Enhancing Adverse Drug Reaction Classification of Attention Deficit Hyperactivity Disorder Diagnosis Data Using Deep Learning with Optimization Algorithm

N. Deepaletchumi[1,*], R. Mala[2]

[1]Research Scholar, Department of Computer Applications, Alagappa University, Karaikudi, India

[2]Asst.Prof & Head, Department of Computer Science, Government Arts and Science College for Women, Paramakudi, India

Emails: deepawaran86@gmail.com; murugan.dcdrf@gmail.com

Abstract

Adverse Drug Reaction (ADR) is a significant global public health issue and the main cause of death. Generally, the effects of ADR are complex. Clinically, they can cause major patient damage and, in some cases, death. Besides, this outcome in significant healthcare costs financially owing to enlarged hospital visits, extra treatments, and harm to productivity. Therefore, early recognition and mitigation of ADRs are vital for the patients. Enhancing the early detection of ADRs and deadliness could severely reduce the harm to patients, improve patient safety, decrease healthcare costs, and increase the efficacy of the drug development procedure. Conventional pre-clinical toxicity tests are expensive, time-consuming, and frequently fail to forecast human-specific toxic effects. Artificial Intelligence (AI)-based deep learning (DL) has been quickly adopted in numerous areas, with healthcare, for its latent to manage huge datasets, find out patterns, and generate predictions. This study presents a new Adverse Drug Reaction Detection through Deep Learning and Improved Red-Tailed Hawk Algorithm (ADRD-DLIRTHA). The main intention of the ADRD-DLIRTHA model is to enhance the detection and classification process of ADR using advanced hybrid and optimization techniques. At first, the data normalization stage applies z-score normalization for converting input data into a beneficial set-up. Furthermore, the proposed ADRD-DLIRTHA method designs a convolutional neural network and long short-term memory (CNN-LSTM) technique for the classification process. At last, the improved red-tailed hawk (IRTH) algorithm-based hyperparameter selection process has been applied to optimize the classification results of the CNN-LSTM system. A wide range of experimentation was led to authorize the performance of the ADRD-DLIRTHA system. The simulation results specified that the ADRD-DLIRTHA model emphasized advancement over other existing techniques

Keywords: Adverse Drug Reaction; Deep Learning; Improved Red-Tailed Hawk Algorithm; Data Normalization; Artificial Intelligence

1. Introduction

An adverse drug reaction (ADR) can be viewed as some noxious, undesired, and unintended impact of drug, without therapeutic failures, drug abuse, or accidental and intentional poisoning [1]. ADR comprising drug interactions, in elderly people has greater significance in medical care concerns, often admittance to hospital, arises generally in hospitalization, and is a significant cause of mortality and morbidity [2]. Old peoples are mainly vulnerable to ADRs not only age-associated variations in pharmacodynamics and pharmacokinetics but also due to greater incidence of disability, co-morbidity, and numerous drug regimens [3]. Furthermore, ADRs in older people may have serious clinical and financial concerns [4]. These effects comprise an increased risk of severe illness, with institutionalization, possible longer-term disability, and even death that improves expenditure for healthcare [5].

Currently, the recognition of ADR has dual models namely passive monitoring depends upon the spontaneous reporting system (SRS) and active surveillance depends on electronic clinical data mining like case reports,

Doi: https://doi.org/10.54216/FPA.190104
Received: October 25, 2024 Revised: January 03, 2025 Accepted: January 30, 2025

biomedical literature, and the mining of electronic health records (EHRs) [6]. The government-backed SRSs agree on voluntary reports from pharmaceutical companies, healthcare professionals, and consumers. Among these studies, user reports comprise a huge proportion of all reports [7]. Furthermore, owing to untrained customers and time lag, SRSs suffered from difficulties like heavy underestimation, the over-reporting of known ones, duplicate reporting, and incomplete data. Unfortunately, as a complementary and alternative model, the secondary usage of clinical data is plagued by complications in the admittance of patient data and in evading privacy and legal concerns [8]. Generally, SRS data have a tendency to focus nearby signal recognition utilizing the reported ratio of the statistical model [9]. Nevertheless, statistical models are restricted on the subject of inspecting free textual or structure of chemical data for signal recognition. Thus, Deep Learning (DL) and Machine Learning (ML) methodologies are developed to make investigation of these kinds of data possible for ADR signal recognition [10]. DL and ML-based models are favorable to upgrade our capability to predict and detect ADR and toxicity.

This study presents a new Adverse Drug Reaction Detection Through Deep Learning And Improved Red-Tailed Hawk Algorithm (ADRD-DLIRTHA). At first, the data normalization stage applies z-score normalization for converting input data into a beneficial set-up. Furthermore, the proposed ADRD-DLIRTHA method designs a CNN-LSTM technique for the classification process. At last, the improved red-tailed hawk (IRTH) algorithm-based hyperparameter selection process has been carried out to optimize the classification results of the CNN-LSTM system. A wide range of experimentation was led to authorize the performance of the ADRD-DLIRTHA system.

2. Review of Literature

Spandana and Prakash [11] developed to accomplish ADR recognition from social media-related comments. Convention models are accomplished with specific types of drugs for only a few local areas globally. To overwhelm this, a DCNN-based ADR framework is projected with different training for better accuracy and flexibility. Jiang et al. [12] projected an HMGRL method. The RGCN model is utilized to acquire multiple explicit relationships among drugs from these heterogeneous graphs. Furthermore, the MVDSC module is presented. In the MVDSC, diverse DP aspects are employed to create graphs, where nodes specify DPs and edges represent diverse implicit co-relations. Afterward, several DP representations are created over graph cutting, every highlighting different implicit co-relations. Li et al. [13] project an ADR forecasting method that depends on DL and knowledge graph embedding that can forecast stimulated outcomes. This paper develops an ADR knowledge graph derived from drug feature data; the finest embedded approach is chosen to attain sample vectors; and then a CNN method is made to forecast adverse reactions.

Dey et al. [14] projected a classical-quantum hybrid DL method for identifying negative impact of drugs from social media analyses. The VQC encrypts conventional data and categorizes input analyses into feature vectors. In [15], an innovative structure named PreciseADR for patient-level ADR prediction is projected. This method effectually incorporates relationships among ADRs and patients and utilizes the ability of heterogeneous GNN. By utilizing patient embedding which considers their drugs and diseases, possible ADRs might be precisely forecast. Lee et al. [16] present a DL method that classifies drug ADR posts depending on RNN. Posts containing considerably related ADR word pairs and drug names were filtered over analysis of association, and training data was created over physical labeling tasks. Subsequently, the region under the curve is assessed with other ML approaches. Moreover, the overall process is verified utilizing the non-steroidal anti-inflammatory drug aceclofenac. Ahmed et al. [17] developed the consequences of 3 different methods: RNN, DNN, and RF in the early filtration round to discover a proper candidate method to forecast the severity and risk of negative outcomes that might be lifted while drug A is integrated with B.

3. Methodology

In this study, we have introduced a new ADRD-DLIRTHA methodology. The main intention of the ADRD-DLIRTHA model is to enhance the detection and classification process of ADR using advanced hybrid and optimization techniques. It contains three various kinds of processes involved data normalization, hybrid classification, and hyperparameter selection. Fig. 1 represents the complete procedure of ADRD-DLIRTHA method.

Doi: https://doi.org/10.54216/FPA.190104
Received: October 25, 2024 Revised: January 03, 2025 Accepted: January 30, 2025

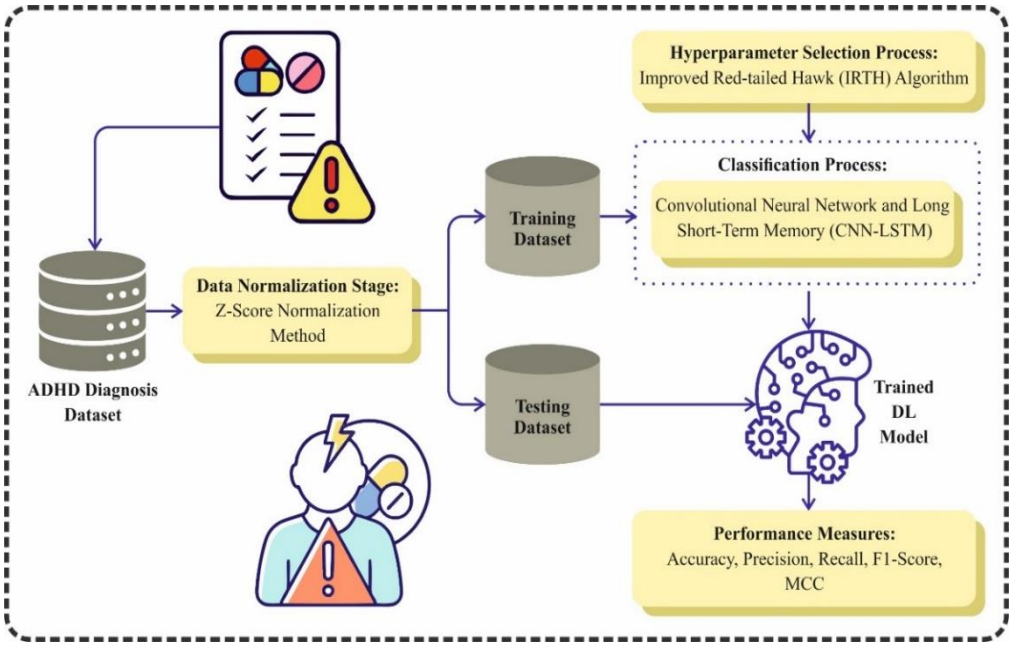

Figure 1. Overall process of ADRD-DLIRTHA system

A. Z-score Normalization

At first, the data normalization stage applies z-score normalization for converting input data into a beneficial format. The Z-score is a measure that characterizes how much a data point is from the standard deviation regarding mean [18]. It mathematically estimates the comparative location of a particular value inside the distribution and can identify abnormalities according to deviations also scaling permits the column's normalization from dissimilar classes, making it assured to average or sum the scores. The benefit of this score is that it streamlines data types, enabling calculation and comparison. It further allows the discovery of relation locations inside the data and is applied in statistical study. This equation is as presented in Eq. (1). Z characterizes the outcome of utilizing the Z-score, while X denotes data value, μ stands for mean, and σ refers to standard deviation.

$$Z = \frac{(X - \mu)}{\sigma} \tag{1}$$

B. Hybrid Classification Model

Furthermore, the proposed ADRD-DLIRTHA method designs the CNN-LSTM technique for the classification process. CNN leads to the DL method for various application-solving purposes [19]. It comprises 4 diverse layers: activation, convolution, fully connected, and pooling layers. These layers are combined to remove aspects from the input data and categorize them into diverse classes. Convolution layer that extracts spatial characteristics, activation layer that enhances the learning within pooling, non-linear layers, which decrease the dimension counts, and fully connected layers, that integrate local and global aspects, form the CNN structure. The convolution layer connecting the input x_i^k and output y_i^k, executes matrix multiplication among the preceding input layer and filter bank w_{ij}^k, it also contains the addition of a regularization term (bias) b_j^k.

$$y_i^k = \sum_i (x_i^k * w_{ij}^k) + b_j^k \tag{2}$$

Within the non-linear layer, a leaky rectified linear unit (LeakyRelu) is employed as the function of activation. LeakyRelu varies from Relu by precluding the removal of negative inputs, therefore addressing the dying Relu concern. The mathematical of LeakyRelu is given.

$$f(x)LeakyRelu = \begin{cases} x, if\ x & > 0 \\ mx, x & \leq 0 \end{cases} \tag{3}$$

The leak factor specified by m is a small value repeatedly set to 0.001, presenting a slight negative slope for the function of activation. It enables a minor flow of data while the input is negative, assisting in process of learning. During the maximal pooling layer, the superior value in a given matrix is chosen. The fully connected (FC) layer may combine a non-linear activation function such as LeakyRelu. During classification tasks, the output layer is

Doi: https://doi.org/10.54216/FPA.190104
Received: October 25, 2024 Revised: January 03, 2025 Accepted: January 30, 2025

connected to activation function of SofttMax, although in regression tasks, a linear regression function is employed. The linear regression and SofttMax functions are shown by.

$$\sigma(z)i = \frac{e^{zi}}{\Sigma_{j=1}^{K} e^{zj}} \; for \; i = 1, \ldots\ldots\ldots, k \; and \; z = (z1, \ldots\ldots\ldots zk) \in R^K \quad (4)$$

Recurrent neural networks (RNN) have a framework named LSTM. Data is transferred in diverse cells of the hidden layer over some controllable gates. It permits the memory and forgetting range of the preceding and existing information to be inspected. In comparison with conventional LSTM, RNN has a longer-term memory function and evades the gradient disappearance concern. The LSTM integrates dual gates that control the state of memory cell. The forget gate measures the range to which memory from the preceding time step is recollected, whereas the input gate establishes the number of existing inputs that are stored in the cell state. Eventually, the LSTM system comprises an output gate that controls the number of information that is emitted relating to the status of cell.

$$f(t) = \sigma(wf.[ht-1, xt] + bf) \quad (5)$$

$$I(t) = \sigma(wf.[ht-1, xt] + bi) \quad (6)$$

$$\tilde{c}(t) = Tanh(wc.[ht-1, xt] + bc) \quad (7)$$

$$c(t) = fi * ct - 1 + It * \tilde{c}t \quad (8)$$

$$o(t) = \sigma(wo.[ht-1, xt] + bo) \quad (9)$$

$$h(t) = ot * \tanh(c1) \quad (10)$$

$$sigmoid\ (x) = \frac{1}{1+e^{-x}} \quad (11)$$

$$tanh\ (x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \quad (12)$$

The hyperbolic tangent function in eq. (12) is utilized to overwhelm the gradient disappearance problem. Eqs. (5)-(12) define the output and input of the network framework of the LSTM. Fig. 2 describes the structure of CNN-LSTM.

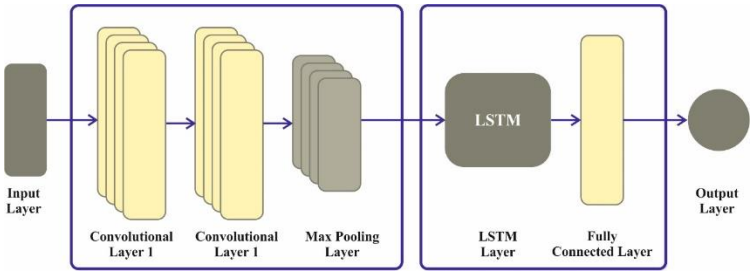

Figure 2. Structure of CNN-LSTM

C. Parameter Tuning using IRTH

At last, the IRTH algorithm-based hyperparameter selection process is performed to optimize the classification results of the CNN-LSTM system. The unique RTH model implements better unique-peaked functions and possesses basic architectures; still, it is inclined to difficulties like dropping into local optimizer after meeting composite real-time optimizer problems [20]. To overwhelm these issues, an enhanced form of the RTH model according to the trusted domain is presented in this section. The particulars are as demonstrated.

1) A Stochastic Reverse Learning Approach-based Bernoulli Mapping

During this RTH model, the primary population is achieved by arbitrary activation. This arbitrary primary population frequently results in a further distributed spreading of individuals within the solution area that lack target explorations of the possible best solution area. Owing to the individual's chance in the arbitrary primary population, that is probable that the majority of the folks accumulate in the area neighboring the local best solution at the start creating the exploration low effective. This mapping is presented to enhance the RTH model that maintains the new uncertainty while improving the use of previous information to improve the implementation of the model.

Doi: https://doi.org/10.54216/FPA.190104
Received: October 25, 2024 Revised: January 03, 2025 Accepted: January 30, 2025

Bernoulli transition mapping is the probability-based transition method, which converts the specific condition according to the Bernoulli dispersion. Computationally this dispersion is a distinct likelihood delivery, and they enhance it by utilizing dual-part linear mappings, as presented in Eq. (13).

$$X_{n+1} = \begin{cases} \frac{x_n}{1-\sigma}' 0 < x_n \leq 1-\sigma \\ \frac{x_n - (1-\sigma)}{\sigma}, 1-\sigma < x_n \leq 1' \end{cases} \quad (13)$$

Whereas σ is assigned to 0.4. Reverse learning is an approach, which increases the searching ability of a model by reflecting the present solution and its reverse. Besides manipulating and assessing the standard solutions, the reverse solutions of this solution are completed, and by opposing the values of fitness of the new and the reverse solutions, the best solution has been chosen to continue to the upcoming iteration rounds. When the reverse solution fitness is superior to that of the novel solutions, therefore the new solution was substituted by the reverse solution that permits the model to possibly leap out of the local optimal as the reverse solution might be discovered in the further best area outer the present searching area. The explanation of the inverse solution is presented in Eq. (14):

$$OBL_i = K \odot (MAX + MIN) - X_i \quad (14)$$

Whereas OBL_j signifies the opposing solution gained by the element X_i Later refractive opposing mutual learn; K refers to matrix of single row dim columns, while the components within the matrix are randomly generated values among (0,1); and MIN and MAX represent minimal and maximal amounts of the individual, individually. Once the value of fitness of the oppositional solution is superior to the new solutions, the novel solution is upgraded to the oppositional solution; or else, no upgrade is implemented. This can be presented in Eq. (15).

$$X_i = \begin{cases} OBL_j, f(OBL_i) < f(X_i) \\ X_i, other \end{cases} \quad (15)$$

Whereas $f(OBL_i)$ represents the fitness value of the oppositional solution and $f(X_i)$ symbolizes the novel solution's fitness value.

2) Dynamic Position Update Optimization Approach for Stochastic Mean Fusion

During this RTH model, $randperm$ has been applied for randomizing the individual orders within the population, and this arbitrariness aids in presenting a particular extent of range but additionally presents ambiguity. All times the model is running, the change in the random plan might result in a larger change in the convergence route and last outcome of the model. Besides this, the precise summation of step size might not be well suited to various optimizer problem states. For some composite objective functions with dissimilar scale features, an additional adaptable stepsize fine-tuning approach can be required to guarantee that the model can meet more generally. Then, presented a dynamic location updated optimizer approach for random mean fusion that guarantees that the model might well aggregate to the global optimum solution over dual dissimilar stepsize calculation approach, and utilizes adaptive parameter to enhance the speed of convergence of the model responding to the quick varying as atmospheres and issues. The enhanced location upgrade is presented in Eq. (16).

$$X(t) = \begin{cases} X_{best} + \alpha * step1.* R, rand > 0.5 \\ X_{best} + \alpha * step2.* R, other \end{cases} \quad (16)$$

Whereas R means $randn(1, dim)$; α refers to adaptive parameters that can be designed by Eq. (17); step1, and step2 are dual dissimilar dynamically changing step sizes to allow the model to well familiarize itself with dissimilar optimizer problem setups, which are computed by Eqs. (18) and (19).

$$\alpha = (1 - \frac{t}{T_{\max}})^{2*\frac{t}{T_{\max}}} \quad (17)$$

$$step1 = X_{mean1} - X(t) \quad (18)$$

$$step2 = X_{mean2} - X(t) \quad (19)$$

whereas X_{mean1} is gained by capturing a randomly generated number p among (5,10) and then arbitrarily choosing p individuals within the population to make a novel sub-population X_p, and then discover the mean of X_p as X_{mean1}. X_{mean2} is gained by seizing an arbitrary number q amongst 15 and N and then arbitrarily choosing q individuals within the population to create a novel sub-population X_q, and then discover the mean of X_q as X_{mean2}.

3) Optimization Model for Frontier Location Update-based Trust Domain

During this swooping and stooping phase, it is important for the model to aggregate to the best solution as soon as promising however further retain a few of the explorations to stop the model from dropping into the local optimal.

Doi: https://doi.org/10.54216/FPA.190104
Received: October 25, 2024 Revised: January 03, 2025 Accepted: January 30, 2025

During this RTH model, it is comparatively challenging to utilize mixtures of G, and so on. During unique-peak function optimizer issues, these composite groupings are over-complicated to calculate, resulting in an excess of computation power; in multifaceted multiple-peak function optimizer problems, they might not be efficiently adjusted to the modifications in the function land, succeeding challenging to precisely director the searching towards the globally best solutions. The trusted domain model can successfully stop the problem of the model divergence attributable to too big a stepsize by estimating the objective function in a trusted domain and discovering the optimum solution in these areas. Since the model repeats, the trusted domain's radius is enthusiastically attuned based on the alteration in the fitness value to enhance the global convergence performance of the model. Furthermore, the adapted change of the radius can further adjust to various setups earlier, such that the model can well perform in all types of problems and functions. During this IRTH method, the trusted domain-based optimizer model for the border location upgrade is established by Eq. (20).

$$X(t) = \begin{cases} X(t) + \left(X_{mean1} - TDX(R1)\right).* rand, & rand > 0.5 \\ X(t) + \left(X_{mean2} - TDX(R1)\right).* rand, & other \end{cases} \quad (20)$$

Whereas $R1$ denotes a randomly generated number, which represents an arbitrarily chosen individual in the trusted domain for the location upgrade of the supplementary model. TDX represents the population of the trusted domain that comprises all individuals situated in the trusted domain. In all rounds of location upgrade, we arbitrarily choose an individual in the trusted domain for location upgrade, which guarantees the speed of convergence of the model however maintaining the uncertainty, and better increases the model's performance. Once the fitness value of this novel solution is superior to the fitness value of new solutions, this designates that there is an issue by the location upgrade, and they are involved in improving the radius of the trusted domain to create the model superior for exploring. After the fitness value of the novel solution is lower than the value of fitness of new solutions, it denotes that the location upgrade of the model is efficient, and then they decrease the radius of the trusted domain to generate the model meet quicker.

The IRTH system improves a fitness function (FF) to take on developed classifier performance. It concludes with an optimistic number to signify the improved efficacy of candidate solution. At this point, the classifier rate of error minimization is reflected as FF, as formulated in Eq. (21).

$$fitness(x_i) = ClassifierErrorRate(x_i) = \frac{no.of\ misclassified\ samples}{Total\ no.of\ samples} * 100 \quad (21)$$

4. Experimental Validation

In this portion, the investigational validation of the ADRD-DLIRTHA technique is verified below ADHD diagnosis database [21]. The dataset contains 103 instances of dual class labels as presented in Table 1.

Table 1: Dataset details

Classes	No. of Instances
Class_ADHD (Attention Deficit Hyperactivity Disorder)	51
Class_Normal (Clinically Controls)	52
Total number of Instances	103

Fig. 3 establishes the confusion matrix made by the ADRD-DLIRTHA approach under 80:20 and 70:30 of TRASE/TESSE. The outcomes identify that the ADRD-DLIRTHA approach is capable detection and identification of dual classes specifically.

Doi: https://doi.org/10.54216/FPA.190104
Received: October 25, 2024 Revised: January 03, 2025 Accepted: January 30, 2025

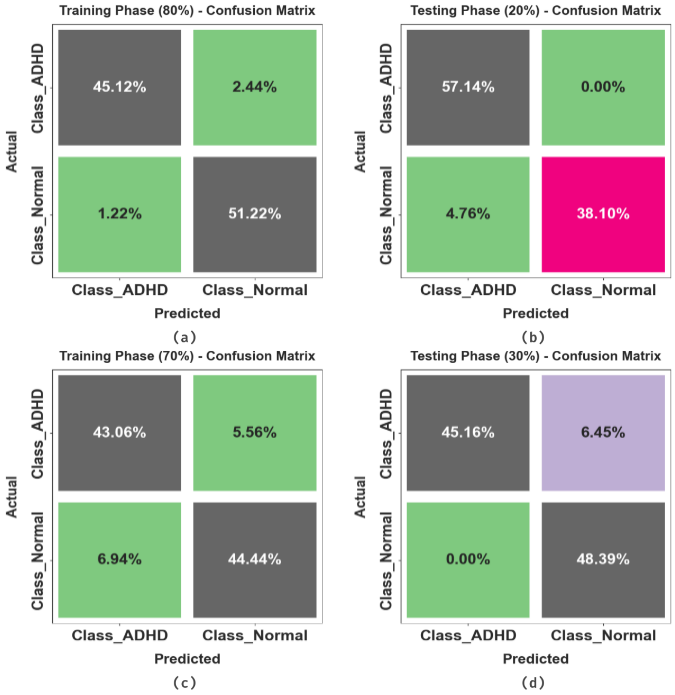

Figure 3. 80% and 20% of (a-b) confusion matrix and (c-d) curves of PR and ROC

Table 2 and Fig. 4 signify the ADR detection of ADRD-DLIRTHA methodology below 80:20 and 70:30 of TRASE/TESSE. The results suggest that the ADRD-DLIRTHA methodology correctly identified the samples. With 80%TRASE, the ADRD-DLIRTHA system offers average $accu_y$, $prec_n$, $reca_l$, $F1_{score}$, and MCC of 96.34%, 96.41%, 96.27%, 96.33%, and 92.68%, correspondingly.

Table 2: ADR detection of ADRD-DLIRTHA methodology below 80:20 and 70:30 of TRASE/TESSE

Class Labels	$Accu_y$	$Prec_n$	$Reca_l$	$F1_{score}$	MCC
TRASE (80%)					
Class_ADHD	96.34	97.37	94.87	96.10	92.68
Class_Normal	96.34	95.45	97.67	96.55	92.68
Average	96.34	96.41	96.27	96.33	92.68
TESSE (20%)					
Class_ADHD	95.24	92.31	100.00	96.00	90.58
Class_Normal	95.24	100.00	88.89	94.12	90.58
Average	95.24	96.15	94.44	95.06	90.58
TRASE (70%)					
Class_ADHD	88.57	86.11	88.57	87.32	75.03
Class_Normal	86.49	88.89	86.49	87.67	75.03
Average	87.53	87.50	87.53	87.50	75.03
TESSE (30%)					
Class_ADHD	87.50	100.00	87.50	93.33	87.87
Class_Normal	100.00	88.24	100.00	93.75	87.87
Average	93.75	94.12	93.75	93.54	87.87

Doi: https://doi.org/10.54216/FPA.190104
Received: October 25, 2024 Revised: January 03, 2025 Accepted: January 30, 2025

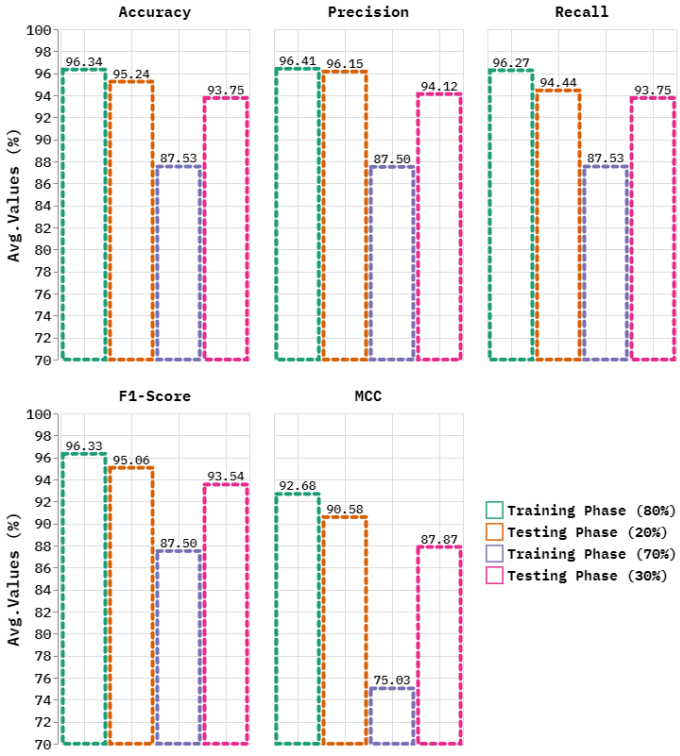

Figure 4. Average of ADRD-DLIRTHA methodology below 80:20 and 70:30 of TRASE/TESSE

Moreover, with 20%TESSE, the ADRD-DLIRTHA approach provides average $accu_y$, $prec_n$, $reca_l$, $F1_{score}$, and MCC of 95.24%, 96.15%, 94.44%, 95.06%, and 90.58.68%, respectively. Also, with 70%TRASE, the ADRD-DLIRTHA system delivers average $accu_y$, $prec_n$, $reca_l$, $F1_{score}$, and MCC of 87.53%, 87.50%, 87.53%, 87.50%, and 75.03%, correspondingly. Besides, with 30%TESSE, the ADRD-DLIRTHA methodology offers average $accu_y$, $prec_n$, $reca_l$, $F1_{score}$, and MCC of 93.75%, 94.12%, 93.75%, 93.54%, and 87.87%, respectively.

In Fig. 5, the training (TRA) $accu_y$ and validation (VAL) $accu_y$ results of the I ADRD-DLIRTHA methodology under 80%TRASE and 20%TESSE are illustrated. The $accu_y$ values are intended across an interval of 0-50 epochs. The result highlights that the TRA and VAL $accu_y$ analysis exhibitions growing tendency that knowledgeable the abilities of the ADRD-DLIRTHA technique with greater outcomes across multiple iterations. Simultaneously, the TRA and VAL $accu_y$ remains closer across the epochs, which indicates inferior overfitting and exhibition of maximum outcomes of the ADRD-DLIRTHA technique, ensuring dependable prediction on hidden samples.

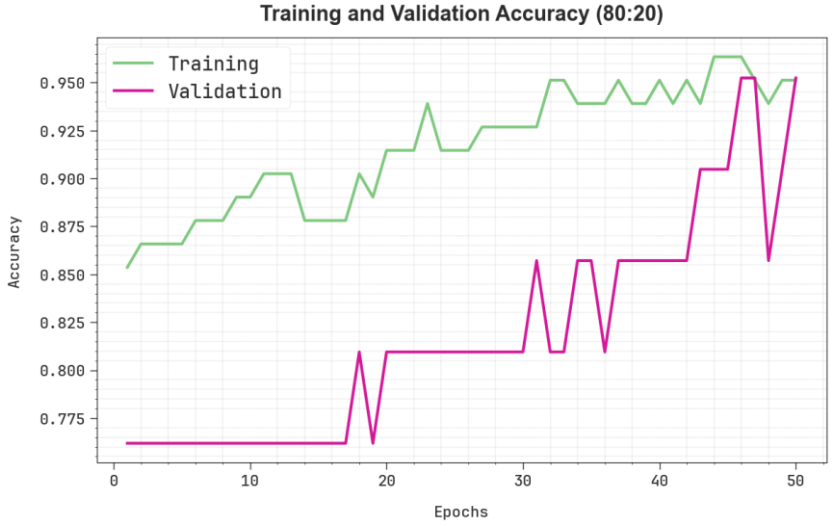

Figure 5. $Accu_y$ curve of ADRD-DLIRTHA model under 80%TRASE and 20%TESSE

Doi: https://doi.org/10.54216/FPA.190104
Received: October 25, 2024 Revised: January 03, 2025 Accepted: January 30, 2025

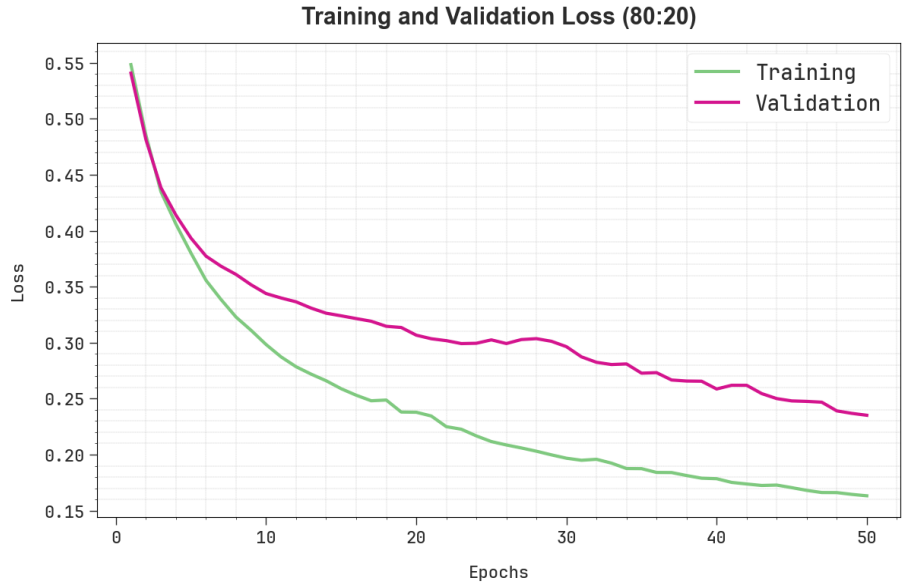

Figure 6. Loss graph of ADRD-DLIRTHA methodology below 80%TRASE and 20%TESSE

In Fig. 6, the TRA loss (TRALOS) and VAL loss (VALLOS) curve of the ADRD-DLIRTHA algorithm under 80%TRASE and 20%TESSE is shown. The values of loss are computed across an interval of 0-50 epochs. It is implied that the TRALOS and VALLOS analysis exemplify a diminishing trend, notifying the proficiency of the ADRD-DLIRTHA approach in balancing a exchange amongst simplification and data fitting. The continuous fall in values of loss assurances the maximal solution of the ADRD-DLIRTHA methodology and adjusts the predictive results in time.

In Fig. 7, the PR graph outcomes of the ADRD-DLIRTHA methodology below 80%TRASE and 20%TESSE offer clarification into its outcomes by plotting Precision beside Recall for 2 classes. The outcome reveals that the ADRD-DLIRTHA system continually attains superior PR analysis over different classes, demonstrating its capacity to keep up an important section of true positive predictions between each positive prediction (precision) which besides picking up a enormous proportion of real positives (recall). The steady upsurge in PR values amongst each class label describes the effectiveness of the ADRD-DLIRTHA system in the classification procedure.

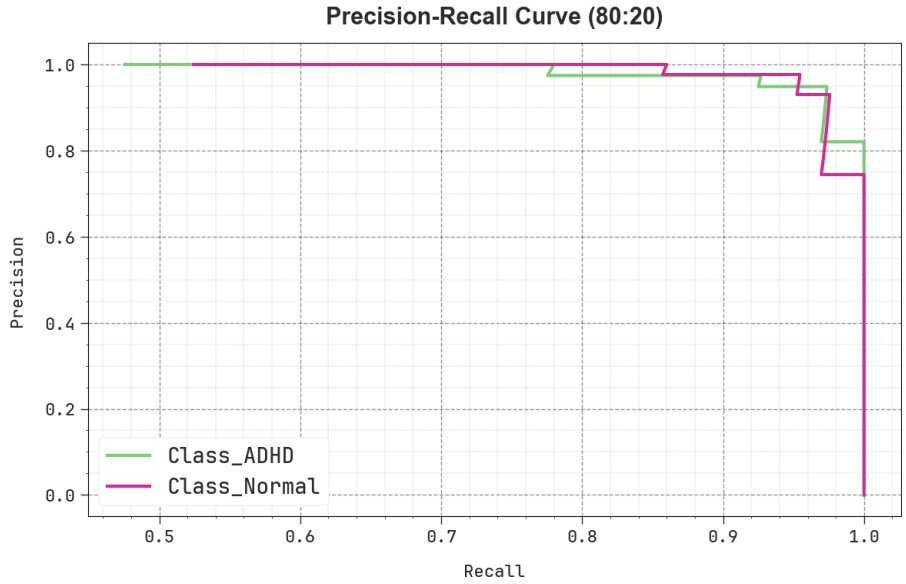

Figure 7. PR analysis of ADRD-DLIRTHA approach below 80%TRASE and 20%TESSE

Doi: https://doi.org/10.54216/FPA.190104
Received: October 25, 2024 Revised: January 03, 2025 Accepted: January 30, 2025

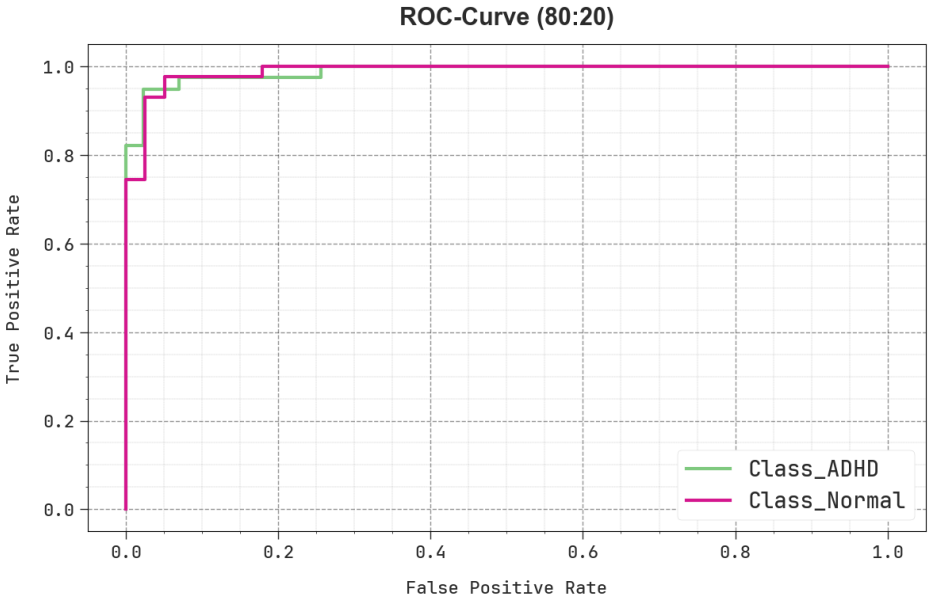

Figure 8. ROC graph of ADRD-DLIRTHA technique below 80%TRASE and 20%TESSE

In Fig. 8, the ROC graph of the ADRD-DLIRTHA algorithm below 80%TRASE and 20%TESSE is examined. The outcomes suggest that the ADRD-DLIRTHA system achieves superior ROC results across all classes, establishing important capacity for differentiating the classes. This dependable inclination of greater ROC analysis across numerous classes means the capable performance of the ADRD-DLIRTHA algorithm in predicting classes, emphasizing the robust nature of the classification process.

Table 3 and Fig. 9 inspect the comparison outcomes of the ADRD-DLIRTHA system with the existing techniques [22-24]. The results highlight that the AdaBoost, GB, RF, ME-TFIFDF, CRNN, FPGNN-SDAJM, and ResNet approaches have reported inferior performance. Furthermore, the proposed ADRD-DLIRTHA method reported optimal performance with maximal $accu_y$, $prec_n$, $reca_l$, and $F1_{score}$ of 96.34%, 96.41%, 96.27%, and 96.33%, respectively.

Table 3: Comparative outcomes of ADRD-DLIRTHA model with existing algorithms

Algorithm	$\boldsymbol{Accu_y}$	$\boldsymbol{Prec_n}$	$\boldsymbol{Reca_l}$	$\boldsymbol{F1_{score}}$
AdaBoost	85.17	91.28	85.86	95.13
Gradient Boosting	85.16	85.68	87.87	90.91
Random Forest	95.16	85.66	86.64	93.03
ME-TFIFDF	90.01	95.37	93.77	95.30
CRNN Model	95.23	92.96	93.61	92.48
FPGNN-SDAJM	92.92	88.32	89.45	89.06
ResNet Method	92.42	91.85	94.38	86.63
ADRD-DLIRTHA	96.34	96.41	96.27	96.33

Doi: https://doi.org/10.54216/FPA.190104
Received: October 25, 2024 Revised: January 03, 2025 Accepted: January 30, 2025

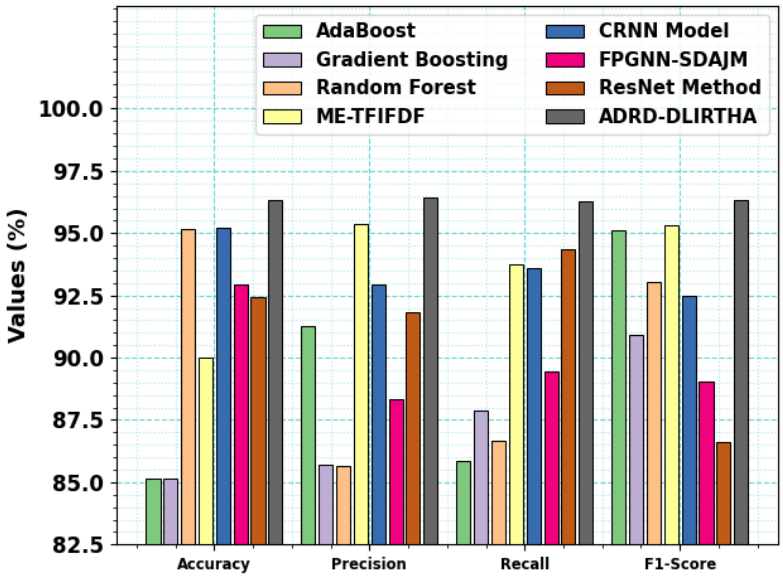

Figure 9. Comparative results of ADRD-DLIRTHA system with existing algorithms

5. Conclusion

In this manuscript, we have introduced a new ADRD-DLIRTHA methodology. The main objective of ADRD-DLIRTHA model is to enhance detection and classification process of ADR using advanced hybrid and optimization techniques. It contains three various kinds of processes involved data normalization, hybrid classification, and hyperparameter selection. At first, the data normalization stage applies z-score normalization for converting input data into a beneficial format. Furthermore, the proposed ADRD-DLIRTHA method designs CNN-LSTM technique for the classification process. At last, the IRTH algorithm-based hyperparameter selection process has been carried out to optimize the classification results of the CNN-LSTM system. A wide range of experimentation was led to authorize the performance of ADRD-DLIRTHA system. The simulation results specified that the ADRD-DLIRTHA model emphasized advancement over other existing techniques.

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] Gupta, R., Srivastava, D., Sahu, M., Tiwari, S., Ambasta, R.K. and Kumar, P., 2021. Artificial intelligence to deep learning: machine intelligence approach for drug discovery. Molecular diversity, 25, pp.1315-1360.

[2] Choudhury, O., Park, Y., Salonidis, T., Gkoulalas-Divanis, A., Sylla, I. and k Das, A., 2020, March. Predicting adverse drug reactions on distributed health data using federated learning. In AMIA Annual symposium proceedings (Vol. 2019, p. 313).

[3] Basiri, M.E., Abdar, M., Cifci, M.A., Nemati, S. and Acharya, U.R., 2020. A novel method for sentiment classification of drug reviews using fusion of deep and machine learning techniques. Knowledge-Based Systems, 198, p.105949.

[4] Khalil, H. and Huang, C., 2020. Adverse drug reactions in primary care: a scoping review. BMC health services research, 20, pp.1-13.

[5] Malki, M.A. and Pearson, E.R., 2020. Drug–drug–gene interactions and adverse drug reactions. The pharmacogenomics journal, 20(3), pp.355-366.

[6] Basile, A.O., Yahi, A. and Tatonetti, N.P., 2019. Artificial intelligence for drug toxicity and safety. Trends in pharmacological sciences, 40(9), pp.624-635.

[7] Zhao, L., Ciallella, H.L., Aleksunes, L.M. and Zhu, H., 2020. Advancing computer-aided drug discovery (CADD) by big data and data-driven machine learning modeling. Drug discovery today, 25(9), pp.1624-1638.

Doi: https://doi.org/10.54216/FPA.190104
Received: October 25, 2024 Revised: January 03, 2025 Accepted: January 30, 2025

[8] D'Souza, S., Prema, K.V. and Balaji, S., 2020. Machine learning models for drug–target interactions: current knowledge and future directions. Drug Discovery Today, 25(4), pp.748-756.

[9] Alqurbi, M.M.A. and Atiah, M.A.Q., 2020. The role of clinical pharmacists in reducing adverse drug reactions. International Journal of Medicine in Developing Countries, 4(1), pp.236-239.

[10] Edmerdash, M., 2021. An Overview of Cloud-Based Secure Services for Enterprise Drug–Drug Interaction Systems. Full Length Article, 2(2), pp.49-9.

[11] Spandana, S. and Prakash, R.V., 2024. Multiple features-based adverse drug reaction detection from social media using deep convolutional neural networks (DCNN). Multimedia Tools and Applications, pp.1-15.

[12] Jiang, M., Liu, G., Su, Y., Jin, W. and Zhao, B., 2025. Hierarchical multi-relational graph representation learning for large-scale prediction of drug-drug interactions. IEEE Transactions on Big Data.

[13] Li, Y., Zhao, W., Dang, B., Yan, X., Gao, M., Wang, W. and Xiao, M., 2024, June. Research on adverse drug reaction prediction model combining knowledge graph embedding and deep learning. In 2024 4th International Conference on Machine Learning and Intelligent Systems Engineering (MLISE) (pp. 322-329). IEEE.

[14] Dey, A., Shrivastava, J.N. and Kumar, C., 2024. Classical-quantum hybrid transfer learning for adverse drug reaction detection from social media posts. Journal of Computational Social Science, pp.1-18.

[15] Gao, Y., Zhang, X., Sun, Z., Chandak, P., Bu, J. and Wang, H., 2025. Precision Adverse Drug Reactions Prediction with Heterogeneous Graph Neural Network. Advanced Science, 12(4), p.2404671.

[16] Lee, C.C., Lee, S., Song, M.H., Kim, J.Y. and Lee, S., 2024. Bidirectional Long Short-Term Memory–Based Detection of Adverse Drug Reaction Posts Using Korean Social Networking Services Data: Deep Learning Approaches. JMIR Medical Informatics, 12(1), p.e45289.

[17] Ahmed, R., Darwish, A. and Hassanien, A.E., 2024. Efficient Prediction Adverse Drug-Drug Interactions with Deep Neural Networks. In Artificial Intelligence for Environmental Sustainability and Green Initiatives (pp. 183-199). Cham: Springer Nature Switzerland.

[18] Lee, D.K., Choi, J.S., Choi, S.J., Choi, M.H. and Hong, M., 2025. Classification of Chronic Obstructive Pulmonary Disease (COPD) Through Respiratory Pattern Analysis. Diagnostics, 15(3), p.313.

[19] Kareem, A.R. and Abdullah, H.S., 2025. Business Intelligence Approach-Based Hybrid Deep Learning Model for QS World University Ranking. Iraqi Journal of Science, pp.375-387.

[20] Wang, M., Yuan, P., Hu, P., Yang, Z., Ke, S., Huang, L. and Zhang, P., 2025. Multi-Strategy Improved Red-Tailed Hawk Algorithm for Real-Environment Unmanned Aerial Vehicle Path Planning. Biomimetics, 10(1), p.31.

[21] https://www.kaggle.com/datasets/arashnic/adhd-diagnosis-data

[22] Wei, J., Lu, Z., Qiu, K., Li, P. and Sun, H., 2020. Predicting drug risk level from adverse drug reactions using SMOTE and machine learning approaches. IEEE Access, 8, pp.185761-185775.

[23] Rawat, A., Wani, M.A., ElAffendi, M., Imran, A.S., Kastrati, Z. and Daudpota, S.M., 2022. Drug adverse event detection using text-based convolutional neural networks (TextCNN) technique. Electronics, 11(20), p.3336.

[24] Yang, J., Hu, Z., Zhang, L. and Peng, B., 2024. Predicting Drugs Suspected of Causing Adverse Drug Reactions Using Graph Features and Attention Mechanisms. Pharmaceuticals, 17(7), p.822.

Doi: https://doi.org/10.54216/FPA.190104
Received: October 25, 2024 Revised: January 03, 2025 Accepted: January 30, 2025