Journal of Intelligent Systems and Internet of Things
Vol. 16, No. 01, PP. 75-85, 2025
DOI: https://doi.org/10.54216/JISIoT.160107

Leveraging Variational Autoencoder with Hippopotamus Optimizer-Based Dimensionality Reduction Model for Attention Deficit Hyperactivity Disorder Diagnosis Data

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 Reactions (ADRs) are very hazardous to patients. Thus, the detection of ADR intends to automatically distinguish, which is an intensive study for public health monitoring functions. Detecting ADRs is the most significant information to determine the patient's opinion on some drugs. As patients can experience projected and occasionally unpredicted negative results from taking some drugs, late detection of ADRs may place life-threatening dangers to patients; posing significant financial, social, and legal consequences to the regulatory agencies and manufacturing companies. The usage of medical data, like states and electronic health records (EHR), became normal in offering a richer understanding of health services and assisting ADR analysis. Developments in deep learning (DL) and machine learning (ML) have made several analytic models have the potential to apply higher-dimensional data to predict adverse effects. In this study, we present a Hippopotamus Optimizer-Based Feature Selection for Adverse Drug Reaction Detection Using a Variational Autoencoder (HOFS-ADRDVAE) model. The main intention of the HOFS-ADRDVAE model is to provide an automatic system for the detection of ADR using state-of-the-art techniques. Initially, the data normalization stage employs min-max normalization for converting input data into a beneficial format. In addition, the feature selection process has been executed by the hippopotamus optimization (HO) algorithm. Besides, the proposed HOFS-ADRDVAE model designs a variational autoencoder (VAE) technique for the classification procedure. At last, the Hunger Games search (HGS) algorithm-based hyperparameter selection process is executed to optimize the classification results of the VAE system. A wide-ranging experiment was implemented to point out the performance of the HOFS-ADRDVAE method. The experimental outcomes specified that the HOFS-ADRDVAE model emphasized improvement over another existing method.

Received: October 22, 2024 Revised: January 05, 2025 Accepted: February 07, 2025

Keywords: Hippopotamus Optimizer; Feature Selection; Adverse Drug Reaction Detection; Variational Autoencoder; Hyperparameter Tuning

1. Introduction

Adverse drug reaction (ADR) is described as the side effect of medications on medical care. A systematic survey of upcoming observational examinations established that 5.3 percent of patients are handling ADRs. Consequently, timely recognition of these actions possibly has a remarkable effect on human health [1]. Societies involve an effectual method to identify ADR connected to scientific medications. Economically, ADRs remarkably raise the costs of hospitalization. Timely recognition of ADRs related to drugs in their post-approval cycles are vital challenge for *pharmacovigilance* models [2]. Pharmacovigilance is described as the activities and science for the valuation, understanding, detection, and avoidance of adverse impacts or some other drug issue. Owing to the several restrictions of pre-approval medical experiments, it is impossible to evaluate the significance of the

utilization of a specific drug before it is liberated [3]. Investigation has exposed that adverse reactions caused by drugs succeeding their discharge into the market is a main healthcare concern: with hospitalizations and deaths numbering in millions. Therefore, post-marketing observation of drugs is of paramount significance for drug manufacturers [4].

At present, owing to the exponentially rising bio-medical literature and the faster growth of media platforms, the source that is created is limitless [5]. Because of great potential and fascinating features, automatically removing entities and their connections from the bio-medical text has captivated much study attention. Many investigations are employed for ADR monitoring like electronic health records and voluntary reporting methods [6]. The fast development of electronic available health-related data and the capability of processing huge amounts of them automatically, utilizing machine learning (ML) models and natural language processing (NLP) are unlocked novel opportunities for pharmacovigilance [7]. Employing sophisticated NLP sources and models, deep linguistic and semantic attributes can be removed from this text [8]. These aspects can be utilized to designate mutually exclusive assets of posts of multiple sizes. The multiple assets can be integrated to create a large number of attributes for every post, consequently, assisting the process of automated classification, and therefore, automated recognition of ADRs [9]. Furthermore, Deep Learning (DL) involves substantial attention in NLP owing to its various benefits like better performances, less feature engineering, and strong representations of data related to other methods [10].

In this study, we present a Hippopotamus Optimizer-Based Feature Selection for Adverse Drug Reaction Detection Using a Variational Autoencoder (HOFS-ADRDVAE) model. Initially, the data normalization stage employs min-max normalization for converting an input data into a beneficial format. In addition, the feature selection procedure has been executed by the hippopotamus optimization (HO) algorithm. Besides, the proposed HOFS-ADRDVAE model designs a variational autoencoder (VAE) technique for the classification process. At last, the Hunger Games search (HGS) algorithm-based hyperparameter selection process is achieved to optimize the classification results of the VAE system. A wide-ranging experiment was implemented to point out the performance of the HOFS-ADRDVAE method.

2. Literature of Works

Das and Mazumder [11] projected a MLCNNF with an image augmentation model to handle this challenge. The major purpose is to establish whether opposing COVID-19 drug reactions might be forecast from the chemical 2D framework. The RGB color channel is utilized to signify the aspects of chemical formulation, and these aspects are removed utilizing MaxPooling2D and Convolution2D layers. In [12], a deep and balanced active learning structure is presented for MedNER to improve the above concerns. Particularly, to define the selection approach accurately, this model initially describes the vagueness of clinical sentences as a labeling loss forecast by a loss-prediction segment and describes variations in the smallest textual distance among pairing of sentences. Jung and Yoo [13] utilized bio-medical studies to project hierarchical attention-based DL techniques to forecast DDIs and their categories. The drug-embedding segment removes representation vectors that effectually take properties of the drug utilizing sequence and sentence embedding approaches. Sentence embedding is a pre-trained bio-medical language method utilized for mapping sentences of drug-related into vector space. Then, it is sequentially fed into bi-LSTM with a hierarchical attention networks. Eventually, a DDI forecast is implemented utilizing a DNN depending on the sequence-embedding vector.

Geng et al. [14] developed MGDDI, a GNN-based method for forecasting possible adverse drug interactions. Particularly, this method utilizes an MGNN for learning drug molecule representations, preventing gradient concerns, and addressing sub-structure size variations. To capture interaction details among drug pairs, a sub-structure interaction-learning segment depending on attention mechanisms is also incorporated. Osheba et al. [15] introduced a novel smart pharmacy method that employs sophisticated LLM to improve pharmacy operational efficiency and drug safety. This method incorporates real-world data from medication databases, patient prescriptions, and electronic health records (EHR) to automate the recognition of possible drug interactions, decreasing ADR-related risks and enhancing medical decision-making. The system's framework is intended to simply handle prescription processing for patients, Pharmacist consultations, inventory control, and management over an intellectual chatbot interface.

Xie et al. [16] developed an innovative method that deviates from the traditional dependence on physically annotated data. It utilizes a transformer-based model named PULearning that integrates adaptive learning. For enhancing the prediction, this model utilizes relative location embedding. Furthermore, dual Kullback-Leibler (KL) regularizers is intended to align with PULearning presumptions. Sun et al. [17] employed multiple-source heterogeneous system data to develop a system model across several network diffusion models and attain compressed lower-dimensional feature vector to predict DTIs. This method is implemented in the MHRW model to enhance the implementation of the RWR model. Moreover, the propagation efficiency of MHRW was utilized in the IMRWR model, assisting deep system sampling.

3. Proposed Methodology

In this paper, we develops a HOFS-ADRDVAE system. The main intention of the HOFS-ADRDVAE model is to provide an automatic method for the recognition of ADR using state-of-the-art techniques. To accomplish that, the proposed HOFS-ADRDVAE technique involves various stages such as data normalization, feature selection, classification, and hyperparameter tuning. Fig. 1 signifies the complete workflow of HOFS-ADRDVAE algorithm.

A. Data Normalization

Initially, the data normalization stage employs min-max normalization for converting input data into a beneficial format. The data normalization utilizing Eq. (1) is of major importance for model training. Normalization not only quickens the model's convergence rate [18]. By scaling feature values to the united range, optimizer methods like gradient descent can iterate most effectively. It additionally avoids the effect of feature-dominance, evading the condition however, features with larger numeric values need a very important influence on the method, therefore improving the contributions balance from dissimilar features. Moreover, normalization participates in enhancing mathematical stability decreasing problems like gradient explosion or numerical overflow, and improving the model is training stability such as neural networks.

$$y = (y_{\max} - y_{\min}) \times \frac{x - x_{\min}}{x_{\max} - x_{\min}} + y_{\min} \tag{1}$$

The normalized value is represented as x, while $x_{\max}$ and $x_{\min}$ represents maximum and minimum values, individually, and the interval of normalization recognized at $[y_{\max}, y_{\min}]$ is selected as $[0,1]$.

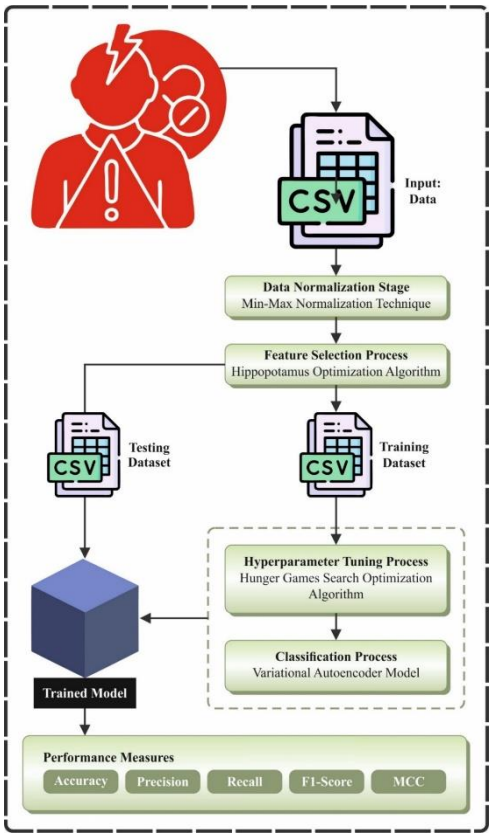

Figure 1. Overall Workflow of HOFS-ADRDVAE algorithm

B. Feature Selection Method

In addition, the HO algorithm has executed the process of feature selection. The Hippo Optimizer (HO) is a new metaheuristic optimization algorithm stimulated by the hippo's natural behaviour on land and in water [18]. The model attains this aim by mimicking the dual-motion tactic of hippos, incorporating local exploitation and global exploration to deal with composite optimizer issues. The fundamental to the model rests in the interaction between these dual stages: swimming in the pool water that signifies the global search, and walking on land that represents the local search.

During this swimming stage, the hippopotamus is designed to freely move over water, discovering the solution area widely. This movement is managed by the equation, which combines the present location of the solution, the distance to the global optimal solution, and dual random features to present diversities within the search procedure. The equation to update the location of the solution in this stage is as shown:

$$X_{new} = X_{old} + \beta \cdot (rand_1 \cdot D - rand_2 \cdot D) \quad (2)$$

Whereas X_{new} denote novel candidate solutions, X_{old} signifies the present solution, D refers to the distance between the present solution and the global optimal solution, and β denotes the coefficient, which controls the speed of the swimming. $rand_1$ and $rand_2$ are random variables present randomness to the movement, guaranteeing that the searching procedure protects an extensive region of the solution area. During this walking stage, the hippopotamus is challenged to move on land, concentrating on local search to improve the solution. This stage tries to optimize the position of solution near the present global optimal, and the position upgrade is directed by the succeeding equation:

$$X_{new} = X_{old} + \alpha \cdot \left(rand_3 \cdot (X_{best} - X_{old})\right) \quad (3)$$

Now, X_{new} characterizes the upgraded solution, X_{best} denotes the global optimal solution discovered thus far which are the parameters, that control the walking's step size. The random variable $rand_3$ has been applied to present variation in the movement, therefore avoiding the search from being deterministic and guaranteeing that the model fails to become trapped in local bests. The method operates by iterating over these dual stages: global exploration over swimming and local refinement over walking. The solutions are estimated utilizing an objective function, and the global optimal solution is upgraded appropriately. The process continues till a halt condition is encountered, like reaching the maximal iteration counts or attaining a particular convergence level. The key parameters of the HOA contain β that controls the speed of the swimming and thus the exploration capability of the model, and α that fine-tunes the step size in the local walking stage, manipulating the search accuracy. $rand_1, rand_2,$ and $rand_3$ are random variables applied to increase the diversity of the search procedure and stop early convergence with lower answers. By integrating the benefits of either local exploitation or global exploration, the HOA establishes important promise in resolving composite optimizer tasks, giving an adjustable and strong model for dealing with real-time difficulties through several fields.

In the HO model, the fitness function (FF) applied is composed to have a balance among the chosen feature counts in all solutions (least) and the classification precision (greatest) gained by employing these chosen features, Eq. (4) characterizes the FF to estimate solution.

$$Fitness = \alpha\gamma_R(D) + \beta\frac{|R|}{|C|} \quad (4)$$

Whereas $\gamma_R(D)$ characterizes the classification rate of error of a specified classifier. $|R|$ represents cardinality of the chosen sub-set and $|C|$ denotes total feature counts in the dataset, α and β denote dual parameters. $\in$ [1,0] and $\beta = 1 - \alpha$.

C. Classification Model

Besides, the proposed HOFS-ADRDVAE model designs the VAE technique for the classification process. AEs are artificial neural networks (ANNs), which are proficient to rebuild their input in a self-directed way [19]. AE is made up of encoder-decoder networks. The encoding gets us to input the sample of data $x \in \mathbb{R}^d$, while d represents data dimensions, and squeezes these data into the latent area $z \in \mathbb{R}^h$, whereas h denotes the encoder dimension, typically $h < d$. Formerly, the decoder attempts for mapping support the latent representations z to the unique input area $\hat{x} \in \mathbb{R}^d$ over reconstruction. The encoding architecture is like a bottleneck, wherein data are condensed into removing an important encoding representation. The decoding performs the reverse. The dual systems are illustrated by the function of the encoder f_φ: $\mathbb{R}^d \to \mathbb{R}^h$ and by the function of the decoder f_θ:$\mathbb{R}^h \to \mathbb{R}^d$, correspondingly. Discover parameters (weights) φ and θ for the dual functions is completed by backpropagation (BP), reducing the loss function. $\mathcal{L}_{AE}(x, \hat{x}) = \|x - \hat{x}\|^2$, named error of reconstruction, specified output $\hat{x}$ and input x. Fig. 2 represents the architecture of VAE model.

VAEs contrasted with basic AEs, assume that the presence of a probability-based method parametrized through the variable of latent $z \in \mathbb{R}^h$, which makes the noted input values $x \in \mathbb{R}^d$. In that z latent (for example: hidden), they might understand its features by calculating the following $p(z|x) = \frac{p(x|z)p(z)}{p(x)}$; miserably, calculating the marginalization $p(x) = \int p\,(x|z)p(z)dz$ toward the denominator is inflexible after z is higher-dimensional. Variation interpretation is applied to overwhelm this problem by estimating $p(z|x)$ together with distribution $q(z|x)$ that is susceptible, and by reducing their divergence of KL $D_{KL}(q(z|x)\|p(z|x))$ to guarantee that $q(z|x)$

is related to $p(z|x)$. By substituting $p(z|x)$ in $D_{KL}(q(z|x)\|p(z|x))$ with $\frac{p(x,z)}{p(x)}$, the reduction problem becomes equal to the maximization of $E_{q(z|x)}[\log p(x|z)] - D_{KL}(q(z|x)\|p(z))$, while the initial term characterizes the probability of the reconstruction (similar to the error of reconstruction in AEs), however, the next term guarantees that the distribution of learned $q(z|x)$ is connected with the true previous distribution $p(z)$ (be the head of regularizing the variable of latent z). The distributions $q(z|x)$ and $p(z|x)$ are parametrized utilizing dual ANNs that are equivalent to a decoder (by weights θ)and encoder (by weights φ), consistently. Discover weights φ and θ is completed by BP, reducing the loss function $\mathcal{L}_{VAE}(x) = D_{KL}(q_\varphi(z|x)\|p_\theta(z)) - E_{q_{(\beta}(z|x)}(\log p_\theta(x|z))$, that characterizes the variation lower limit of the data x based on the inequality of the Jensen. The posterior $q_\varphi(z|x)$ is typically considered as generally dispersed by parameters (μ_z, Σ_z), however the popular option for the previous $p_\theta(z)$ becomes the normal distribution of isotropic $\mathcal{N}(0, I)$.

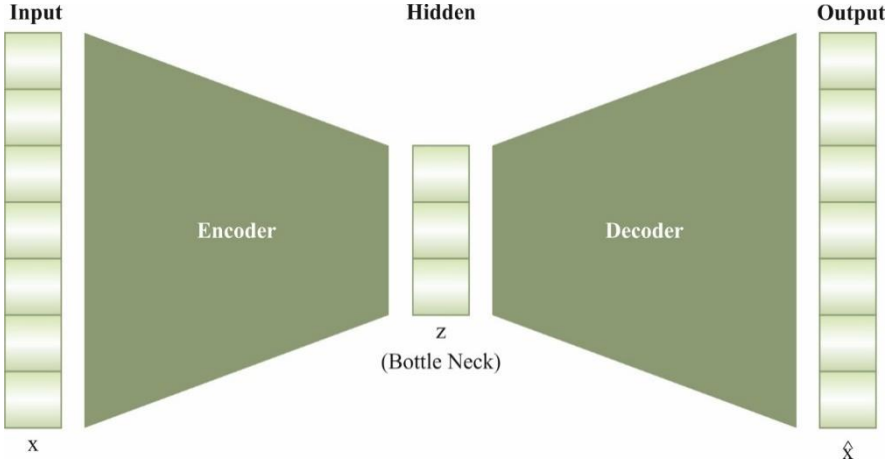

Figure 2. Architecture of VAE model

D. Hyperparameter Tuning Method

At last, the HGS algorithm-based hyperparameter tuning process is achieved to optimize the classification results of the VAE system. The HGS is a new population-based metaheuristic model, that stimulated the cooperative behaviour of the animals, and their movements are motivated by hunger [20]. The mathematic modelling of HGS comprises dual phases the hunger rule and the approach to food. The animals can work together socially throughout the foraging, and Eq. (5) characterizes foraging behaviour and cooperative communication

$$X_{t+1}^i = \begin{cases} Game_1: X_t^i * (1 + rand), & r_1 < l \\ Game_2: W_1 * X_b + R * W_2 * |X_b - X_t^i|, & r_1 > l, r_2 > E \\ Game_3: W_1 * X_b - R * W_2 * |X_b - X_t^i|, & r_1 > l, r_2 < E \end{cases} \quad (5)$$

Whereas r_1 and r_2 represents randomly generated numbers related to[0,1]. $rand$ stands for regular distribution randomly generated numbers, signifies the iteration counts, l means constant, W_1 and W_2 represents hunger weights, X_b indicates the solution position using the optimal fitness and X_t^i characterizes the agent position (i) in the iteration (t).

R signifies a ranging controller that slowly reaches 0 and is stated as Eq. (6) whereas E epitomizes the variation controller of all locations and is approximated as in Eq. (7):

$$R = (2^* rand - 1)^* 2^* \left(1 - \frac{t}{T}\right) \quad (6)$$

$$E = \frac{2}{e^{|F(i) - BF|} + e^{-|F(i) - BF|}} \quad (7)$$

Whereas $F(i)$ denotes fitness of the agent index (i) and BF refers to optimal fitness amongst fitness of all agents. $W_1(i)$ and $W_2(i)$ are expressed as in Eqs. (8) and (9).

$$W_1(i) = \begin{cases} hungry, (i) * \frac{N}{SHungry} * r_4 \text{ '} r_3 < l \\ 1, \qquad r_3 > l \end{cases} \quad (8)$$

$$W_2(i) = (1 - e^{-|Hungry(i) - SHungry|})^* 2^* r_5 \quad (9)$$

Here, N characterizes the agent counts and $SHungry$ characterizes the collective feeling of hunger powered by each individual, described as the addition of their hungry conditions (sum($hungry$)). r_3, r_4 and r_5 denote randomly generated numbers within the interval [0,1].

$Hungry\ (i)$ characterizes the hunger of all agents and is characterized as in Eq. (10):

$$Hungry\ (i) = \begin{Bmatrix} 0, & F(i) == BF \\ hungry(i) + H, & F(i)! = BF \end{Bmatrix} \quad (10)$$

The equation of H is projected according to Eqs. (11) and (12):

$$TH = \frac{F(i)-BF}{WF-BF} * r_6 * 2 * (Ub - Lb) \quad (11)$$

$$H = \begin{Bmatrix} LH * (1 + r), & TH < LH \\ TH, & TH > LH \end{Bmatrix} \quad (12)$$

Whereas r_6 characterizes randomly generated numbers inside the interval [0,1], WF denotes poor fitness within the iteration, LB and UB represent upper and lower limits. LH denotes continuous parameter adjusted in experimental testing. Model 1 defines the process of the HGS model. The major advantages of the HGS model are established in dual ideas: initially, it contains the effective exploration of the search area because of (W_1 and (W_2) and prevents catching in local bests. Nevertheless, the major disadvantage of HGS is the weak exploitation of the searching space.

Algorithm1: Pseudocode of HGS Algorithm
Set the parameters: l, LH, N, T, and $SHungry$
Initialization of the agent's locations $X_i, i = 1{:}N$
While (t<=T)
Calculate the all agent's Fitness $(F(i), i = 1{:}N)$
Upgrade X_b, WF and BF
Calculate the Hungry utilizing Eq. (10).
Calculate W_1 and W_2 for all agents according to Eqs. (8) and (9), correspondingly.
Calculate R utilizing Eq. (6).
Calculate E utilizing Eq. (7)
Upgrade the location of the agent according to Eq. (5)
$t = t + 1$
End While
Return X_b and BF

The HGS model originates a fitness function (FF) to influence boosted performance of classification. It outlines an optimistic number to embody the better outcome of the candidate solution. In this paper, the classification rate of error reduction was reflected as FF. Its mathematical formulation is represented in Eq. (13).

$$fitness(x_i) = ClassifierErrorRate(x_i)$$
$$= \frac{no.of\ misclassified\ samples}{Total\ no.of\ samples} * 100 \quad (13)$$

4. Experimental Analysis

The performance evaluation of the HOFS-ADRDVAE model is inspected under ADHD Diagnosis data [21]. This dataset contains 103 instances under dual class labels as portrayed in Table 1. The total amount of features is 788 but only 356 features are chosen.

Table 1: Dataset details

Class Labels	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 formed by the HOFS-ADRDVAE model below 80:20 and 70:30 of TRAPH/TSPH. The outcomes identify that the HOFS-ADRDVAE approach is capable of the detection and identification of each class.

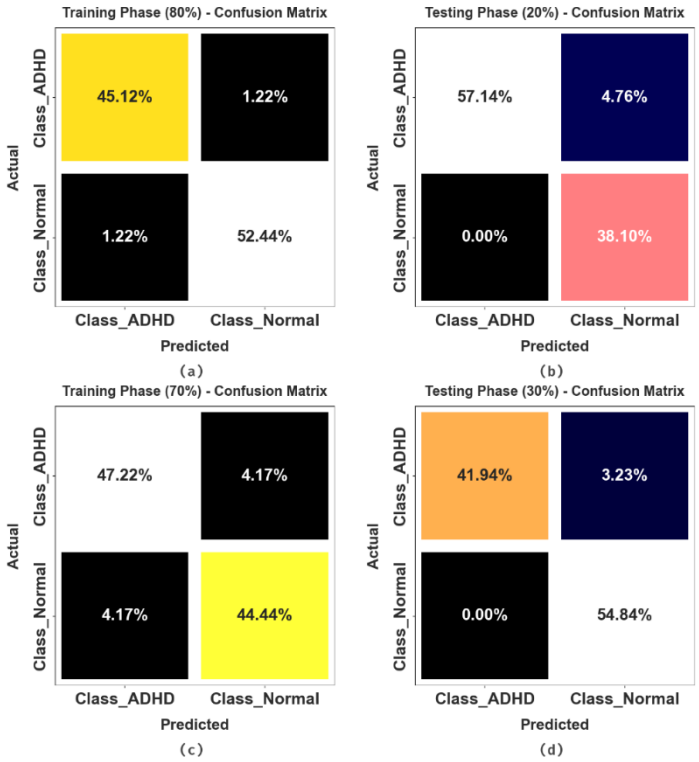

Figure 3. Confusion matrix of HOFS-ADRDVAE methodology (a-b) 80%TRAPH and 20%TESPH and (c-d) 70%TRAPH and 30%TESPH

Table 2: ADR detection of HOFS-ADRDVAE model under 80:20 and 70:30 of TRAPH/TESPH

Class Labels	$Accu_y$	$Prec_n$	$Reca_l$	$F1_{score}$	MCC
TRAPH (80%)					
Class_ADHD	97.37	97.37	97.37	97.37	95.10
Class_Normal	97.73	97.73	97.73	97.73	95.10
Average	97.55	97.55	97.55	97.55	95.10
TESPH (20%)					
Class_ADHD	92.31	100.00	92.31	96.00	90.58
Class_Normal	100.00	88.89	100.00	94.12	90.58
Average	96.15	94.44	96.15	95.06	90.58
TRAPH (70%)					
Class_ADHD	91.89	91.89	91.89	91.89	83.32
Class_Normal	91.43	91.43	91.43	91.43	83.32
Average	91.66	91.66	91.66	91.66	83.32
TESPH (30%)					
Class_ADHD	92.86	100.00	92.86	96.30	93.65
Class_Normal	100.00	94.44	100.00	97.14	93.65
Average	96.43	97.22	96.43	96.72	93.65

Table 2 and Fig. 4 represent the ADR detection of the HOFS-ADRDVAE algorithm below 80:20 and 70:30 of TRAPH/TESPH. The result stated that the HOFS-ADRDVAE approach has accurately classified all the different classes. Based on 80% TRAPH, the proposed HOFS-ADRDVAE approach accomplishes average $accu_y$ of 97.55%, $prec_n$ of 97.55%, $reca_l$ of 97.55%, $F1_{score}$ of 97.55%, and MCC of 95.10%. Also, depending on 20% TESPH, the proposed HOFS-ADRDVAE system achieves average $accu_y$ of 96.15%, $prec_n$ of 94.44%, $reca_l$ of 96.15%, $F1_{score}$ of 95.06%, and MCC of 90.58%. Additionally, for 70% TRAPH, the proposed HOFS-ADRDVAE methodology attains average $accu_y$ of 91.66%, $prec_n$ of 91.66%, $reca_l$ of 91.66%, $F1_{score}$ of 91.66%, and MCC of 83.32%. Likewise, based on 20% TESPH, the proposed HOFS-ADRDVAE system attains average $accu_y$ of 96.43%, $prec_n$ of 97.22%, $reca_l$ of 96.43%, $F1_{score}$ of 96.72%, and MCC of 93.65%.

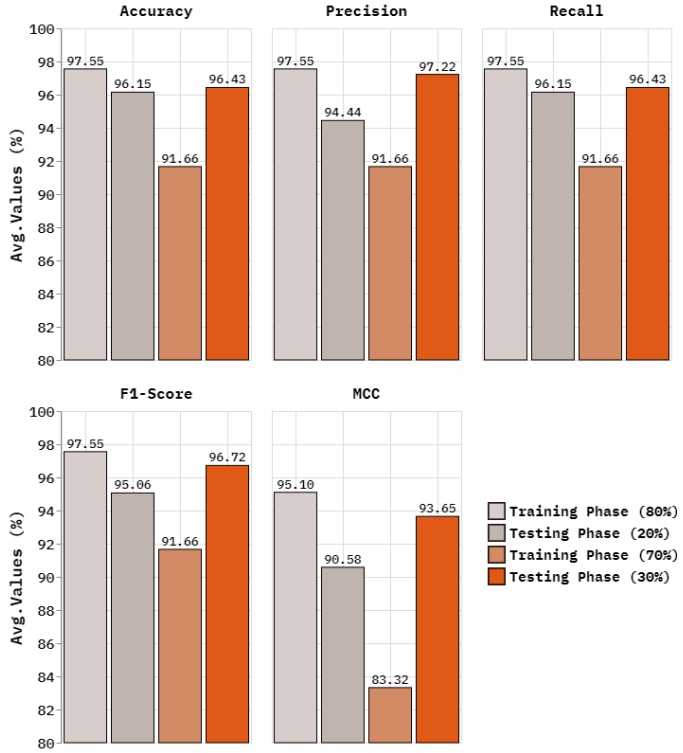

Figure 4. Average of HOFS-ADRDVAE model under 80:20 and 70:30 of TRAPH/TESPH

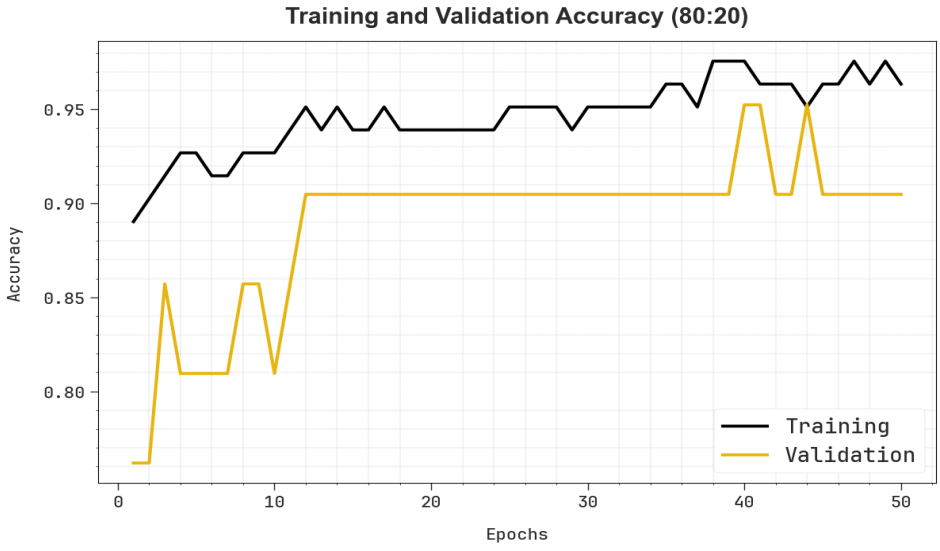

Figure 5. $Accu_y$ analysis of HOFS-ADRDVAE model below 80%TRAPH and 20%TESPH

In Fig. 5, the training (TRA) $accu_y$ and validation (VAL) $accu_y$ outcomes of the HOFS-ADRDVAE system below 80%TRAPH and 20%TESPH is illustrated. The $accu_y$ analysis are calculated across a period time of 0-50 epochs. The figure highlights that the TRA and VAL $accu_y$ analysis exhibitions a rising trend which well-versed the capacity of the HOFS-ADRDVAE algorithm with superior outcomes across multiple iterations. Simultaneously, the TRA and VAL $accu_y$ remains closer across the epochs, which indicates inferior overfitting and exhibits maximum outcomes of the HOFS-ADRDVAE approach, guaranteeing reliable prediction on unseen samples.

In Fig. 6, the TRA loss (TRALOS) and VAL loss (VALLOS) curve of the A HOFS-ADRDVAE methodology below 80%TRAPH and 20%TESPH are shown. The values of loss are calculated within the range of 0-50 epochs. It is indicated that the TRALOS and VALLOS analysis demonstrate a reducing trend, informing the capability of the HOFS-ADRDVAE technique to balance a trade-off. The continuous reduction in values of loss besides promises the optimal performance of the HOFS-ADRDVAE technique.

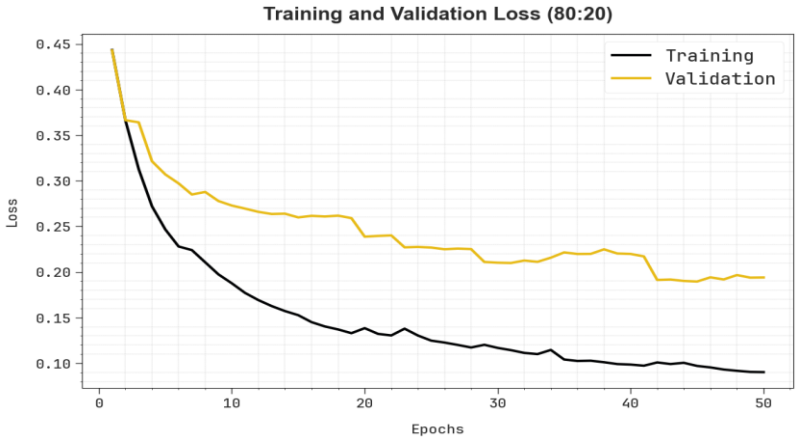

Figure 6. Loss graph of HOFS-ADRDVAE model below 80%TRAPH and 20%TESPH

The comparative results of the HOFS-ADRDVAE algorithm with existing methodologies are demonstrated in Table 3 and Fig. 7 [22-24]. The simulation outcome specified that the HOFS-ADRDVAE approach outperformed better performances. Depend on $accu_y$, the HOFS-ADRDVAE technique has higher $accu_y$ of 97.55% while the AdaBoost, GB, RF, ME-TFIFDF, CRNN, FPGNN-SDAJM, and ResNet approaches have obtained lesser $accu_y$ of 85.17%, 85.16%, 95.16%, 90.01%, 95.23%, 92.92%, and 92.42%, respectively. Moreover, based on $Prec_n$, the HOFS-ADRDVAE approach has maximum $Prec_n$ of 97.55% where the AdaBoost, GB, RF, ME-TFIFDF, CRNN, FPGNN-SDAJM, and ResNet methods have accomplished lesser $Prec_n$ of 91.28%, 85.68%, 85.66%, 95.37%, 92.96%, 88.32%, 91.85%, correspondingly. In addition, depending on $Reca_l$, the HOFS-ADRDVAE methodology has greater $Reca_l$ of 97.55% whereas the AdaBoost, GB, RF, ME-TFIFDF, CRNN, FPGNN-SDAJM, and ResNet techniques have reached the worst $Reca_l$ of 85.86%, 87.87%, 86.64%, 93.77%, 93.61%, 89.45%, 94.38%, respectively. Lastly, for $Reca_l$, the HOFS-ADRDVAE model has better $Reca_l$ of 97.55% while the AdaBoost, GB, RF, ME-TFIFDF, CRNN, FPGNN-SDAJM, and ResNet algorithms have attained minimal $Reca_l$ of 95.13%, 90.91%, 93.03%, 95.30%, 92.48%, 89.06%, 86.63%, correspondingly.

Table 3: Comparative outcomes of HOFS-ADRDVAE algorithm with existing technique

Algorithm	$Accu_y$	$Prec_n$	$Reca_l$	$F1_{score}$
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
HOFS-ADRDVAE	97.55	97.55	97.55	97.55

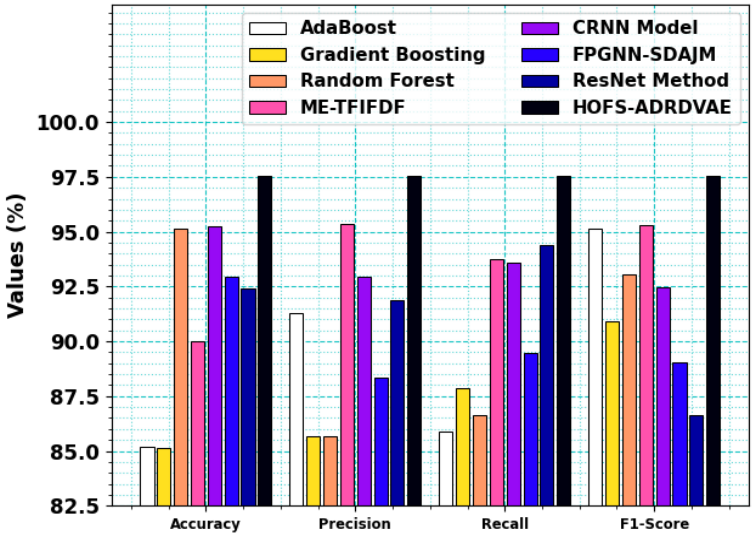

Figure 7. Comparative outcomes of HOFS-ADRDVAE technique with existing methodologies

5. Conclusion

In this paper, we present a HOFS-ADRDVAE model. The main intention of the HOFS-ADRDVAE model is to provide an automatic method for the recognition of ADR sing state-of-the-art techniques. Initially, the data normalization stage employs min-max normalization for converting input data into a beneficial format. In addition, the HO algorithm has executed the feature selection process. Besides, the proposed HOFS-ADRDVAE model designs the VAE technique for the classification process. At last, the HGS algorithm-based hyperparameter tuning process is performed to optimize the classification results of the VAE system. A wide-ranging experiment was implemented to point out the performance of the HOFS-ADRDVAE method. The experimental outcomes specified that the HOFS-ADRDVAE model emphasized improvement over another existing method.

Funding: "This research received no external funding"

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

References

[1] C. S. Wang, P. J. Lin, C. L. Cheng, S. H. Tai, Y. H. Kao Yang, and J. H. Chiang, "Detecting potential adverse drug reactions using a deep neural network model," *J. Med. Internet Res.*, vol. 21, no. 2, p. e11016, 2019.

[2] A. A. Pandit and S. A. Dubey, "A comprehensive review of Adverse Drug Reactions (ADRs) Detection and Prediction Models," in *Proc. 13th Int. Conf. Comput. Intell. Commun. Netw. (CICN)*, Sep. 2021, pp. 123–127.

[3] Z. Rezaei, H. Ebrahimpour-Komleh, B. Eslami, R. Chavoshinejad, and M. Totonchi, "Adverse drug reaction detection in social media by deep learning methods," *Cell J. (Yakhteh)*, vol. 22, no. 3, p. 319, 2019.

[4] B. Fan, W. Fan, and C. Smith, "Adverse drug event detection and extraction from open data: A deep learning approach," *Inf. Process. Manag.*, vol. 57, no. 1, p. 102131, 2020.

[5] H. Khalil and C. Huang, "Adverse drug reactions in primary care: a scoping review," *BMC Health Serv. Res.*, vol. 20, pp. 1–13, 2020.

[6] O. Choudhury et al., "Predicting adverse drug reactions on distributed health data using federated learning," in *AMIA Annu. Symp. Proc.*, vol. 2019, p. 313, Mar. 2020.

[7] R. Prasad, A. Singh, and N. Gupta, "Adverse drug reactions in tuberculosis and management," *Indian J. Tuberc.*, vol. 66, no. 4, pp. 520–532, 2019.

[8] F. Schifano, S. Chiappini, J. M. Corkery, and A. Guirguis, "An insight into Z-drug abuse and dependence: an examination of reports to the European Medicines Agency database of suspected adverse drug reactions," *Int. J. Neuropsychopharmacol.*, vol. 22, no. 4, pp. 270–277, 2019.

[9] J. Wang et al., "Landscape of DILI-related adverse drug reaction in China Mainland," *Acta Pharm. Sin. B*, vol. 12, no. 12, pp. 4424–4431, 2022.

[10] B. R. Del Pozzo-Magaña and C. Liy-Wong, "Drugs and the skin: A concise review of cutaneous adverse drug reactions," *Br. J. Clin. Pharmacol.*, vol. 90, no. 8, pp. 1838–1855, 2024.

[11] P. Das and D. H. Mazumder, "MLCNNF: A Multi-Label Convolutional Neural Network Framework for Predicting Adverse COVID Drug Reactions From the Chemical Structure," *IEEE Trans. Comput. Biol. Bioinf.*, 2025.

[12] Y. Zhuang, J. Zhang, R. Lu, K. He, and X. Li, "MedNER: Enhanced Named Entity Recognition in Medical Corpus via Optimized Balanced and Deep Active Learning," *ACM Trans. Intell. Syst. Technol.*, vol. 15, no. 5, pp. 1–24, 2024.

[13] S. Jung and S. Yoo, "Interpretable prediction of drug-drug interactions via text embedding in biomedical literature," *Comput. Biol. Med.*, vol. 185, p. 109496, 2025.

[14] G. Geng et al., "MGDDI: A multi-scale graph neural networks for drug–drug interaction prediction," *Methods*, vol. 228, pp. 22–29, 2024.

[15] A. Osheba et al., "Leveraging Large Language Models for Smart Pharmacy Systems: Enhancing Drug Safety and Operational Efficiency," in *Proc. 19th Int. Conf. Ubiquitous Inf. Manag. Commun. (IMCOM)*, Jan. 2025, pp. 1–8.

[16] W. Xie et al., "Transformer-based Named Entity Recognition for Clinical Cancer Drug Toxicity by Positive-unlabeled Learning and KL Regularizers," *Curr. Bioinform.*, vol. 19, no. 8, pp. 738–751, 2024.

[17] L. Sun, Z. Yin, and L. Lu, "ISLRWR: A network diffusion algorithm for drug–target interactions prediction," *PLoS One*, vol. 20, no. 1, p. e0302281, 2025.

[18] Y. Cao, W. Wang, and Y. He, "Prediction of Heat-Treated Wood Adhesive Strength Using BP Neural Networks Optimized by Four Novel Metaheuristic Algorithms," *Forests*, vol. 16, no. 2, p. 291, 2025.

[19] D. Azzalini et al., "An empirical evaluation of deep autoencoders for anomaly detection in the electricity consumption of buildings," *Energy Build.*, vol. 327, p. 115069, 2025.

[20] M. Issa, M. A. Elaziz, and S. I. Selem, "Enhanced hunger games search algorithm that incorporates the marine predator optimization algorithm for optimal extraction of parameters in PEM fuel cells," *Sci. Rep.*, vol. 15, no. 1, p. 4474, 2025.

[21] "ADHD diagnosis data," *Kaggle*, 2021. [Online]. Available: https://www.kaggle.com/datasets/arashnic/adhd-diagnosis-data.

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

[23] A. Rawat et al., "Drug adverse event detection using text-based convolutional neural networks (TextCNN) technique," *Electronics*, vol. 11, no. 20, p. 3336, 2022.

[24] J. Yang, Z. Hu, L. Zhang, and B. Peng, "Predicting Drugs Suspected of Causing Adverse Drug Reactions Using Graph Features and Attention Mechanisms," *Pharmaceuticals*, vol. 17, no. 7, p. 822, 2024.