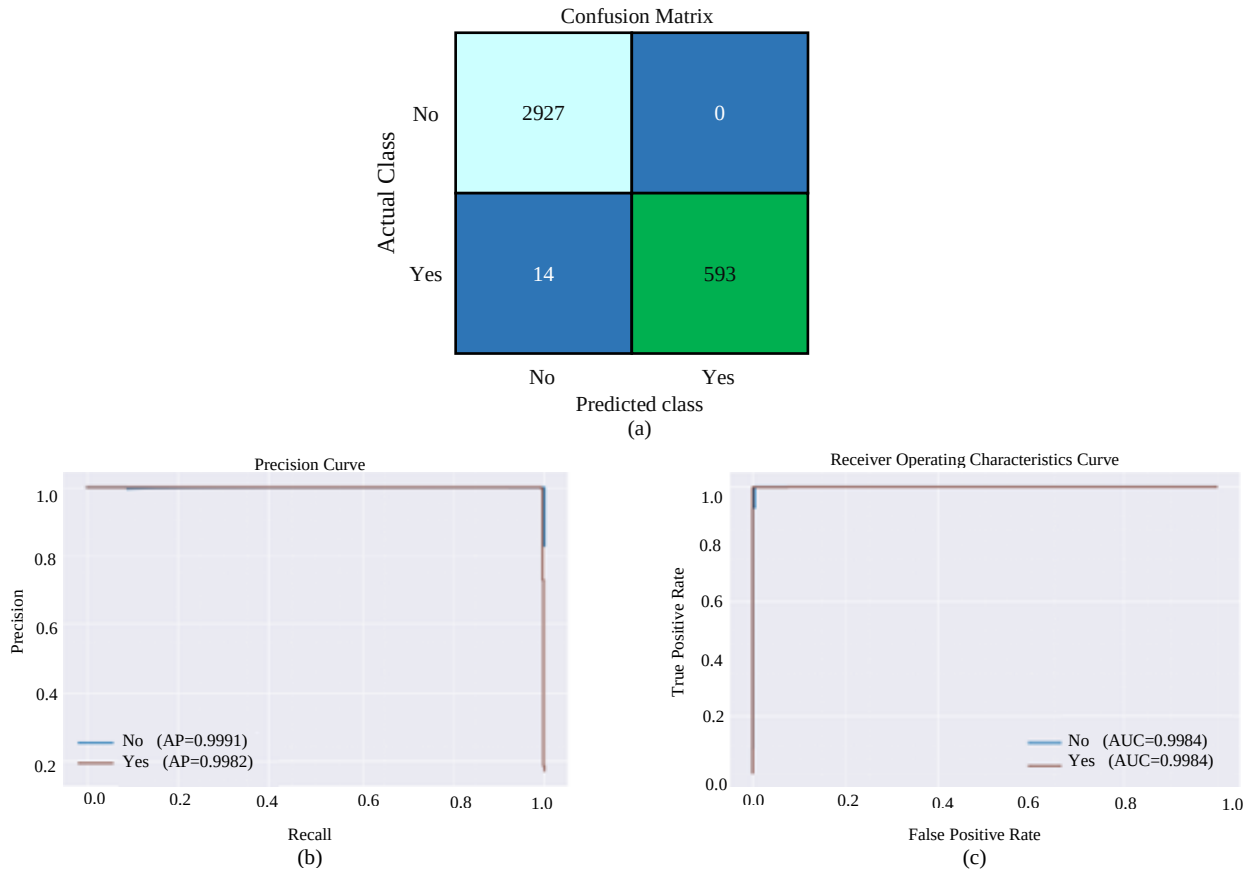

Fig. 4 Training set a) Confusion matrix b) PR-curve c) ROC analysis

Figure 3 shows the convergence analysis of the FHO approach. The results indicate that the FHO method exhibits fast convergence over several iterations of the applied data. Figure 4 exhibits the classifier results of the DHODL-ADRRC method under the training set. Figure 4a depicts the DHODL-ADRRC method's confusion matrix on 70% of TRS. The figure denoted that the DHODL-ADRRC approach has identified 2927 instances under ADR-No and 593 instances under ADR-Yes. Similarly, Figure 4b represents the PR evaluation of the DHODL-ADRRC method. The figures stated that the DHODL-ADRRC approach achieved maximal PR under two classes. At last, Figure 4c represent the ROC examination of the DHODL-ADRRC approach. The figure illustrated that the DHODL-ADRRC model had given an outcome in advanced results with higher values of ROC under 2 class labelling.

Figure 5 depicts the classifier outputs of the DHODL-ADRRC approach under the testing set. Figure 5a depicts the confusion matrices offered by the DHODL-ADRRC method on 30% of TSS. The figure denoted that the DHODL-ADRRC approach has identified 1244 instances under ADR-No and 261 instances under ADR-Yes. Similarly, Figure 5b illustrates the PR evaluation of the DHODL-ADRRC technique. The figures showed that the DHODL-ADRRC approach had achieved maximal PR under two classes. At last, Figure 5c depicts the ROC examination of the DHODL-ADRRC technique. The figure revealed that the DHODL-ADRRC model has productive outcomes with greater values of ROC under 2 class labelling.

To accentuate the ADR results of the DHODL-ADRRC method, the outputs are examined under two classes on 70 and 30 percent of TRS/TSS in Table 1 and Figure 6. The figure exhibited that the DHODL-ADRRC method has properly recognized two class labels. As a case, with TRS, the DHODL-ADRRC method attains an average $accu_y$ of 98.84%, $prec_n$ of 99.76%, $reca_l$ of 98.84%, $spec_y$ of 98.84%, F_{score} of 99.30%, and MCC of 98.60%. On the other hand, with TSS, the DHODL-ADRRC techniques approach average $accu_y$ of 98.44%, $prec_n$ of 99.30%, $reca_l$ of 98.44%, $spec_y$ of 98.44%, F_{score} of 98.86%, and MCC of 97.73%.

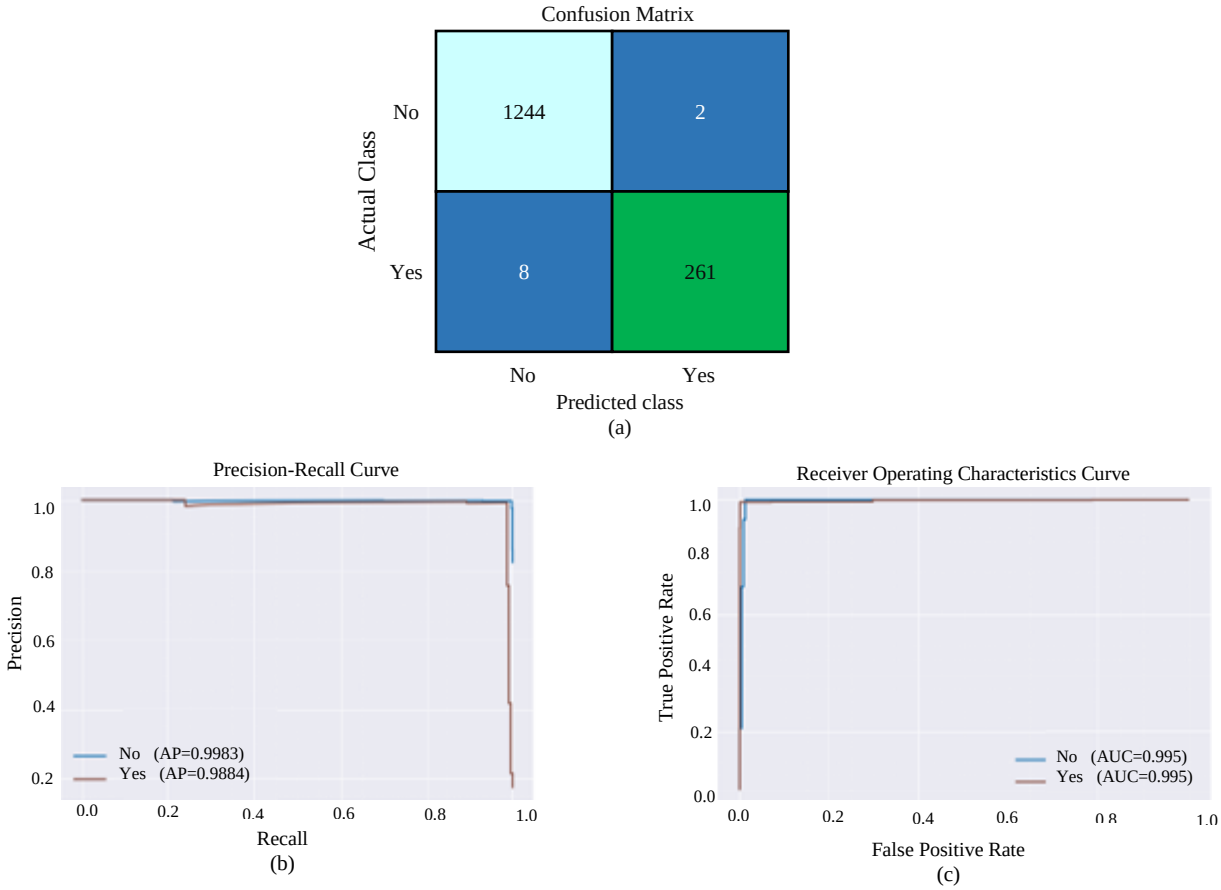

Fig. 5 Testing set a) Confusion Matrix b) PR-Curve c) ROC Analysis

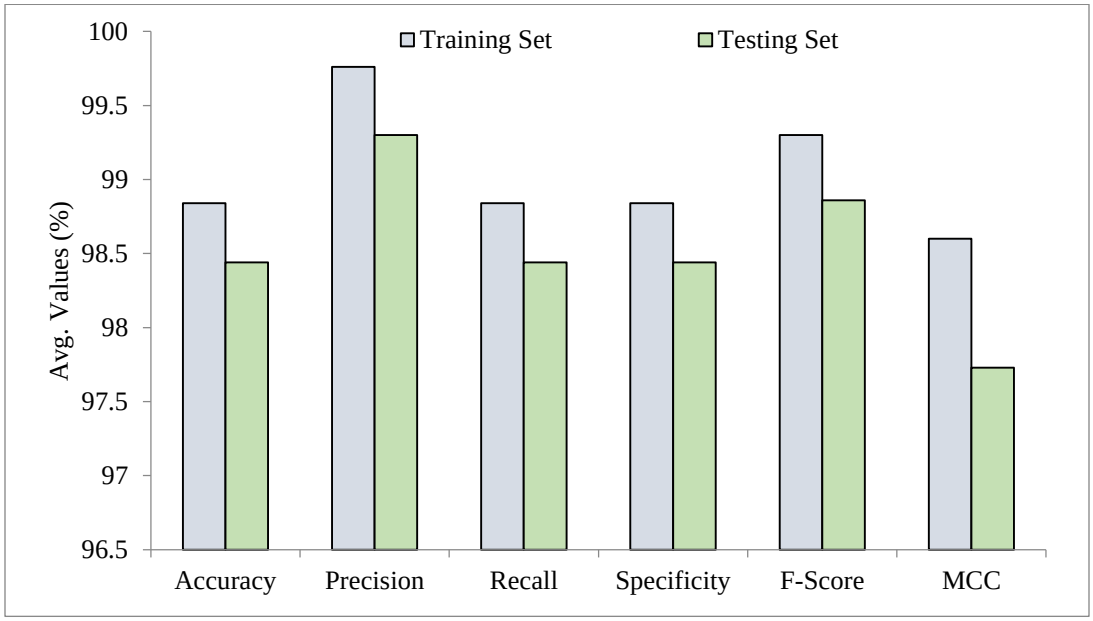

Fig. 6 Average output of DHODL-ADRRC method on 70:30 of TRS/TSS

Table 1. ADR output of DHODL-ADRRC technique on 70:30 of TRS/TSS

Class	$Accu_y$	$Prec_n$	$Reca_l$	$Spec_y$	F_{Score}	MCC
Training Set						
ADR-No	100.00	99.52	100.00	97.69	99.76	98.60
ADR-Yes	97.69	100.00	97.69	100.00	98.83	98.60
Average	98.84	99.76	98.84	98.84	99.30	98.60
Testing Set						
ADR-No	99.84	99.36	99.84	97.03	99.60	97.73
ADR-Yes	97.03	99.24	97.03	99.84	98.12	97.73
Average	98.44	99.30	98.44	98.44	98.86	97.73

Table 2. Comparative outcome of DHODL-ADRRC algorithm with other models

Methods	$Accu_y$	$Prec_n$	$Reca_l$	$Spec_y$
CNN Model	88.00	47.00	50.00	48.00
CRNN Model	85.00	38.00	53.00	44.00
RCNN Model	89.00	50.00	44.00	46.00
BERT-BCE	90.00	56.00	50.00	53.00
BERT-MSE	91.00	62.00	45.00	52.00
Balanced Bagging	81.00	32.00	67.00	43.00
Balanced Random Forest	73.00	26.00	76.00	39.00
DHODL-ADRRC	98.84	99.76	98.84	99.30

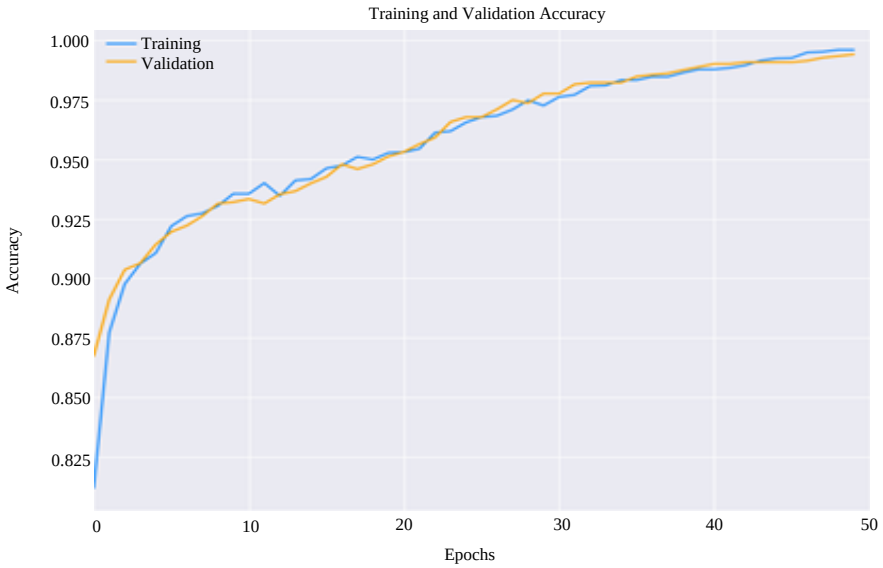

Fig. 7 Accuracy curve of the DHODL-ADRRC approach

Figure 7 investigates the $accu_y$ Of the DHODL-ADRRC method at the time of training and validation of the trial data. The figure indicates that the DHODL-ADRRC approach attains growing $accu_y$ values over increasing epochs. Moreover, the increasing validation $accu_y$ overtraining $accu_y$ depicts that the DHODL-ADRRC approach learns effectively on the trial data.

The loss examination of the DHODL-ADRRC approach at the time of training and validation is depicted in the trial data in Figure 8. The figure indicates that the DHODL-ADRRC technique attains closer training and validation loss values. The DHODL-ADRRC method learns effectively on the trial data.

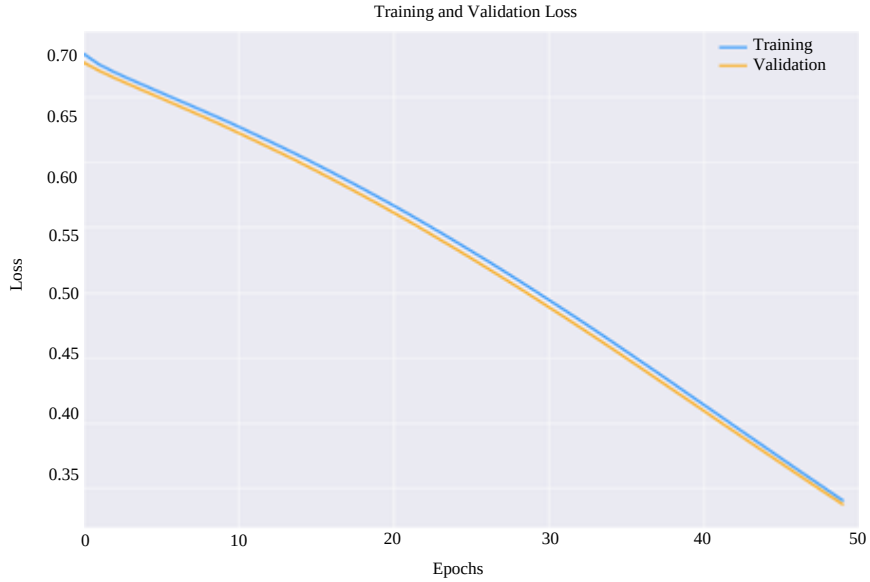

Fig. 8 Loss curve of the DHODL-ADRRC approach

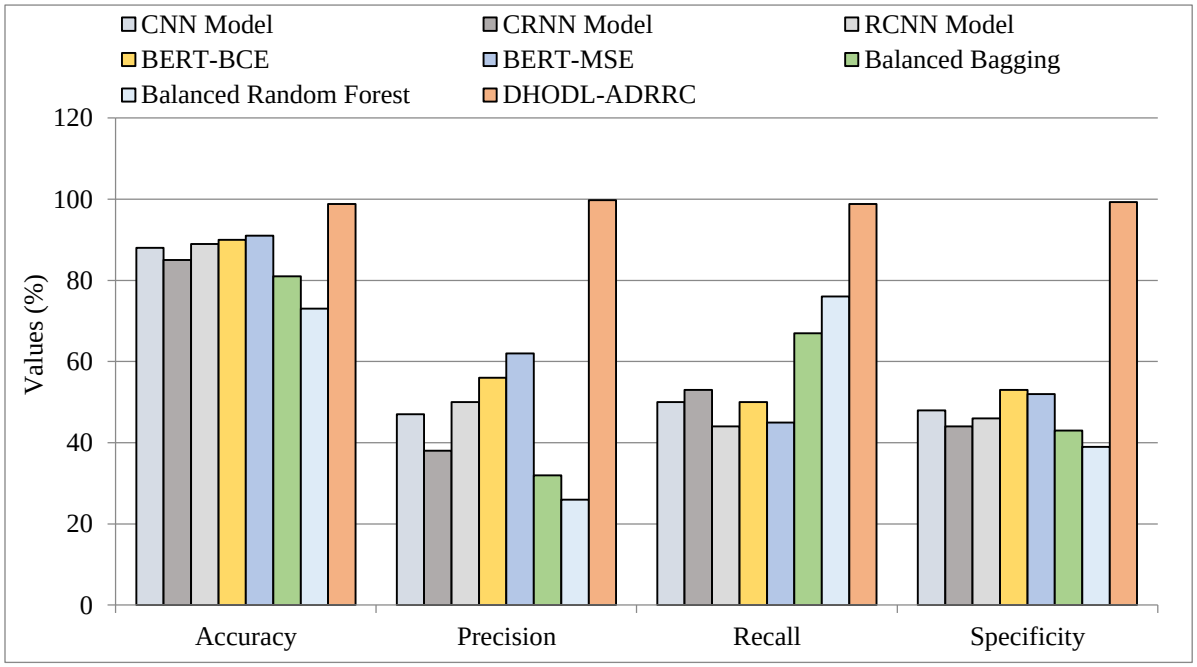

Fig. 9 Comparative outcome of DHODL-ADRRC algorithm with other methods

Table 2 and Figure 9 show a complete relative result of the DHODL-ADRRC technique is well studied [26-28]. The results identified the ineffectual performance of the CNN, CRNN, BB, and BRF models. At the same time, the RCNN, BERT-BCE, and BERT-MSE methods have reported indeed boosted ADR detection results. But the DHODL-ADRRC technique outperformed the other classifier models with an increased $accu_y$ of 98.84%, $prec_n$ of 99.76%, $reca_l$ of 98.84%, and F_{score} of 99.30%. These outputs show the advancement of the DHODL-ADRRC models in the ADR detection process.

5. Conclusion

In this article, we have presented the DHODL-ADRRC method for the ADR detection and classification process. The DHODL-ADRRC technique comprises four stages of operations, namely data preprocessing, BFHO-based FS, ACBLSTM-based ADR detection, and DHO-based hyperparameter tuning. Here, the DHODL-ADRRC technique involves diverse phases of data preprocessing to normalize the information. Next, the BFHO-based FS technique is used to elect an optimal set of features. Meanwhile, the ACBLSTM model is utilized for ADR

detection. Finally, the DHO technique is exploited for the tuning process adjusting of the ACBLSTM algorithm, which in turn advances the classification outputs. The experimental validation of the DHODL-ADRRC approach is investigated on the ADR data, and the extensive outputs highlighted the advanced achievement of the DHODL-ADRRC algorithm over other current approaches in terms of different measures.

References

[1] Zehong Zhang et al., "Graph Neural Network Approaches for Drug-Target Interactions," *Current Opinion in Structural Biology*, vol. 73, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[2] Chang Sun, "A Deep Neural Network-based Co-coding Method to Predict Drug-protein Interactions by Analyzing the Feature Consistency between Drugs and Proteins," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[3] Hossein Sharifi-Noghabi et al., "MOLI: Multi-Omics Late Integration with Deep Neural Networks for Drug Response Prediction," *Bioinformatics*, vol. 35, no. 14, pp. i501-i509, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[4] Natarajan Yuvaraj et al., "Analysis of Protein-Ligand Interactions of SARS-Cov-2 Against Selective Drug using Deep Neural Networks," *Big Data Mining and Analytics*, vol. 4, no. 2, pp. 76-83, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[5] Yang Li et al., "Drug Target Interaction Predication via Multi-Channel Graph Neural Networks," *Briefings in Bioinformatics*, vol. 23, no. 1, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[6] X. Xiao et al., "MPCDDI: A Secure Multiparty Computation-Based Deep Learning Framework for Drug-Drug Interaction Predictions," *Bioinformatics Research and Applications*, vol. 13760, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[7] Rabia Javed et al., "An Efficient Pattern Recognition Based Method for Drug-Drug Interaction Diagnosis," *2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)*, Riyadh, Saudi Arabia, pp. 221-226, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[8] Shichao Liu *et al*., "Enhancing Drug-Drug Interaction Prediction using Deep Attention Neural Networks," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 20, no. 2, pp. 976-985, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[9] E. G. Swetala, P. Sujatha, and P. Bharath, "Automatic Load Frequency Control for Wind-Thermal Micro Grid Based on Deep Reinforcement Learning," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 8, no. 8, pp. 1-8, 2021. [CrossRef] [Publisher Link]

[10] G. P. Dimf, P. Kumar, and K. Paul Joshua, "CNN with BI-LSTM Electricity Theft Detection based on Modified Cheetah Optimization Algorithm in Deep Learning," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 10, no. 2, pp. 35-43, 2023. [CrossRef] [Publisher Link]

[11] An Huang et al., "A Multimodal Data Fusion-Based Deep Learning Approach for Drug-Drug Interaction Prediction," *Bioinformatics Research and Applications: 18th International Symposium*, pp. 275-285, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[12] Xiaorui Su *et al*., "Biomedical Knowledge Graph Embedding with Capsule Network for Multi-Label Drug-Drug Interaction Prediction," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 6, pp. 5640-5651, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[13] Jaesub Park et al., Large-Scale Prediction of Adverse Drug Reactions-Related Proteins with Network Embedding," *Bioinformatics*, vol. 39, no. 1, pp. 1-9, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[14] Shanchen Pang et al., "AMDE: A Novel Attention-Mechanism-Based Multidimensional Feature Encoder for Drug–Drug Interaction Prediction," *Briefings in Bioinformatics*, vol. 23, no. 1, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[15] Nair Bini Balakrishnan, P. S. Sreeja, and Jisha Jose Panackal, "Alzheimer's Disease Diagnosis using Machine Learning: A Review," *International Journal of Engineering Trends and Technology*, vol. 71, no. 3, pp. 120-129, 2023. [CrossRef] [Publisher Link]

[16] Chi-Shiang Wang et al., "Detecting Potential Adverse Drug Reactions using A Deep Neural Network Model," *Journal of medical Internet Research*, vol. 21, no. 2, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[17] Brandon Fan et al., "Adverse Drug Event Detection and Extraction from Open Data: A Deep Learning Approach," *Information Processing & Management*, vol. 57, no. 1, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[18] Jiajing Zhu et al., "DAEM: Deep Attributed Embedding Based Multi-Task Learning for Predicting Adverse Drug–Drug Interaction," *Expert Systems with Applications*, vol. 215, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[19] Arnold K. Nyamabo et al., "Drug–Drug Interaction Prediction with Learnable Size-Adaptive Molecular Substructures," *Briefings in Bioinformatics*, vol. 23, no. 1, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[20] Christopher McMaster et al., "Developing a Deep Learning Natural Language Processing Algorithm for Automated Reporting of Adverse Drug Reactions," *Journal of Biomedical Informatics*, vol. 137, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] Chun Yen Lee, and Yi-Ping Phoebe Chen, "Prediction of Drug Adverse Events Using Deep Learning in Pharmaceutical Discovery," *Briefings in Bioinformatics*, vol. 22, no. 2, pp. 1884–1901, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[22] Fenia Christopoulou et al., "Adverse Drug Events and Medication Relation Extraction in Electronic Health Records with Ensemble Deep Learning Methods," *Journal of the American Medical Informatics Association*, vol. 27, no. 1, pp. 39-46, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[23] Mohammed Alonazi, and Mrim M. Alnfiai, "Fire Hawk Optimizer with Deep Learning Enabled Human Activity Recognition," *Computer Systems Science and Engineering*, vol. 45, no.3, pp. 3135–3150, 2023. [CrossRef] [Publisher Link]

[24] Jianqing Wu et al., "Towards Attention-Based Convolutional Long Short-Term Memory for Travel Time Prediction of Bus Journeys," *Sensors*, vol. 20, no. 12, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[25] S. R. Kanna, K. Sivakumar, and N. Lingaraj, "Development of Deer Hunting Linked Earthworm Optimization Algorithm for Solving Large Scale Traveling Salesman Problem," *Knowledge-Based Systems*, vol. 227, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[26] S. Sandhiya, and U. Palani, "*A Novel Hybrid PSBCO Algorithm for Feature Selection," International Journal of Computer and Organization Trends*, vol. 10, no. 3, pp. 21-26, 2020. [Google Scholar] [Publisher Link]

[27] Jhih-Yuan Huang,Wei-Po Lee, and King-Der Lee, "Predicting Adverse Drug Reactions from Social Media Posts: Data Balance, Feature Selection and Deep Learning," *Healthcare*, vol. 10, no. 4, pp. 618, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[28] Wongpanya S. Nuankaew et al., "Forecasting Graduation Schedule Model of Higher Education Learners using Feature Selection Techniques," *International Journal of Engineering Trends and Technology*, vol. 71, no. 4, pp. 354-358, 2023. [CrossRef] [Publisher Link]