



Article

Clinical Decision Support Systems to Predict Drug–Drug Interaction Using Multilabel Long Short-Term Memory with an Autoencoder

Fadwa Alrowais ¹, Saud S. Alotaibi ², Anwer Mustafa Hilal ^{3,*}, Radwa Marzouk ⁴, Heba Mohsen ⁵, Azza Elneil Osman ³, Amani A. Alneil ³ and Mohamed I. Eldesouki ⁶

- ¹ Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
² Department of Information Systems, College of Computing and Information System, Umm Al-Qura University, Makkah 24211, Saudi Arabia
³ Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, Al Kharj 16436, Saudi Arabia
⁴ Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
⁵ Department of Computer Science, Faculty of Computers and Information Technology, Future University in Egypt, New Cairo 11835, Egypt
⁶ Department of Information System, College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Al Kharj 16436, Saudi Arabia
* Correspondence: a.hilal@psau.edu.sa



Citation: Alrowais, F.; Alotaibi, S.S.; Hilal, A.M.; Marzouk, R.; Mohsen, H.; Osman, A.E.; Alneil, A.A.; Eldesouki, M.I. Clinical Decision Support Systems to Predict Drug–Drug Interaction Using Multilabel Long Short-Term Memory with an Autoencoder. *Int. J. Environ. Res. Public Health* **2023**, *20*, 2696. <https://doi.org/10.3390/ijerph20032696>

Academic Editors: Fabio Mendonca, Morgado Dias and Sheikh Shanawaz Mostafa

Received: 21 October 2022

Revised: 24 January 2023

Accepted: 26 January 2023

Published: 2 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Big Data analytics is a technique for researching huge and varied datasets and it is designed to uncover hidden patterns, trends, and correlations, and therefore, it can be applied for making superior decisions in healthcare. Drug–drug interactions (DDIs) are a main concern in drug discovery. The main role of precise forecasting of DDIs is to increase safety potential, particularly, in drug research when multiple drugs are co-prescribed. Prevailing conventional method machine learning (ML) approaches mainly depend on handcraft features and lack generalization. Today, deep learning (DL) techniques that automatically study drug features from drug-related networks or molecular graphs have enhanced the capability of computing approaches for forecasting unknown DDIs. Therefore, in this study, we develop a sparrow search optimization with deep learning-based DDI prediction (SSODL-DDIP) technique for healthcare decision making in big data environments. The presented SSODL-DDIP technique identifies the relationship and properties of the drugs from various sources to make predictions. In addition, a multilabel long short-term memory with an autoencoder (MLSTM-AE) model is employed for the DDI prediction process. Moreover, a lexicon-based approach is involved in determining the severity of interactions among the DDIs. To improve the prediction outcomes of the MLSTM-AE model, the SSO algorithm is adopted in this work. To assure better performance of the SSODL-DDIP technique, a wide range of simulations are performed. The experimental results show the promising performance of the SSODL-DDIP technique over recent state-of-the-art algorithms.

Keywords: healthcare; decision making; big data; drug–drug interaction; deep learning; predictive models

1. Introduction

In the digital era, the velocity and volume of public, environmental, health, and population data from a wider variety of sources are rapidly developing. Big Data analytics technologies such as deep learning (DL), statistical analysis, data mining (DM), and machine learning (ML) are used to create state-of-the-art decision models [1]. Decision making based on concrete evidence is crucial and has a dramatic effect on program implementation and public health. This highlights the significant role of a decision model under

uncertainty, involving health intervention, disease control, health services and systems, preventive medicine, quality of life, health disparities and inequalities, etc. A drug–drug interaction (DDI) can occur when more than one drug is co-prescribed [2]. Even though DDIs might have positive impacts, sometimes they have serious negative impacts and result in withdrawing a drug from the market. DDI prediction could assist in reducing the possibility of adverse reactions and improve the post-marketing surveillance and drug development processes [3]. Medical trials are time consuming and impracticable with respect to dealing with largescale datasets and the limitations of experimental conditions. Hence, researchers have presented a computation method to speed up the process of prediction [4]. The present computation DDI prediction method is divided into five classes of models: DL-based, network-based, similarity-based, literature extraction-based, and matrix factorization-based models.

ML techniques are an emerging area which are employed in large datasets for extracting hidden concepts and relationships amongst attributes [5]. An ML model can be used to forecast outcomes. Since it is extremely complex for humans to process and handle a large amount of data [6], hence, an ML model can play a major role to forecast healthcare outcomes with high quality and cost minimization [7]. ML algorithms are based primarily on rule-based, probability-based, tree-based, etc. methods. Large quantities of data gathered from a variety of sources are applied in the data preprocessing stage. During this stage, data dimension is minimized by eliminating redundant data. As the amount of data increases, a model is not capable of making a decision. Hence, various methods must be developed so that hidden knowledge or useful patterns are extracted from previous information [8]. Then, a model using a ML algorithm is tested under test data to discover the model's performance, which can be augmented again by considering some rules or parameters. Generally, ML is utilized in the area of prediction, data classification, and pattern recognition [9]. Numerous applications such as disease prediction, face detection, fraud detection, traffic management, and email filtering, use the ML concept. The DL method is part of ML algorithms, which makes use of supervised and unsupervised models for feature classification [10]. The various elements of DL approaches are utilized in the field of recommender systems, disease prediction, and image segmentation such as restricted Boltzmann machines (RBM), convolution neural networks (CNN), and autoencoders (AEs).

In this study, we develop a sparrow search optimization with deep learning-based DDI prediction (SSODL-DDIP) technique for healthcare decision making in big data environments. The presented SSODL-DDIP technique applies a multilabel long short-term memory with an autoencoder (MLSTM-AE) model for the DDI prediction process. Moreover, a lexicon-based approach is involved in determining the severity of interactions among the DDIs. To improve the prediction outcomes of the MLSTM-AE model, the SSO algorithm is adopted in this work. For ensuring better performance of the SSODL-DDIP technique, a wide range of simulations are performed.

2. Related Works

In [10], the authors proposed a positive unlabeled (PU) learning model which utilized a one-class support vector machine (SVM) model as the learning algorithm. The algorithm could learn the positive distribution from the unified feature vector space of drugs and targets, and regarded unknown pairs as unlabeled rather than labeling them as negative pairs. Wang et al. [11] introduced a novel technique, multi-view graph contrastive representative learning for DDI forecasting, MIRACLE for brevity, for capturing intra-view interactions and inter-view molecular structure among molecules concurrently. MIRACLE treated a DDI network as a multi-view graph in which all nodes in the interaction graph were a drug molecule graph sample. The author employed a bond-aware attentive message propagating algorithm for capturing drug molecular structured data and a graph convolution network (GCN) for encoding DDI relations in the MIRACLE learning phase. Along with that, the author modeled an innovative unsupervised contrastive learning element to integrate and balance multi-view data. In [12], the author devised a deep neural networks

(DNNs) method that precisely identified the protein–ligand interactions with particular drugs. The DNN could sense the response of protein–ligand interactions for the particular drugs and could find which drug could effectively combat the virus.

Lin et al. [13] modeled an end-to-end structure, named a knowledge graph neural network (KGNN), for resolving DDI estimation. This structure could capture a drug and its neighborhood by deriving their linked relations in a knowledge graph (KG). For extracting semantic relations and high-order structures of the KG, the author studied the neighborhoods for all entities in KG as its local receptive, and then compiled neighborhood data from representations of the current entities. Pang et al. [14] presented a new attention-system-related multidimensional feature encoder for DDI estimation, called attention-related multidimensional feature encoders (AMDEs). To be specific, in an AMDE, the author encoded drug features from multidimensional features, which included data from an atomic graph of the drug and a simplified molecular-input line-entry system sequence. Salman et al. [15] modeled a DNN-oriented technique (SEV-DDI: Severity-DDI) that included certain integrated units or layers for attaining higher accuracy and precision. The author moved a step further and used the techniques for examining the seriousness of the interaction, after outpacing other methods in the DDI classifier task successfully. The capability to determine DDI severity helps in clinical decision aid mechanisms for making very precise and informed decisions, assuring the patient’s safety.

Liu et al. [16] presented a deep attention neural network-related DDI predictive structure (DANN-DDI), for forecasting unnoticed DDIs. Firstly, by utilizing the graph embedding technique, the author framed multiple drug feature networks and learned drug representation from such networks; after that, the author concatenated learned drug embeddings and implemented an attention neural network for learning representation of drug–drug pairs; finally, the author devised a DNN to precisely estimate DDIs. Zhang et al. [17] introduced a sparse feature learning ensembled approach with linear neighborhood regularization (SFLN), for forecasting DDIs. Initially, the authors compiled four drug features, i.e., pathways, chemical structures, enzymes, and targets, by mapping drugs in distinct feature spaces into general interaction spaces by sparse feature learning. Then, the authors presented the linear neighborhood regularizations for describing the DDIs in the communication space by utilizing known DDIs.

3. The Proposed Model

In this study, we introduce a novel SSODL-DDIP technique for DDI predictions in big data environments. The presented SSODL-DDIP technique accurately determines the relationship and drug properties from various sources to make predictions. It encompasses data preprocessing, MLSTM-AE-based DDI prediction, SSO hyperparameter tuning, and severity extraction.

3.1. Data Preprocessing

Standard text cleaning and preprocessing operations were carried out on sentences involving but not constrained to lemmatization. Every drug discussed in a sentence was considered and labeled to interact with others [18]. The number of drug pairs (DP) in a sentence is evaluated as follows:

$$\text{Drug Pairs (DP)} = \max \left(0, \sum_{i=1}^n (i - 1) \right) \quad (1)$$

where n indicates the number of drugs in a sentence.

In addition, drug blinding was used, whereby all the drug names were allocated to the label, for a sentence, “Aspirin might reduce the effect of probenecid”, labeled sentence was “Drug^A might reduce the effect of Drug^B”. The drug blinding method assists a technique to identify this label as “subject” and “object” that ultimately assist an approach during classification. Then, the processed sentence is given to the approach for classification and detection of DDI.

During word embedding, every word was converted into a real value vector. This word mapping into the matrix can be performed using Word2Vec and embedding data using the abstract of PubMed comprising the drugs.

$$s_i^{\rightarrow} = W_{EMB} \cdot v_i^s \quad (2)$$

Every sentence is preprocessed and constitutes “ s_i ” and “ d_j ”, where d_j represents drug labels and s_i is another word in the sentence. Every word “ s_i ” is transformed to the word vectors using the word embedding matrices. Word embedding (WEMB) is an embedding matrix and $W_{EMB} \in \mathbb{R}^{d_s \times |V|}$ whereas V denotes the vocabulary in the training dataset, d_s signifies the count of dimensions, and v_i^s denotes the index of word embedding.

3.2. DDI Prediction Process

To predict the DDI accurately, the MLSTM-AE model is applied in this study. The MLSTM-AE model learns to recreate a time flipped version of input [19]. Every input electricity signal is denoted as $\chi_i = \{\chi_i^1, \chi_i^2, \dots, \chi_i^T\}$, and is of length T . The hidden state vectors of long short-term memory (LSTM) encoding at P^h instant are represented as h_i^t . The encoder captures relevant data to recreate the input signals. Once it encodes the final point in the input, the h_i^T hidden state of the encoder is the vector depiction for the input χ_i . The decoding has a similar network architecture as the encoding; however, it learns to recreate a flipped version of e input, viz., $\{\chi_i^T, \chi_i^{T-1}, \chi_i^1\}$. The last hidden state h_i^T of the encoder can be utilized as the first hidden state of decoding input. The targeted output acts as a flipped version of input, viz., $\{\chi_i^T, \chi_i^{T-1}, \dots, \chi_i^1\}$ and the actual recreated one is $\{\hat{\chi}, \hat{\chi}, \dots, \hat{\chi}\}$. The presented method has been demonstrated. Now, the encoder and the decoders are LSTM for modeling dynamic signals. The depiction from the deep layer of the encoder is interconnected with the output label through fully connected networks (FCNs). The reconstruction utilized for training the MLSTM-AE model is formulated by:

$$L_{rec} = \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^T (x_i^t - \hat{x}_i^t)^2 \quad (3)$$

where N denotes the overall sample count. Because, the final objective of the study is to learn to categorize, the embedding from the h_i^T hidden layer is passed via a fully connected (FC) layer, the output of which is the class label. The class label is one-hot encoded. The size of the label vector is equivalent to the number of appliances; once an appliance is ON, the corresponding location of the label vector is 1 or else 0. This can be denoted by $y_i = \{y_i^1, y_i^2, \dots, y_i^c\}$, considering C appliances. Figure 1 represents the structure of MLSTM.

When the C^{th} appliance is ON, the corresponding y_i^c is 1; otherwise it is 0. The ground-truth probability vector of i^{th} samples are described as $\hat{p}_i = y_i / \|y_i\|_1$. The predicted probability vector can be represented as p_i .

$$L_{cls} = \frac{1}{N} \sum_{i=1}^N \sum_{c=1}^c (p_i^c - \hat{p}_i^c)^2 \quad (4)$$

This algorithm has been trained collectively with the reconstruction loss and multilabel classification loss, hence, the overall loss function is formulated by Equation (5):

$$L = L_{cls} + \gamma L_{rec} \quad (5)$$

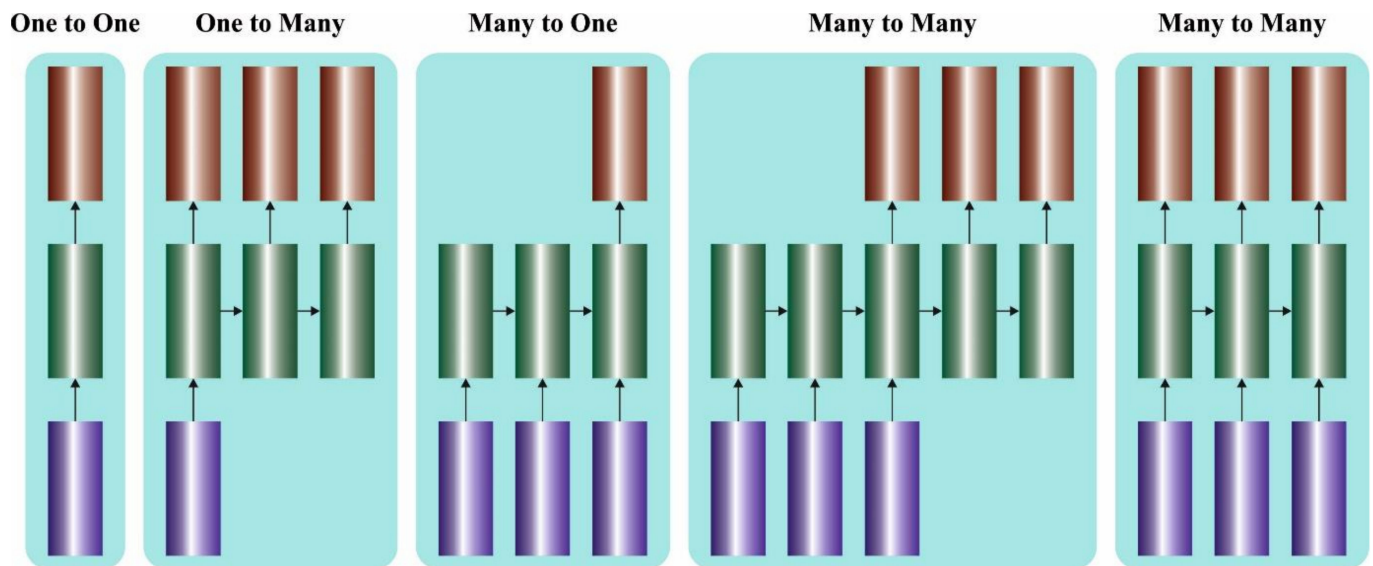


Figure 1. Structure of MLSTM.

3.3. Hyperparameter Tuning Process

For the hyperparameter tuning process, the SSODL-DDIP technique uses the SSO algorithm. The SSO is a recent metaheuristic approach which stimulates the anti-predatory and predation actions of the sparrow population [20], particularly, in foraging, individual sparrows act in two roles: joiner and discoverer. The discoverer is responsible for searching the food and guiding others, and the joiner forages by following the discoverers. A specific percentage of sparrows has been carefully chosen as the guarder that transmits alarm signals and carries out anti-predation behavior while they realize the danger. The discoverer position can be redeveloped as follows:

$$X_{ij}^{t+1} = \begin{cases} X_{ij}^t \cdot \exp\left(-\frac{i}{\alpha \cdot T}\right) & R_2 < ST \\ X_{ij}^t + O \cdot G & R_2 \geq ST \end{cases} \quad (6)$$

In Equation (6), t is the existing value of update. T presents the maximal value of update. X_{ij}^t defines the present position of the i – th agent. X_{ij}^{t+1} denotes the upgraded position of the i – th sparrow in the j – th dimension $\alpha \in (0, 1]$ refers to a random number. $ST \in (0.5, 1]$ signifies a safety value. $R_2 \in (0, 1]$ defines a warning value. G denotes a $1 \times d$ matrix where each value is 1. O represents a random variable.

The joiner position is regenerated as follows:

$$X_{ij}^{t+1} = \begin{cases} O \cdot \exp\left(\frac{x_w - X_{ij}^t}{i^2}\right) & i > n/2 \\ X_b + |X_{ij}^t - X_b| \cdot B \cdot G & \text{otherwise} \end{cases} \quad (7)$$

In Equation (7), X_b signifies the existing optimum position of the discoverer. X_w describes the worst position of the sparrow, B denotes the $1 \times d$ matrix where every value is equivalent to 1 or -1 , and $A^+ = A^T(AA^T)^{-1}$. Figure 2 demonstrates the steps involved in the SSO algorithm.

The position regeneration for the guarder can be defined as follows:

$$X_{ij}^{t+1} = \begin{cases} X_{best}^t + \beta \cdot \left| X_{ij}^t - X_{best}^t \right| & ft_j > ft_g \\ X_{ij}^t + K \cdot \left(\frac{X_{ij}^t - X_{worst}^t}{(ft_i - ft_{tw}) + \epsilon} \right) & ft_i = ft_g \end{cases} \quad (8)$$

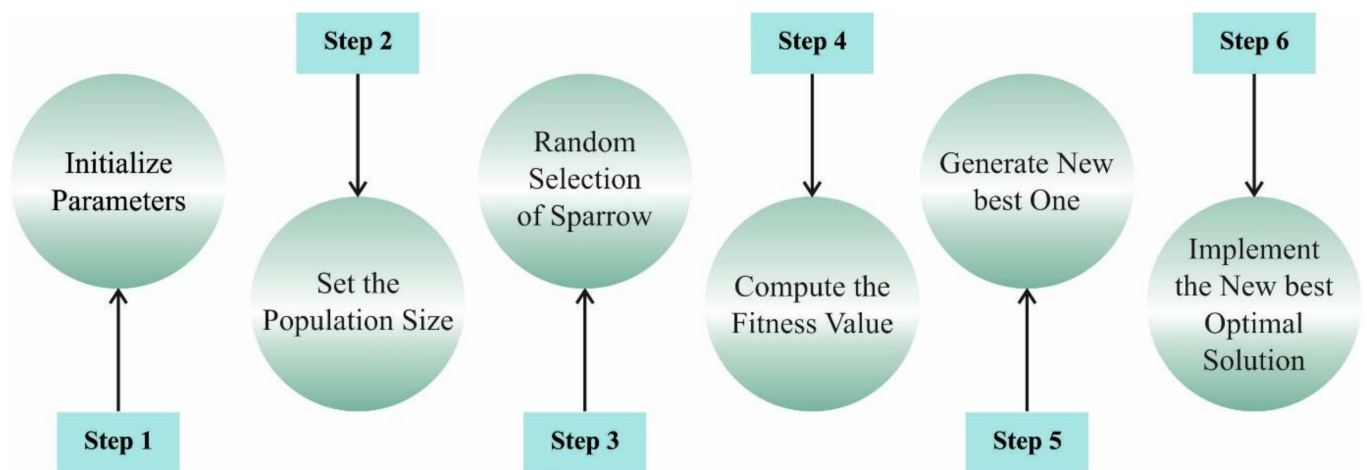


Figure 2. Steps involved in the SSO algorithm.

In Equation (8), X_{best} stand for the best global location. β and $K \in [-1, 1]$ represent two random integers; ft_i defines the fitness value. ft_w and ft_g are the present worst and best fitness values in the population, correspondingly; ϵ indicates a minimal number that is closer to zero as explained in Algorithm 1.

Algorithm 1: Pseudocode of SSO algorithm

Define $Iter_{max}$, NP , n , P_{dp} , sf , G_c , FS_U and FS_L

Arbitrarily initializing the flying squirrels places

$$FS_{i,j} = FS_L + rand() * (FS_U - FS_L), i = 1, 2, \dots, NP, j = 1, 2, \dots, n$$

Compute fitness value

$$f_i = f_i(FS_{i,1}, FS_{i,2}, \dots, FS_{i,n}), i = 1, 2, \dots, NP$$

while $Iter < Iter_{max}$

$$[sorted - f, sorte - index] = sort(f)$$

$$FS_{ht} = FS(sorte_index(1))$$

$$FS_{at}(1 : 3) = FS(sorte_index(2 : 4))$$

$$FS_{nt}(1 : NP - 4) = FS(sorte_index(5 : NP))$$

Create novel places

for $t = 1 : n1$ ($n1$ = entire count of squirrels on acorn trees)

if $R_1 \geq P_{dp}$

$$FS_{at}^{new} = FS_{at}^{old} + d_g G_c (FS_{ht}^{old} - FS_{at}^{old})$$

else

$$FS_{at}^{new} = \text{random location}$$

for $t = 1 : n2$ ($n2$ = entire count of squirrels on normal trees moving to acorn trees)

if $R_2 \geq P_{dp}$

$$FS_{nt}^{new} = FS_{nt}^{old} + d_g G_c (FS_{at}^{old} - FS_{nt}^{old})$$

else

$$FS_{nt}^{new} = \text{random location}$$

end

end

for $t = 1 : n3$ ($n3 =$ entire count of squirrels on normal trees moving to hickory trees)
 if $R_3 \geq P_{dp}$

$$FS_{nt}^{new} = FS_{nt}^{old} + d_g G_c (FS_{ht}^{old} - FS_{nt}^{old})$$

else

$$FS_{nt}^{new} = \text{random location}$$

end

end

$$S_c^t = \sqrt{\sum_{k=-1}^n (FS_{atk}^t - FS_{htk})^2}, S_{c \min} = \frac{10B - 6}{365^{Iter/(Iter_{\max})/2.5}}$$

if $s_c^t < s_{c \min}$

$$FS_{nt}^{new} = FS_L + \text{Lévy}(n) \times (FS_U - FS_L)$$

end

Compute fitness value of novel places

$$f_i = f_i(FS_{i,1}^{new}, FS_{i,2}^{new}, \dots, FS_{i,n}^{new}), i = 1, 2, \dots, NP$$

$$Iter = Iter + 1$$

end

The SSO algorithm derives a fitness function (FF) for reaching maximum classifier performance. It determines positive values for signifying the superior outcome of the candidate solutions. In this article, the reduction of the classifier error rate is the FF, as presented below in Equation (9):

$$\text{fitness}(x_i) = \text{ClassifierErrorRate}(x_i) = \frac{\text{number of misclassified samples}}{\text{Total number of samples}} * 100 \quad (9)$$

3.4. Severity Extraction Process

Lexicons such as Sent WordNet and WordNet Affect are common lexicons that are utilized for extracting common sentiments of texts, for instance, movies and social reviews. The subjectivity lexicon has been utilized for extracting subjective expression in arguments or text statements. Several common and subjectivity lexicons have been changed in medicinal study to distinct healthcare tasks. A wide pharmaceutical lexicon has also progressed specifically to the biomedical and healthcare domains and has been used for extracting the sentiments of clinical and pharmaceutical text. It can extract the polarity of sentences by executing Sent WordNet, and the interface has been classified as low, moderate, or high levels, as dangerous and advantageous DDIs are dependent upon the polarity of candidate sentences.

4. Results and Discussion

The experimental validation of the SSODL-DDIP technique was tested using drug target datasets [10,21]. We used four different datasets to examine the performance of the SSODL-DDIP technique. Table 1 presents the details of the datasets. The distribution of samples under drug, target, and interactions is given in Figure 3.

Table 2 and Figure 4 present the performance of the SSODL-DDIP technique under unlabeled and labeled samples on the top k% values. The results indicate that the SSODL-DDIP technique effectively labeled the samples. For instance, on the top 10% of the enzyme dataset, the SSODL-DDIP technique labeled 317 samples under 29,036 unlabeled samples. Likewise, on the top 10% of the G protein-coupled receptors (GPCR) dataset, the SSODL-DDIP technique labeled 311 samples under 1916 unlabeled samples. Similarly, on

the top 10% of the ion channel dataset, the SSODL-DDIP technique labeled 395 samples under 4026 unlabeled samples. Lastly, on the top 10% of the nuclear receptor dataset, the SSODL-DDIP technique labeled 34 samples under 110 unlabeled samples.

Table 1. Details on the datasets.

Dataset	Drug	Target	Interactions
Enzyme dataset	445	664	2926
Ion channel dataset	210	204	1467
GPCR dataset	223	95	635
Nuclear receptor dataset	54	26	90

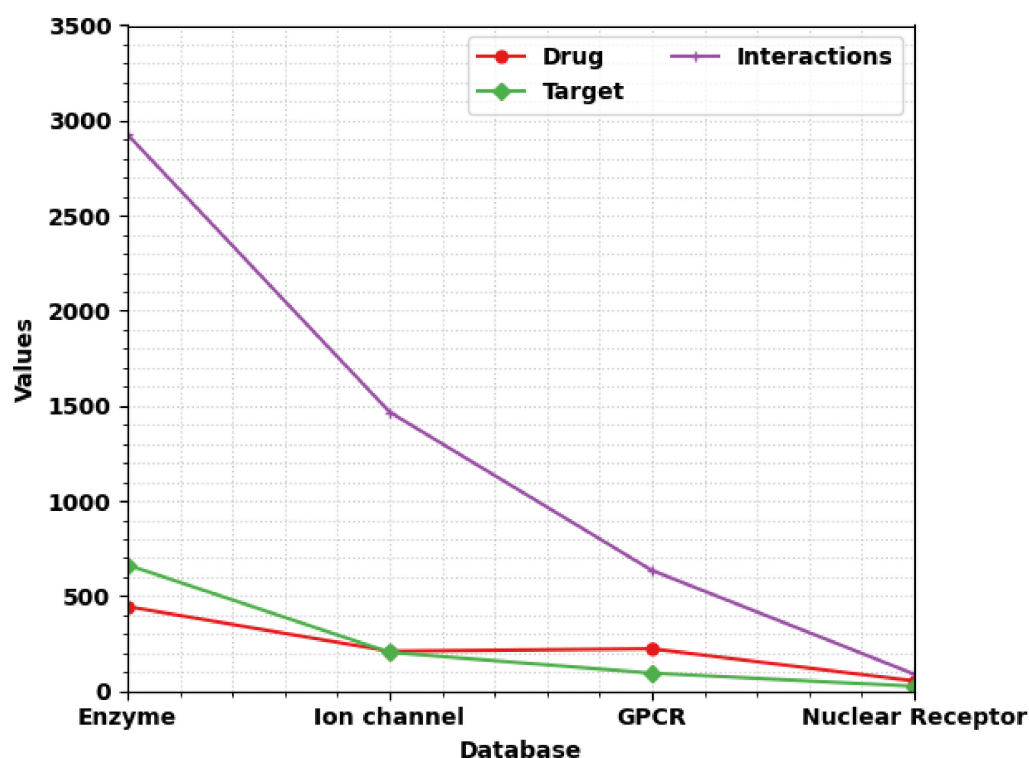


Figure 3. Sample distribution.

Table 3 presents the overall results of the area under the ROC curve (AUC) and the area under the precision-recall curve (AUPR) analysis of the SSODL-DDIP technique on four datasets.

Figure 5 shows the comprehensive AUC values of the SSODL-DDIP technique under different coefficient of variation (CV)_seed values. The figure shows that the SSODL-DDIP technique reached maximum AUC values under all datasets. For instance, on the enzyme dataset, the SSODL-DDIP technique attained higher AUC values of 93.46%, 97.32%, 96.72%, 88.33%, and 97.78% under CV_SEED values of 3201, 2033, 5179, 2931, and 9117, respectively. On the GPCR dataset, the SSODL-DDIP technique attained higher AUC values of 87.09%, 84.82%, 87.17%, 88.50%, and 92.95% under CV_SEED values of 3201, 2033, 5179, 2931, and 9117, respectively.

Figure 6 presents the comprehensive AUPR values of the SSODL-DDIP technique under different CV_seed values. The figure implied that the SSODL-DDIP technique attained maximum AUPR values under all datasets. For example, on the enzyme dataset, the SSODL-DDIP technique attained higher AUPR values of 60.29%, 68.78%, 66.59%, 54.85%, and 71.31% under CV_SEED values of 3201, 2033, 5179, 2931, and 9117 respectively. On the GPCR dataset, the SSODL-DDIP technique attained higher AUPR values of 63.21%,

62.04%, 63.58%, 66.67%, and 68.97% under CV_SEED values of 3201, 2033, 5179, 2931, and 9117 respectively.

Table 2. Analysis results of the SSODL-DDIP technique applied to distinct datasets.

Enzyme Dataset			GPCR Dataset		
Top k (%)	Unlabeled	Labeled	Top k (%)	Unlabeled	Labeled
10	29,036	317	10	1916	311
20	58,173	478	20	3901	461
30	87,431	510	30	5966	497
40	116,727	516	40	8020	541
50	145,973	547	50	10,043	607
60	175,216	578	60	12,097	640
70	204,442	639	70	14,107	690
80	233,718	645	80	16,163	699
90	262,967	682	90	18,214	703
100	292,205	727	100	20,292	719

Ion Channel Dataset			Nuclear Receptor Dataset		
Top k (%)	Unlabeled	Labeled	Top k (%)	Unlabeled	Labeled
10	4026	395	10	110	34
20	8052	736	20	241	36
30	12,219	802	30	374	36
40	16,418	803	40	503	38
50	20,358	1090	50	630	41
60	24,453	1189	60	760	41
70	28,596	1232	70	888	43
80	32,713	1277	80	1017	44
90	36,823	1346	90	1143	48
100	40,911	1422	100	1271	50

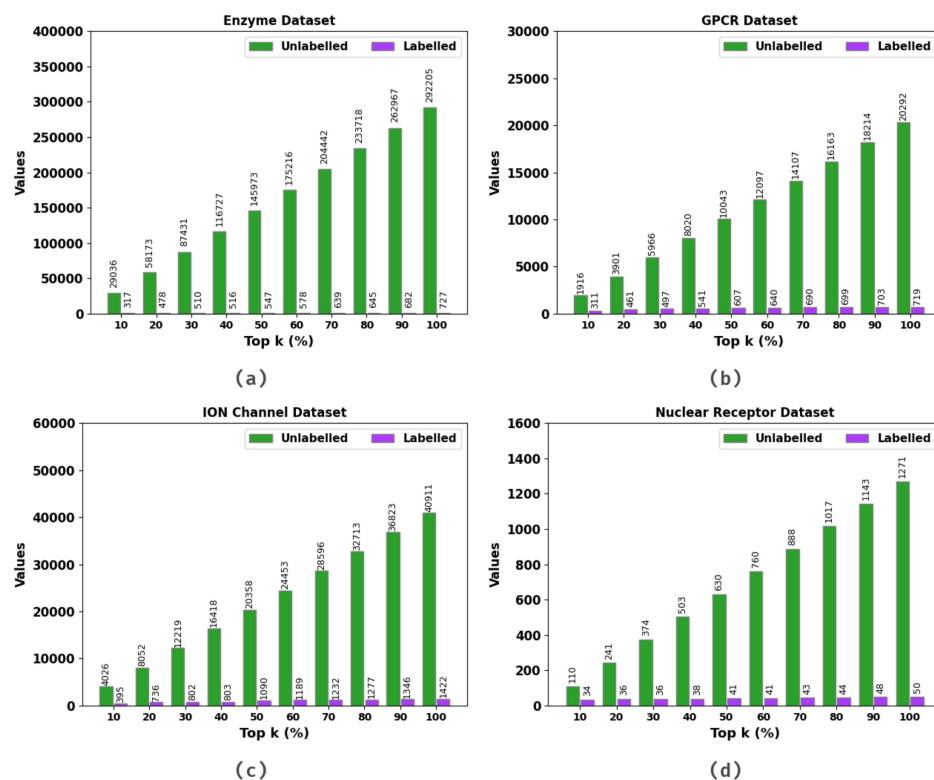


Figure 4. Result analysis of the SSODL-DDIP system: (a) Enzyme; (b) GPCR; (c) ion channel; (d) nuclear receptor.

Table 3. AUC and AUPR analysis of the SSODL-DDIP system under distinct datasets.

Enzyme Dataset			GPCR Dataset		
CV_SEED	AUC	AUPR	CV_SEED	AUC	AUPR
3201	93.46	60.29	3201	87.09	63.21
2033	97.32	68.78	2033	84.82	62.04
5179	96.72	66.59	5179	87.17	63.58
2931	88.33	54.85	2931	88.50	66.67
9117	97.78	71.31	9117	92.95	68.97

Ion Channel Dataset			Nuclear Receptor Dataset		
CV_SEED	AUC	AUPR	CV_SEED	AUC	AUPR
3201	83.71	61.46	3201	91.67	75.35
2033	88.11	66.96	2033	94.79	76.65
5179	91.94	67.55	5179	98.08	82.93
2931	83.98	63.18	2931	98.13	86.05
9117	92.00	70.34	9117	98.85	87.94

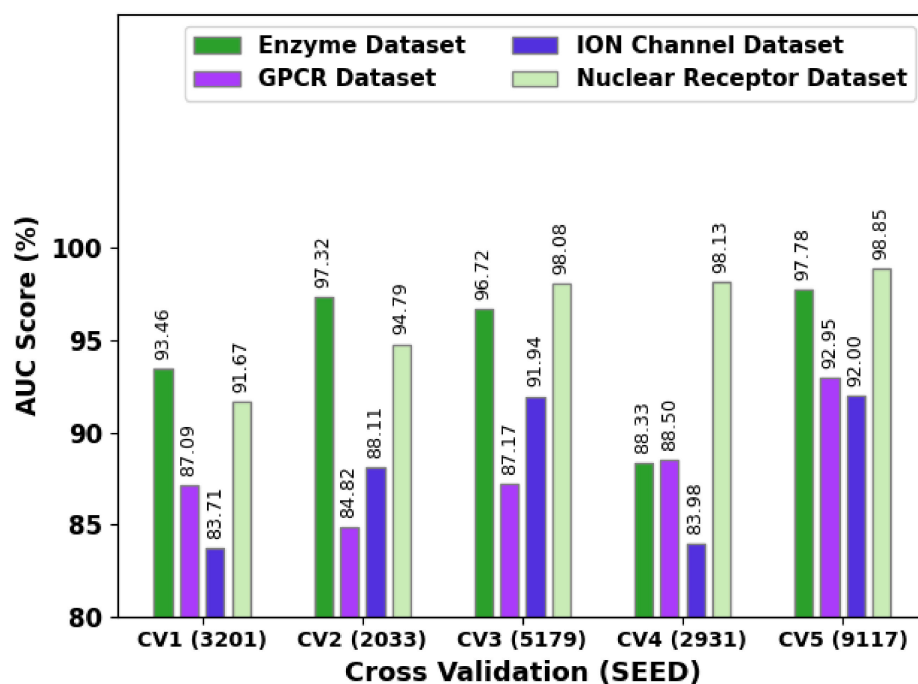
**Figure 5.** AUC analysis of the SSODL-DDIP technique under different CV_seed values.

Table 4 and Figure 7 show the results of a comparison study of the SSODL-DDIP technique on four datasets in terms of AUC [22–25]. The experimental values indicate that the SSODL-DDIP technique attained maximum AUC values under all datasets. For instance, on the enzyme dataset, the SSODL-DDIP technique attained a higher AUC value of 97.78%. In contrast, the bigram position-specific scoring matrix (PSSM), neural network (NN), IFB, kernelized Bayesian matrix factorization with twin kernels' (KBMF2K), drug-based similarity inference (DBSI), and drug–target interaction prediction model using optimal recurrent neural network (DTIP-ORNN) technique attained lower AUC values of 86%, 94.80%, 89.80%, 84.50%, 83.20%, 80.60%, and 96.10% respectively. On the GPCR dataset, the SSODL-DDIP technique attained a higher AUC value of 92.95%. Conversely, the bigram PSSM, NN, IFB, KBMF2K, DBSI, and DTIP-ORNN technique attained lower AUC values of 86%, 87.60%, 88.90%, 88.90%, 81.20%, 85.70%, 80.30%, and 91.53%, respectively.

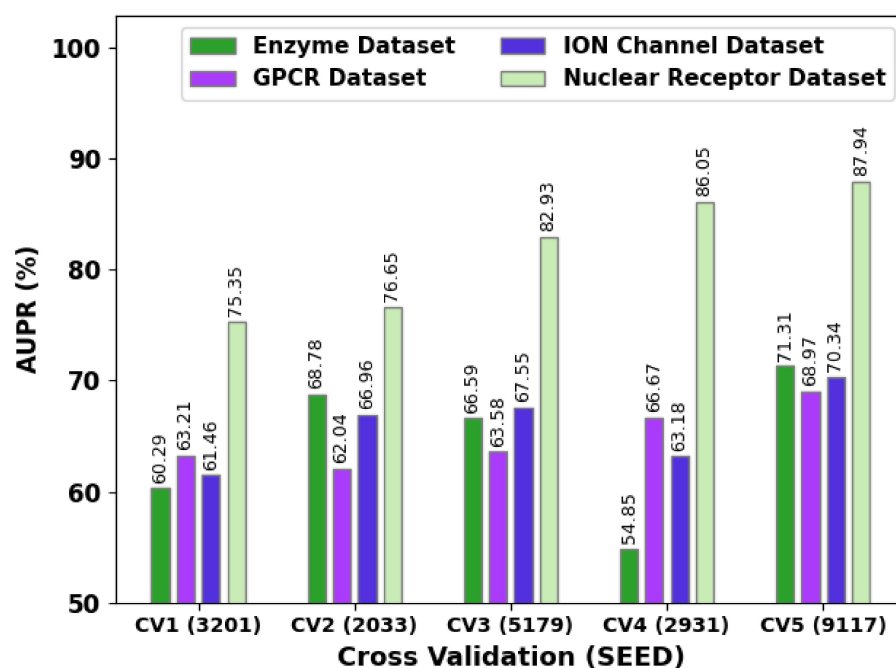


Figure 6. AUPR analysis of the SSODL-DDIP system under different CV_seed values.

Table 4. Comparative analysis of the SSODL-DDIP technique on different datasets in terms of AUC.

Methods	Enzyme	GPCR	ION Channel	Nuclear Receptor
UDTPP	86.00	87.60	77.50	80.00
Bi-gram PSSM	94.80	88.90	87.20	86.90
Nearest neighbor	89.80	88.90	85.20	82.00
IFB model	84.50	81.20	73.10	83.00
KBMF2K	83.20	85.70	79.90	82.40
DBSI	80.60	80.30	80.30	75.90
DTIP-ORNN	96.10	91.53	90.14	98.72
SSODL-DDIP	97.78	92.95	92.00	98.85

Table 5 and Figure 8 present a comparative inspection of the SSODL-DDIP technique on four datasets in terms of AUPR. The simulation values indicate that the SSODL-DDIP technique attained maximum AUPR values under all datasets. For instance, on the enzyme dataset, the SSODL-DDIP technique attained a higher AUPR value of 71.31%. In contrast, the bipartite local model (BLM), self-training support vector machine with BLM (SELF-BLM), positive-unlabeled learning with BLM (PULBLM)-3, PULBLM-5, PULBLM-7, and DTIP-ORNN technique attained lower AUPR values of 57.00%, 63.00%, 67.00%, 67.00%, 66.00%, and 69.01% respectively. In addition, on the GPCR dataset, the SSODL-DDIP technique attained a higher AUPR value of 68.97%. In contrast, the bigram BLM, SELF-BLM, PULBLM-3, PULBLM-5, PULBLM-7, and DTIP-ORNN technique attained lower AUPR values of 55.00%, 60.00%, 64.00%, 64.00%, 65.00%, and 67.20%, respectively. These results confirmed the effective DDI prediction results of the SSODL-DDIP technique.

Table 5. Comparative analysis of the SSODL-DDIP technique on different datasets in terms of AUPR.

Methods	Enzyme	GPCR	ION Channel	Nuclear Receptor
BLM	57.00	55.00	47.00	42.00
SELF-BLM	63.00	60.00	51.00	45.00
PULBLM-3	67.00	64.00	60.00	58.00
PULBLM-5	67.00	64.00	61.00	59.00
PULBLM-7	66.00	65.00	63.00	59.00
DTIP-ORNN	69.01	67.20	68.12	86.38
SSODL-DDIP	71.31	68.97	70.34	87.94

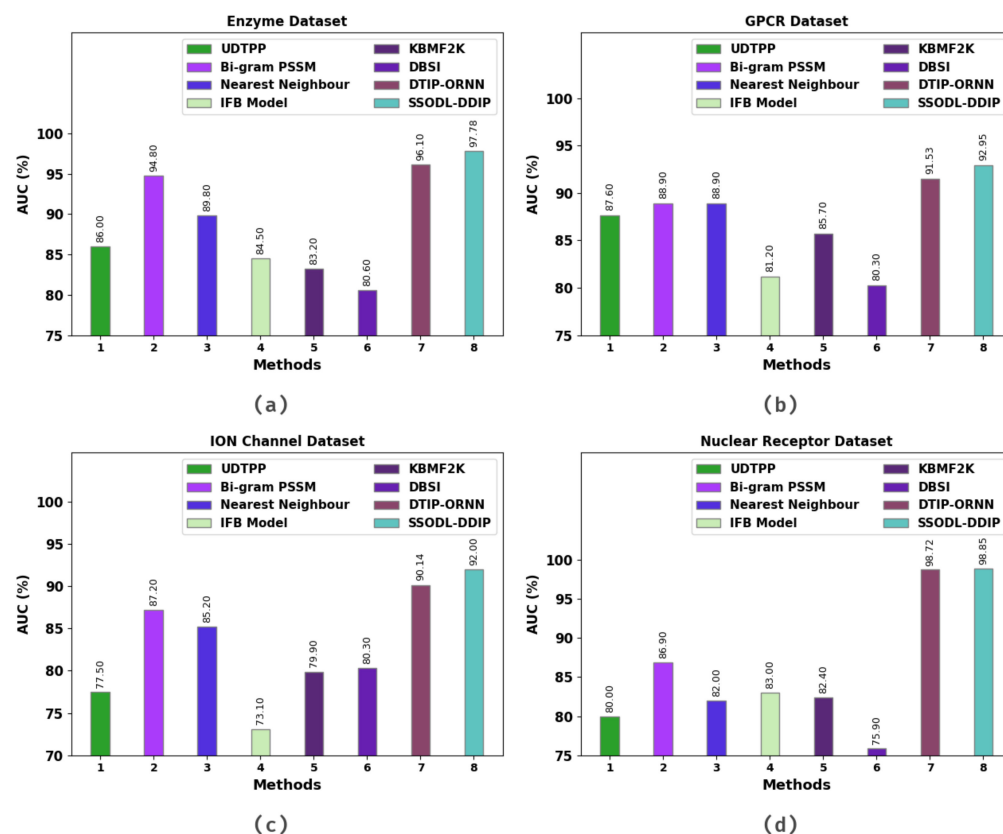


Figure 7. AUC analysis of the SSODL-DDIP technique: (a) Enzyme; (b) GPCR; (c) ion channel; (d) nuclear receptor.

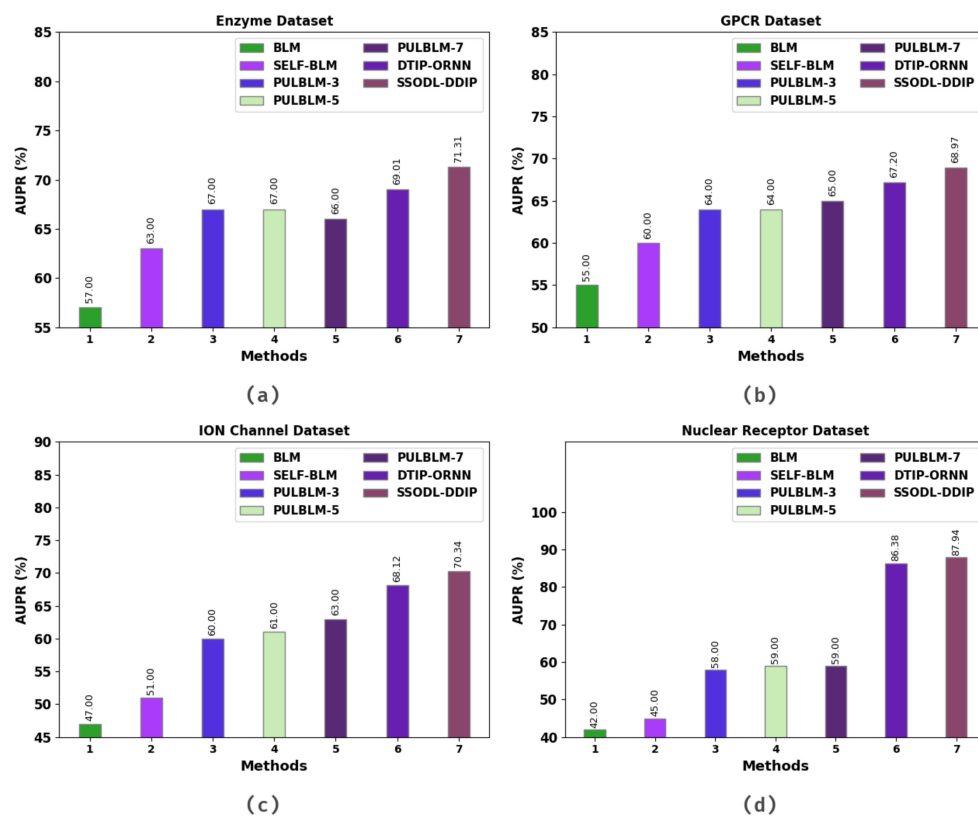


Figure 8. AUPR analysis of the SSODL-DDIP technique: (a) Enzyme; (b) GPCR; (c) ion channel; (d) nuclear receptor.

5. Conclusions

In this study, we introduced a novel SSODL-DDIP technique for DDI predictions in big data environments. The presented SSODL-DDIP technique accurately determined the relationship and drug properties from various sources to make a prediction. In addition, the MLSTM-AE model was employed for the DDI prediction process. Furthermore, a lexicon-based approach was involved in determining the severity of interactions among the DDIs. To improve the prediction outcomes of the MLSTM-AE model, the SSO algorithm was adopted in this work. To assure better performance of the SSODL-DDIP technique, a wide range of simulations were performed. The experimental outcomes show the promising performance of the SSODL-DDIP technique over recent state-of-the-art methodologies. Thus, the SSODL-DDIP technique can be employed for improved DDI predictions. In the future, hybrid metaheuristics could be designed to improve the prediction performance. In addition, outlier detection and clustering techniques could be integrated to enhance the predictive results of the proposed model.

Author Contributions: Conceptualization, A.M.H. and M.I.E.; methodology, F.A.; software, M.I.E.; validation, S.S.A., R.M., A.A.A. and M.I.E.; formal analysis, H.M.; investigation, A.A.A.; resources, H.M.; data curation, R.M. and M.I.E.; writing—original draft preparation, A.M.H., F.A., S.S.A., R.M. and H.M.; writing—review and editing, A.A.A., A.E.O. and M.I.E.; visualization, M.I.E. and A.E.O.; supervision, F.A.; project administration, A.M.H.; funding acquisition, F.A. and S.S.A. All authors have read and agreed to the published version of the manuscript.

Funding: Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R77), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The authors would like to thank the Deanship of Scientific Research at Umm Al-Qura University for supporting this work by Grant Code: (22UQU4210118DSR53). This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2023/R/1444).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing is not applicable to this article as no datasets were generated during the current study.

Conflicts of Interest: The authors declare that they have no conflict of interest. The manuscript was written through contributions of all authors. All authors have given approval for the final version of the manuscript.

Abbreviations

Abbreviation	Meaning
DDI	Drug–drug interactions
ML	Machine learning
DL	Deep learning
SSODL-DDIP	Sparrow search optimization with deep learning-based DDI prediction
MLSTM-AE	Multilabel long short-term memory with an autoencoder
DM	Data mining
RBM	Restricted Boltzmann machines
CNN	Convolution neural networks
AE	Autoencoder
PU	Positive unlabeled
SVM	Support vector machine
GCN	Graph convolution network
DNN	Deep neural networks
KGNN	Knowledge graph neural network
KG	Knowledge graph
AMDE	Attention-related multidimensional feature encoders
SEV-DDI	Severity DDI

DANN-DDI	Deep attention neural network-related DDI
SFLLN	Sparse feature learning ensembled approach with linear neighborhood regularization
DP	Drug pairs
WEMB	Word embedding
LSTM	Long short-term memory
FCN	Fully connected networks
FC	Fully connected
FS	Flying Squirrels
GPCR	G protein-coupled receptors
AUPR	Area under the precision-recall curve
AUC	Area under the ROC Curve
CV	Coefficient of variation
PSSM	Position-specific scoring matrix
NN	Neural network
KBMF2K	Kernelized Bayesian matrix factorization with twin kernels
DBSI	drug-based similarity inference
DTIP-ORNN	Drug–target interaction prediction model using optimal recurrent neural network
BLM	Bipartite local model
SELF-BLM	Self-training support vector machine with BLM
PULBLM	Positive unlabeled learning with BLM

References

1. Ryu, J.Y.; Kim, H.U.; Lee, S.Y. Deep learning improves prediction of drug–drug and drug–food interactions. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E4304–E4311. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Hung, T.N.K.; Le, N.Q.K.; Le, N.H.; Van Tuan, L.; Nguyen, T.P.; Thi, C.; Kang, J.H. An AI-based Prediction Model for Drug-drug Interactions in Osteoporosis and Paget’s Diseases from SMILES. *Mol. Inform.* **2022**, *41*, 2100264. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Wang, N.N.; Wang, X.G.; Xiong, G.L.; Yang, Z.Y.; Lu, A.P.; Chen, X.; Liu, S.; Hou, T.J.; Cao, D.S. Machine learning to predict metabolic drug interactions related to cytochrome P450 isozymes. *J. Cheminform.* **2022**, *14*, 1–16. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Kastrin, A.; Ferk, P.; Leskošek, B. Predicting potential drug-drug interactions on topological and semantic similarity features using statistical learning. *PLoS ONE* **2018**, *13*, e0196865. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Deng, Y.; Xu, X.; Qiu, Y.; Xia, J.; Zhang, W.; Liu, S. A multimodal deep learning framework for predicting drug–drug interaction events. *Bioinformatics* **2020**, *36*, 4316–4322. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Kumar, R.; Saha, P. A review on artificial intelligence and machine learning to improve cancer management and drug discovery. *Int. J. Res. Appl. Sci. Biotechnol.* **2022**, *9*, 149–156.
7. Lim, S.; Lee, K.; Kang, J. Drug drug interaction extraction from the literature using a recursive neural network. *PLoS ONE* **2018**, *13*, e0190926. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Chen, Y.; Ma, T.; Yang, X.; Wang, J.; Song, B.; Zeng, X. MUFFIN: Multi-scale feature fusion for drug–drug interaction prediction. *Bioinformatics* **2021**, *37*, 2651–2658. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Vilar, S.; Friedman, C.; Hripcsak, G. Detection of drug–drug interactions through data mining studies using clinical sources, scientific literature and social media. *Brief. Bioinform.* **2018**, *19*, 863–877. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Yamanishi, Y.; Araki, M.; Gutteridge, A.; Honda, W.; Kanehisa, M. Prediction of drug–target interaction networks from the integration of chemical and genomic spaces. *Bioinformatics* **2008**, *24*, i232–i240. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Wang, Y.; Min, Y.; Chen, X.; Wu, J. Multi-view graph contrastive representation learning for drug-drug interaction prediction. In Proceedings of the Web Conference, Online, 19–23 April 2021; pp. 2921–2933.
12. Yuvaraj, N.; Srihari, K.; Chandragandhi, S.; Raja, R.A.; Dhiman, G.; Kaur, A. Analysis of protein-ligand interactions of SARS-Cov-2 against selective drug using deep neural networks. *Big Data Min. Anal.* **2021**, *4*, 76–83. [\[CrossRef\]](#)
13. Lin, X.; Quan, Z.; Wang, Z.J.; Ma, T.; Zeng, X. KGNN: Knowledge Graph Neural Network for Drug-Drug Interaction Prediction. *IJCAI* **2020**, *380*, 2739–2745.
14. Pang, S.; Zhang, Y.; Song, T.; Zhang, X.; Wang, X.; Rodriguez-Patón, A. AMDE: A novel attention-mechanism-based multidimensional feature encoder for drug–drug interaction prediction. *Brief. Bioinform.* **2022**, *23*, bbab545. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Salman, M.; Munawar, H.S.; Latif, K.; Akram, M.W.; Khan, S.I.; Ullah, F. Big Data Management in Drug–Drug Interaction: A Modern Deep Learning Approach for Smart Healthcare. *Big Data Cogn. Comput.* **2022**, *6*, 30. [\[CrossRef\]](#)
16. Liu, S.; Zhang, Y.; Cui, Y.; Qiu, Y.; Deng, Y.; Zhang, Z.M.; Zhang, W. Enhancing drug-drug interaction prediction using deep attention neural networks. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2022**. [\[CrossRef\]](#) [\[PubMed\]](#)

17. Zhang, W.; Jing, K.; Huang, F.; Chen, Y.; Li, B.; Li, J.; Gong, J. SFLLN: A sparse feature learning ensemble method with linear neighborhood regularization for predicting drug–drug interactions. *Inf. Sci.* **2019**, *497*, 189–201. [[CrossRef](#)]
18. Rastegar-Mojarad, M.; Boyce, R.D.; Prasad, R. UWM-TRIADS: Classifying Drug-Drug Interactions with Two-Stage SVM and Post-Processing. In Proceedings of the SEM 2013-2nd Joint Conference on Lexical and Computational Semantics, Atlanta, GA, USA, 12 June 2013; Volume 2, pp. 667–674.
19. Verma, S.; Singh, S.; Majumdar, A. Multi-label LSTM autoencoder for non-intrusive appliance load monitoring. *Electr. Power Syst. Res.* **2021**, *199*, 107414. [[CrossRef](#)]
20. Luan, F.; Li, R.; Liu, S.Q.; Tang, B.; Li, S.; Masoud, M. An Improved Sparrow Search Algorithm for Solving the Energy-Saving Flexible Job Shop Scheduling Problem. *Machines* **2022**, *10*, 847. [[CrossRef](#)]
21. Available online: <http://web.kuicr.kyoto-u.ac.jp/supp/yoshi/drugtarget/> (accessed on 12 September 2022).
22. Rajpura, H.R.; Ngom, A. Drug target interaction predictions using PU-Leaming under different experimental setting for four formulations namely known drug target pair prediction, drug prediction, target prediction and unknown drug target pair prediction. In Proceedings of the 2018 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), Saint Louis, MO, USA, 30 May–2 June 2018; pp. 1–7.
23. Lan, W.; Wang, J.; Li, M.; Liu, J.; Li, Y.; Wu, F.X.; Pan, Y. Predicting drug–target interaction using positive-unlabeled learning. *Neurocomputing* **2016**, *206*, 50–57. [[CrossRef](#)]
24. Haddadi, F.; Keyvanpour, M.R. PULBLM: A Computational Positive-Unlabeled Learning Method for Drug-Target Interactions Prediction. In Proceedings of the 10th International Conference on Information and Knowledge Technology (IKT 2019), Tehran, Iran, 31 December 2019.
25. Kavipriya, G.; Manjula, D. Drug–Target Interaction Prediction Model Using Optimal Recurrent Neural Network. *Intell. Autom. Soft Comput.* **2023**, *35*, 1677–1689. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.