International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 04, 2020
ISSN: 1475-7192

An Intelligent Technique for Predicting Adverse Drug Reactions (ADRs) Using Modified Shrinkage Function Based Extreme Gradient Boost (MSXGB) Classifier

C. Rathika and Dr.V. Umarani

Abstract*—It has turned out to be highly significant for the prediction of Adverse Drug Reactions (ADRs) in consequence of the massive universal health troubles and malfunction of drugs. This signifies the necessity for earlier prediction of possible ADRs in preclinical phases that could recover failures of drug and decrease the duration and expense of progress and therefore offering effective and harmless therapeutic decisions for patients. Despite the fact that different techniques have been proposed for prediction of ADR, still there exists need to enhance the classification and precision in predicting adverse drug reactions. This research work presents a framework for discovering the side effects using the best feature selection and classification methods. The proposed research work presents a new approach to measure the features (attributes) in drug prediction dataset using the Modified Orthogonal Opposition Learning based Cuckoo Search Algorithm (MOOLCSA). Finally in this work, Modified Extreme Gradient Boost (MXGB)is used as a classifier that acts as a supervised machine learning approach and could predict Adverse Drug Reactions. This proposed approach holds the benefits of better predictability, interpretability and is an intelligent technique that is beneficial for both patients and also the medical researchers. The simulation result shows that the proposed Modified Shrinkage function based Extreme Gradient Boost (MSXGB) classifier model provides best side effect prediction compared to the other prediction techniques.*

***Keywords*—** *Drug Side Effect, Predictive Modeling, Machine Learning, Feature Selection, Modified Orthogonal Opposition Learning based Cuckoo Search Algorithm (MOOLCSA), Classification; Modified Shrinkage Function based Extreme Gradient Boost (MSXGB) Classifier Model.*

I. INTRODUCTION

For the general business achievement of drugs, predicting and neutralizing the side effects of a drug in the course of its developmental period continues to be very vital. Side Effects are accountable for a major count of instances in which pre-marketed at some point in clinical trials the drugs are unsuccessful. It remains a difficult job to spot out the processes lying behind the side effects and this is due to the pleiotropic effects of the drugs on a biological system [1]. Many of the drugs are tiny composites which aim and act together with proteins to provoke perturbations in the network of proteins. This emphasizes the necessity for system-wide mechanisms for predicting side effects of

C. Rathika, Research Scholar, Sri Ramakrishna College of Arts and Science for Women, Tamilnadu, India.
E-mail: rathikaprakash@gmail.com
Dr.V. Umarani, Assistant Professor(Sr.Gr),PSG College of Technology, Tamilnadu, India.
E-mail: v_umarani@yahoo.com

DOI: 10.37200/IJPR/V24I4/PR201596
Received: 17 Jan 2020 | Revised: 19 Feb 2020 | Accepted: 28 Feb 2020

drug by associating various degrees of drug actions, relationship between drug–protein and interaction existing among drugs and side effects [2].

In contemporary innovations in drug, Adverse Drug Reactions (ADRs) are one among the crucial troubles. Uncommon and severe ADRs are accountable for pipelines of disastrous discovery of drug and for abandonments of drug market. Aggregate expenses of the maintenance of ADRs have been assessed as around 30 billion per year [3]. Trustworthy methodologies for well advanced prediction of side effects of drug are absolutely required, and they will take the advantage from structures of systems analysis which are widely predictive over tissues and forms of human cell. To distinguish the likely drug side effects, the most extensively applied technique is to employ information of its chemical structure, on the basis of the thought that chemical structures of drug could lead in the direction of protein targets, the ligand promiscuity [4].

Identifying potential ADRs of drug candidates in the early stage of the drug development pipeline can improve drug safety, reduce risks for the patients and save money for the pharmaceutical companies [5]. The information available in the early stages of drug development is mainly the chemical structure of the drug candidate. Many existing studies on ADR prediction have been devoted to analyzing the chemical properties of drug molecules [6]. Though the mechanisms of ADRs are complicated and may not be well understood, machine learning techniques are promising solutions to understand and analyze such complicated problems.

In general, the basic steps of ADR prediction based on structural information can be broken down into two stages. First, each drug molecule is represented in a suitable feature vector based on its chemical structure [7]. Second, a machine learning algorithm is applied on the resulting feature space to predict ADRs. So far, most of the existing studies focused on the second step, or the method development, to improve the prediction power. However, how to represent the drug molecules by a useful set of features and how to interpret their effects on the final ADR predictions remain relatively less explored [8]. Note that finding the specific sub-structures of the drug molecule that is related to an ADR can be particularly useful for finding the mechanism of actions of the drug and thus, can be utilized in the early phase of drug design.

Many factors that directly or indirectly cause ADRs varying from pharmacological, immunological and genetic factors to ethnic, age, gender, social factors as well as drug and disease related ones [9]. It is worth noting that ADR detection and prediction methods intrinsically relate to understanding of the causal factors. Different from ADR detection and prediction methods in pre-market surveillance that are done in vivo or in vitro, the methods in post-market surveillance are mostly data-driven with data collected from the patient drug usage [10]. All data-driven methods depend on two components of data sources and computational methods.

In early days the ADR detection methods mainly exploited data from spontaneous adverse event reporting systems or administrative databases with conventional techniques of statistics [11]. On one hand, due to the technology progress, in the last decade several new kinds of data, notably omics data, social media data and electronic medical records (EMRs), have been generating and offering more

DOI: 10.37200/IJPR/V24I4/PR201596
Received: 17 Jan 2020 | Revised: 19 Feb 2020 | Accepted: 28 Feb 2020

chances to detect and predict ADRs. On the other hand, advanced methods of statistics, machine learning and data mining allow the users to more effectively analyze the data for descriptive and predictive purposes.

To address the deficiency, this research work presents a framework for discovering the side effects using the best feature selection and classification methods. The proposed research work presents a new approach to measure the features (attributes) in drug prediction dataset using the Modified Orthogonal Opposition Learning based Cuckoo Search Algorithm (MOOLCSA).

Finally in this work, Modified Shrinkage function based Extreme Gradient Boost (MSXGB) model uses a classifier that exists as a supervised machine learning approach and the ADRs' side effects could be predicted. This proposed model has improved interpretability and predictability. The proposed intelligent technique is beneficial for both patients and medical investigators.

The remaining of this research work is structured as, Section 1 describes for discovering the side effects of a drug and the importance of drug side effect prediction model. In section 2 review the some recent technique of drug side effect prediction model. In section 3 explains the predictive model using the feature selection and classification technique. In section 4 the results and discussion are submitted. In section 5, the conclusion and future work are presented.

II. LITERATURE REVIEW

Tatonetti et al [12]presented an adaptive data-driven technique for amending the aspects under circumstances where the covariates are not measured or unidentified and link this methodology with current systems to enhance assessment of impacts of drug exploiting three sets of test data. A wide-ranging drug effects database and aside effects for drug-drug interaction database is also represented.

In order to illustrate the biological application of these new reserves, utilised them to recognize the targets of drug, predict indications of drug, and uncover interactions of drug class. This study recommends that a collective treatment with choosy serotonin reuptake inhibitors and thiazides is correlated with considerably amplified occurrence of lengthy QT periods. To end with, it complete with the facts that perplexing consequences from covariates in observational medical data could be regulated in analysis of data and hence enhance the prediction and detection of adversative drug interactions and impacts..

Yamanishi et al [13] developed a novel approach to predict possible side-effect descriptions of drug candidate molecules depending on their chemical organizations and target protein data in huge amount. The paper propose various expansions of kernel regression representation for different reactions to handle the dissimilar sources of data.

The uniqueness rests on the combination of the chemical space of drug chemical structures and the biological space of drug target proteins in an integrated structure. Thus it illustrates the effectiveness of the proposed approach on the concurrent prediction of 969 side-effects for permitted drugs from their target protein profiles and chemical

DOI: 10.37200/IJPR/V24I4/PR201596
Received: 17 Jan 2020 | Revised: 19 Feb 2020 | Accepted: 28 Feb 2020

substructure and reveal that the precision of prediction constantly progresses as a result of the proposed regression model and chemical and biological information integration.

Zhang et al [14] proposed an integrative label propagation framework for predicting Drug-Drug Integration (DDIs) by incorporating Side Effects (SEs) obtained from prescribed drugs' package inserts, SEs derived from FDA Adverse Event Reporting System and chemical structures obtained from PubChem. Research outcomes depending on hold-out authentication determined the efficiency of the algorithm proposed.

Besides, the algorithm also ordered sources of drug information on the basis of their influences on the prediction, thereby assuring not only that SEs are significant aspects for prediction of DDI and laying the route for construction still trustworthy DDI prediction techniques also by placing in order the different sources of data. The DDIs that are predicted would assist the clinicians to prevent threatening interactions of drug in their recommendations and also would help pharmacological corporations to sketch medical experiment in large-scale by analysing hypothetically harmful blends of drugs.

Gottlieb et al [15] presented an innovative approach for the prediction of drug indications (PREDICT) in large-scale which could tackle both unique molecules and accepted drugs. This technique is grounded on the examination that analogous drugs are specified for diseases alike, and uses different drug–drug and disease–disease resemblance procedures for the prediction activity. By means of cross-validation, it attains excessive sensitivity and specificity (AUC¼0.9) in drug indication prediction, outstanding current approaches.

The work also endorses the predictions by their overlay with indications of drug which are presently under medical examinations, and by their understanding with the information on tissue-specific representation over the targets of drug. It indicates that genetic signatures which are disease specific could be applied to predict precisely the drug indications for diseases that emerge recently (AUC¼0.92).

Cheng et al [16] described a novel framework for computation to predict poly pharmacological drugs profiles by the incorporation of therapeutic, side effect and chemical space. Based on the database of drug side effects that was developed already which was termed as Meta ADEDB, a drug side effect similarity inference (DSESI) approach was formed for drug–target interaction (DTI) prediction on an accepted DTI network linking 621 drugs that were official and 893 target proteins.

The zone below the curve of receiver operating feature was 0.882 ± 0.011 be an average of 100 simulated tests of 10-fold cross-validation for the DSESI methodology that is proportional with drug structural similarity inference and drug therapeutic similarity inference approaches. Seven different predicted candidate target proteins for seven accepted drugs were established by experimentations published, with the productive success degree greater than 15.9%. The outcomes reveal that the methods proposed can be effective for poly pharmacological drug profile prediction.

Zhao et al [17] presented a unique computational method for drug combinations prediction by collaborating pharmacological and molecular information. Particularly, drugs are denoted by a group of their attributes like their indications and targets.

DOI: 10.37200/IJPR/V24I4/PR201596
Received: 17 Jan 2020 | Revised: 19 Feb 2020 | Accepted: 28 Feb 2020

By means of combining many of these traits, express that characteristic models augmented in drug combinations that are approved are not only predictive for latest combinations of drug, however offer understandings into processes of fundamental combinatorial treatment also. Additional investigation established that between the high rated predictions of effectual blends, 69% whereas the remaining signifies the new possible combinations of drug. It is considered that the proposed methodology can assist to constraint the drug combinations' search space and show a different path to efficiently exploit current drugs for further reasons.

Shaked et al [18] described the Array of Model-Based Phenotype Predictors (AMPP), a methodology which influences resources on medical informatics and a human genome-scale metabolic model (GSMM) in order to predict the side effects of drug. AMPP is significantly predictive (AUC > 0.7) for more than 70 side effects of drugs, together with highly dangerous ones like extra pyramidal and interstitial nephritis complaints.

AMPP overtakes an existing biochemical structure-based technique for metabolic dependent side effects prediction. Notably, AMPP empowers the detection of principal metabolic activities and biomarkers which are predictive of particular side effects. Collectively considered, this research sets groundwork for forthcoming identification of metabolically based side effects for the period of primary phases of development of drug.

Pauwels et al [19] proposed a different approach to predict possible drug side-effects of candidate molecules grounded on their chemical structures, appropriate for bigger databanks of molecules. An exceptional characteristic of the proposed technique is its capability to derive associated chemical substructures set and side-effects. This feasibility is by applying sparse canonical correlation analysis (SCCA). The outputs, depicts the effective nature of the method proposed by predicting 1385 side-effects in the SIDER database from the chemical structures of 888 permitted drugs.

These predictions are carried out with parallel derivation of associated collections created by a group of chemical substructures common for drugs which are possible to possess a collection of side-effects. It direct a detailed prediction of side-effect for several uncharacterized drug molecules gathered in Drug Bank, and was capable of substantiating predictions that are attractive utilizing information from autonomous source.

Mizutani et al [20] developed an approach for side effects prediction from data of chemical structure applying canonical correlation analysis (CCA). This work was revolutionary based on the concurrent prediction of several side effects. This work put forth a process for linking fragments of drug chemical with side effects by means of sparse CCA (SCCA), and utilised the chemical fragments to predict profiles of side effect.

This research has done an evaluation on large-scale to extract side effects and correlated sets of targeted proteins on the basis of the co-incidence of profiles of side effect and profiles of drugs in protein-binding, by means of scarce canonical correlation assessment. The examination of 658 drugs with the two profiles for 1368 proteins and 1339 side effects directed to the derivation of 80 correlated sets. Probably the proposed approach beneficial for possible side effects of new drug candidate compounds prediction founded on their profiles of protein-binding.

Huang et al [21] devised a framework meant for predicting profiles of Adverse Drug Reaction (ADR) by incorporating protein–protein interaction (PPI) networks with drug structures. The work also linked operations of

DOI: 10.37200/IJPR/V24I4/PR201596
Received: 17 Jan 2020 | Revised: 19 Feb 2020 | Accepted: 28 Feb 2020

ADR prediction through 18 ADR classes over groups of four feature drug targets only, drug targets having PPI networks, drug structures, and drug targets with PPI networks in addition with drug structures. The outcomes revealed that the PPI networks integration and drug structures could considerably enhance the performance of ADR prediction. The average AUC quantities for the four groups were 0.59, 0.61, 0.65, and 0.70. Protein attributes in the best two models, “Cardiac disorders” and “Psychiatric disorders” are used to construct ADR-specific PPI networks. Apart from the three drugs holding ADRs in the groups did not predict 25 out of 27 inhibited drugs with acute ADRs were predicted by this method effectively.

Takarabeet al [22] proposed an approach for predicting unidentified drug–target correlations on a large scale from the association of pharmacological resemblance of drugs and genomic sequence correspondence of target proteins in the structure of a pharmacogenomic method.

The method proposed is appropriate to a huge count of drugs and it was helpful for unknown drug–target interactions prediction in particular, which cannot be anticipated from structures of drug chemicals. A thorough prediction for possible off-targets of 1874 drugs having identified targets and probable target profiles of 2519 drugs deprived of identified targets, which recommends different possible drug–target interactions which were not predicted by existing chemo genomic or pharmacogenomic methodologies.

Atias et al [23] introduced a new technique for side effects prediction of a particular drug. Commencing from a drug under inquiry, a mixture of network-based diffusion and canonical correlation analysis are used for prediction of its side effects. Both methods are trained using data on known drug-side effect relations. They exploit molecular data on the drugs like their chemical structure or the reaction to various cell lines of medication with the drugs. This method is evaluated by computing its functioning in cross validation utilizing a full-fledged data set with 692 drugs and their identified side effects inherited from package inserts.

III. PROPOSED METHODOLOGY

The evolution of automated techniques utilize a computational approach making use of the drug data sets that are accessible openly for the prediction of side effects is investigated. This research work proposes the application of a machine learning technique to design the classifiers for side effect, utilizing the suitable collection of characteristics of data. This research work proposed a framework for discovering the side effects using the best feature selection and classification methods.

The proposed research work presents a new approach to measure the features (attributes) in drug prediction dataset using the Modified Orthogonal Opposition Learning based Cuckoo Search Algorithm (MOOLCSA). Finally in this work, Modified Shrinkage function based Extreme Gradient Boost (MSXGB) system functions as a classifier that stands as a supervised machine learning technique and for ADRs, the side effects can be predicted. This work, focus on the chemical formats of the drug, substituent’s for drug and drug target proteins which are derived from Drug Bank.

In the prediction of side effect, systems for prediction are built on the training drugs which are then directed over the testing drugs. This proposed intelligent technique is beneficial for both patients and medicinal scientists and

DOI: 10.37200/IJPR/V24I4/PR201596
Received: 17 Jan 2020 | Revised: 19 Feb 2020 | Accepted: 28 Feb 2020

possess better interpretability and predictability. The figure.1.shows the overall process of the Prediction of Drug Side Effects using machine learning techniques.

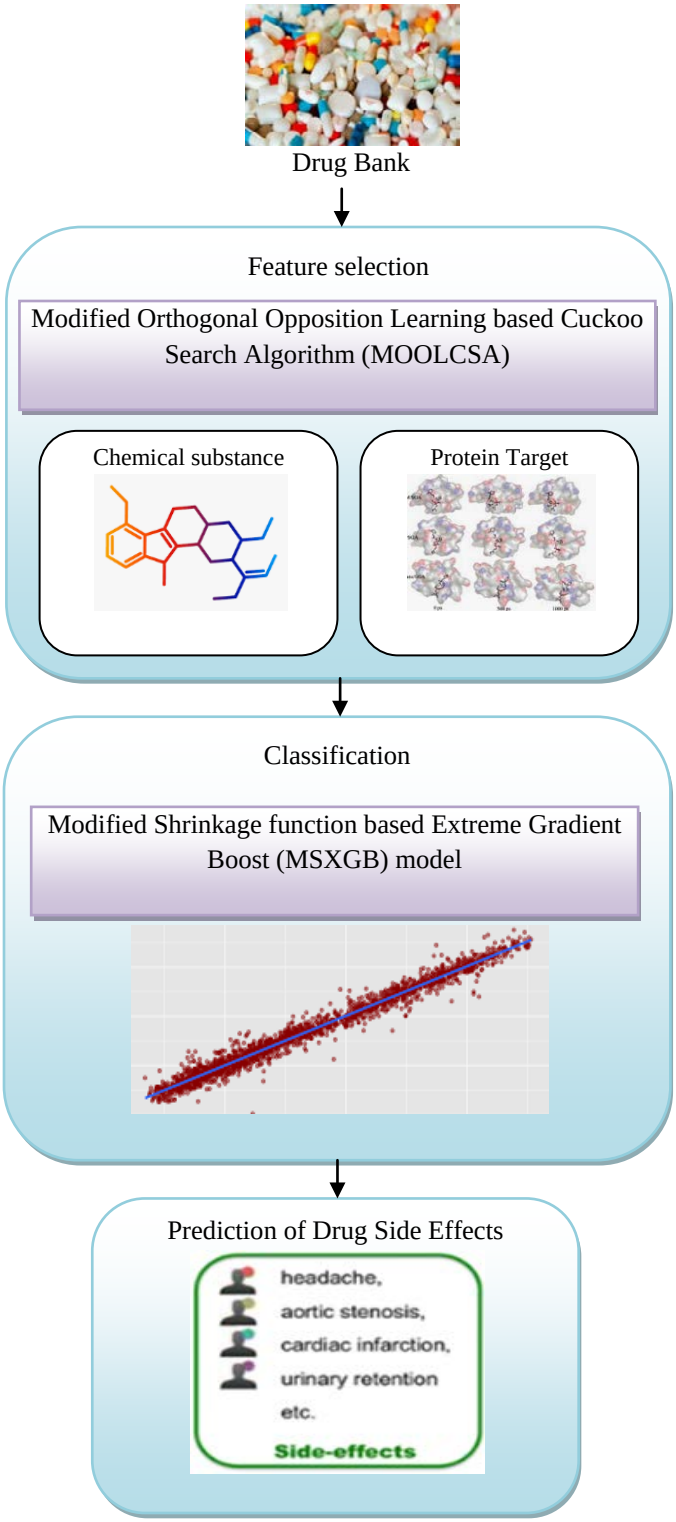

Figure 1: The Overall Process of Prediction of Drug Side Effects

DOI: 10.37200/IJPR/V24I4/PR201596
Received: 17 Jan 2020 | Revised: 19 Feb 2020 | Accepted: 28 Feb 2020

3.1. Data Representation

In this research work, the data on drugs was utilised to define drugs from the biological and chemical viewpoints, together with their chemical organizations and correlated proteins. The proteins were classified still more into a target, in accordance with their consequent operations [24]. That is, these data kinds focus on the characteristics utilised for defining the drugs. In spite of the pathways that might act as beneficial biological aspects, they were not incorporated in the binary vector illustration described for prediction, since the presence of them will include the usage of drug troubled profiles of gene expression, and thus restraining the appropriateness in large-scale of the prediction model. The Fig.2.illustrated as the depiction of drug data, where n1 to n2 are the figures of various kinds of attributes, and m being the total count of side effects taken in to consideration.

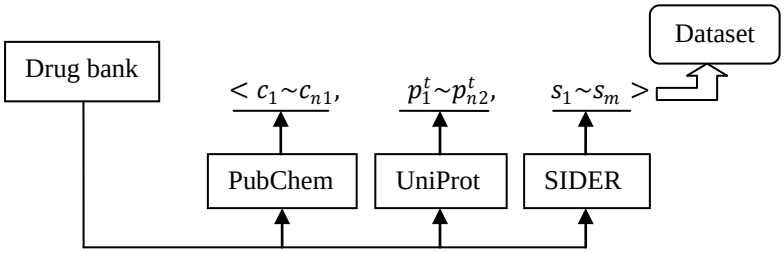

Figure 2: The Overall Data Representation of Drug Data

To retrieve chemical information regarding the substructures of the drugs, the well-known public database of drug information called Drug Bank is used to gather the data on FDA-approved small-molecule drugs and link them to PubChem. SIMES (Simplified Molecular Input Line Entry Specification) is applied to interpret data concerning the drugs' substructures. On the basis of the segment rules, the substructures were expressed and transformed to PF2 structure. In this manner, the chemical substructures described in PubChem can be encoded as binary features as mentioned here: In case the consequent PubChem substructure is existing in the drug, then the entry would be 1; Else the entry would be 0. In the data of retrieved protein for every drug that were gathered from Drug Bank. The proteins were linked to UniProt Knowledgebase that encompasses the highly inclusive and whole set of data on proteins. In analogous to the above defined illustration of binary feature applied for substructures of chemicals, proteins were translated as binary features for every drug to denote the existence or non-existence of the consequent proteins.

On top of systematizing the above defined data, derived information about side effect from the SIDER database that includes information on medicines available in market and their related drug reactions that are adverse [24]. SIDER applies STITCH compound identifiers in order to denote drugs that can be linked with PubChem compound identifiers to make sure the reliability with further drug pertinent information. All the side effects were considered as a binary target class to signify and predict independently its presence in connection with the data for every drug.

3.2. Feature Selection using Modified Orthogonal Opposition Learning based Cuckoo Search Algorithm (MOOLCSA)

In this section, a modified cuckoo search algorithm is presented in order to enhance the optimization competence of the original cuckoo search (CS) algorithm. Opposition learning and the orthogonal design are included in the CS

DOI: 10.37200/IJPR/V24I4/PR201596
Received: 17 Jan 2020 | Revised: 19 Feb 2020 | Accepted: 28 Feb 2020

algorithm to improve the capability of exploitation search. The opposition learning (OL) is infused into CS for population initialization and creating more candidate solutions in evolutionary generation that can direct the population in the path of the highly favorable domains and distribute it through the searching space as much as feasible.

3.2.1. Basic CS Algorithm

The fundamental CS algorithm is grounded on the brood parasitism of certain cuckoo type by laying their eggs in the nests of different host birds.

For ease in defining the fundamental CS, the succeeding three model guidelines are applied: (1) One egg is laid by every cuckoo at a time, and leaves it in set that is arbitrarily selected; (2) the finest nests having good-quality eggs would be transferred to the subsequent generations [26];(3) the existing host nests 'quantity is permanent, and the egg placed by a cuckoo is found by the host bird with a probability $p_a \in [0,1]$. Under this situation, the host bird can either throw away the egg or merely leave the nest and construct a complicated new nest.

This probability reflects the impact of substituting cuckoo eggs with different eggs at every generation. A solution is denoted by each egg. These suppositions guarantee that the superior solutions would persist from one generation to the subsequent generation as a procedure for selection for the optimization algorithm.

Consequently, the objective of the CS algorithm is to substitute solutions existing in the nests with fresh solutions which possess superior characteristic. A new solution X_i^{t+1}forcuckoo *i* is given by:

$$X_i^{t+1} = X_i^t + \alpha \otimes L'evy(\lambda) \quad (1)$$

$$\alpha = \alpha_0 \otimes (X_j^t - X_i^t) \quad (2)$$

where α is the step size (α > 0) with dimension equal to the dimension of the problem; the product ⊗ represents entry-wise multiplications; $X_j^t$is a solution randomly selected; and the *Lévy*(λ) is Lévy flights accidental walks. The attribute α0 is selected equals 0.01, as commended in, to improve the ability of search operation.

Lévy flights is one among the random walks in which its step distance is obtained from *Levy* distribution. This distribution is symbolized by a sequence of immediate leaps produced by a probability density function which possess a power law tail defined by:x

$$L'evy(\lambda) \approx S = t^{-\lambda}, (1 < \lambda \leq 3) (3)$$

The step length *S* of Lévy flights is obtained from a uniform distribution that *conforms Lévy* distribution. Moreover, the algorithm applied a stable blend of a local random walk and the global explorative random walk, governed by a switching attributep_a. The local random walk can be expressed as

$$X_i^{t+1} = X_i^t + \alpha s \otimes H(p_a - \varepsilon) \otimes (X_j^t - X_k^t) \ (4)$$

Where X_j^t and X_k^t are two various solutions chosen arbitrarily by random permutation, H is a function of Heaviside, ε is an arbitrary number chosen from a uniform distribution, and *s*is the size of the step.

Conversely, the global random walk is performed by applying L´evy flights:

$$X_i^{t+1} = X_i^t + \alpha \oplus L'evy(s, \lambda) \ (5)$$

DOI: 10.37200/IJPR/V24I4/PR201596
Received: 17 Jan 2020 | Revised: 19 Feb 2020 | Accepted: 28 Feb 2020

Here, $\alpha > 0$ is the scaling factor of step size; L´evy(s, λ) is the step-lengths which are disseminated as per the subsequent distribution of probability defined in (5) that possess an infinite variance with an infinite average:

$$L'evy(s,\lambda) = \frac{\lambda\Gamma(\lambda)\sin(\frac{\pi\lambda}{2})}{\pi}\frac{1}{s^{1+\lambda}} (6)$$

3.2.2. Modified Orthogonal Design and Opposition Learning based CSA

For enhancing the searching capability of the algorithm still more, the opposition learning function and orthogonal design are incorporated into the CS algorithm [26].The fundamental notion of the orthogonal structure is to use the features of the fractional experiment to effectively decide the superior grouping of levels. An orthogonal array of K factors with Q levels and M combinations is represented as $L_M(Q^K)$,where Q is the prime number,$M = Q^J$, and J is a positive integer fulfilling $K = (Q^J - 1)/(Q - 1)$.The concise process of building the orthogonal array $L_M(Q^K) = [a_{i,j}]_{M.K}$ is defined in the following process. The orthogonal design procedure is described in algorithm 1.

Procedure 1: Steps for Orthogonal Design

Step 1: Create the fundamental columns
For k=1 to J
j $= \frac{Q^{k-1}-1}{Q-1} + 1$
For i =1to Q^J
$a_{i,j} = \left|\frac{i-1}{Q^J-k}\right| \bmod Q$
End for
End for
Step 2: Create the non-fundamental columns
For k = 2 to J
j $= \frac{Q^{k-1}-1}{Q-1} + 1$
For s =1 to j-1
For t = 1 to Q-1
$a_{j+(s-1)(Q-1)+1} = (a_s \times t + a_j) mod\ Q$
End for
End for
Step 3: Increment $a_{i,j}$ by one for $1 \le i \le M, 1 \le j \le N$

Algorithm 1: Orthogonal Design

Begin
(1) Create the orthogonal array adhering the preceding steps
(2) Choose two solutions arbitrarily from the population
(3) Quantize the area created by the two solutions
(4) Arbitrarily produce $k - 1$ integers $p_1 \ldots p_{k-1}$
(5) Apply $L_M(Q^K)$to produce M possible offspring
(6) From the population, arbitrarily choose a solution
(7) Measure up the solution compared with the finest solution from the orthogonal M offspring and achieve the superior solution
End

DOI: 10.37200/IJPR/V24I4/PR201596
Received: 17 Jan 2020 | Revised: 19 Feb 2020 | Accepted: 28 Feb 2020

In computational intelligence, opposition Learning (OL) is a novel thought. The core opinion at the back of OL is to reflect both the solution and its subsequent conflicting solution so as to attain a finerguesstimate of the existing candidate solutions. It has been established to be an efficient approach to improve different techniques of optimization. Hence, the OL notion is fused into the algorithm proposed, so as to upsurge variety furthermore and accelerate the convergence.

In case, $X = (x_1, x_2, \ldots, x_n)$ is a solution in an n-dimensional space, where $x_i \epsilon [Lx_i, Ux_i], (i = 1,2, \ldots, n)$..The opposite solution $X' = (x'_1, x_2, \ldots, x'_n)$ is expressed by:

$$x'_i = Lx_i + Ux_i - x_i \quad (7)$$

Let f (.) be a function for fitness through which the value of fitness could be calculated. As per the previously mentioned description of X and X', if $f(X') \leq f(X)$, then X is substituted with X', or else X is retained. Thus the solution and its contradictory solution are assessed concurrently with the intention of obtaining the better one. OL is put into operation for initializing the population and generate new solutions in the course of evolution procedure as explained in algorithm 2.

Algorithm 2: Modified Orthogonal Opposition Learning based Cuckoo Search Algorithm (MOOLCSA)

Input: Drug dataset
Output: Optimized Features of given drug
(1) Initialize the parameter values for the algorithm, create the arbitrary initial values of vector and fix the iteration number t = 1.
(2) Calculate fitness values of every individual and decide the existing finest individual with the finest objective value. Examine whether the ending condition is achieved. In case it is attained, then the best solution is given as output; Else update the iteration number $t = t + 1$ and proceed the process of iteration till the ending condition is attained.
(3) Hold the best solution of the final iteration, and group of new solutions
$Xnew = \left[x_1^{(t+1)}, \ldots, x_i^{(t+1)}, \ldots x_K^{(t+1)}\right]$ *by* L´evy flight,
(4) Assess the fitness value $F_i^{(t+1)}$ of the new solution $x_i^{(t+1)}$, and relate $F_i^{(t+1)}$ with $F_i^{(t)}$that denotes the solution of the t^{th} iteration.
(5) A division (p_a) of poorer quality nests are discarded and fresh ones are constructed.
(6) Execute the processes of orthogonal design scheme.
(7) Retain the superior solution.
(8) Explore for a new solution applying Opposition Learning methodology using Eqs.
(9) Hold the fine nest with quality solution
(10) Grade the nests and identify the best one from the present nests
(11)Forward the present best nest to the subsequent generation
(12) Go to step (2)
End

3.3. Modified Shrinkage function based Extreme Gradient Boost (MSXGB) classifier for drug prediction

XGBoost (Extreme Gradient Boosting) is a machine learning approach for classification and regression hitches based on the Gradient Boosting Decision Tree (GBDT) [27].

In the regression tree, the inward nodes denote the values for a feature analysis and the leaf nodes with scores denotes a conclusion. Since XG Boost implements an additive learning method with a second-order approximation,

DOI: 10.37200/IJPR/V24I4/PR201596
Received: 17 Jan 2020 | Revised: 19 Feb 2020 | Accepted: 28 Feb 2020

the first-order derivative and second-order derivative of the loss functions with respect to the prediction are needed for fixing the approach.

To illustrate a clear mechanism, first derive the second-order approximation of additive tree boosting. In this m denotes the number of data and n for the number of features.

The 'raw prediction' before the sigmoid function will be denoted as z_i, and the probabilistic prediction will be $\hat{y}_i = \sigma(z_i)$, where $\sigma(.)$ is used to represent the sigmoid function. It is important to keep in mind that there is adiscrepancy between the notations and as the $\hat{y}_i$ in their analysis isdenoted as z here. y_i is used to denote the true label, $\alpha$and γ and are used for the parameters for the two loss functions, respectively [28].

The expressions of the gradients/hessians are noted in a merged format independent from the value of y_i,as this can simplify the program implementation and help vectorization in other related programs.

According to the additive learning objective used in practice is:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l\left(y_i, z_i{}^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \quad (8)$$

where t denotes the t-th iteration of the training process. Notice that the replacement of the notations has been applied in the equation. Applying second-order Taylor expansion on equation 8, one will get:

$$\mathcal{L}^{(t)} \approx \sum_{i=1}^{n}[l\left(y_i, z_i{}^{(t-1)} + g_i f_t(x_i)\right) + \frac{1}{2} h_i (f_t(x_i))^2 + \Omega(f_t) \quad (9)$$

$$\propto \sum_{i=1}^{n}[g_i f_t(x_i) + \frac{1}{2} h_i (f_t(x_i))^2] + \Omega(f_t) \quad (10)$$

The last line comes from the fact that the $l(y_i, z_i^{(t-1)})$ term can be removed from the learning objective as it is unrelatedto the fitting of the model in the t-th iteration.

Since XG Boost does not provide automatic differentiation, the hand-derived derivatives will be essential. Meanwhile, the derived expressions have further potentials to be applied into other machine learning tasks.

For both loss functions, sigmoid is selected as activation, and the following basic property of sigmoid will be consistently used in the derivatives:

$$\frac{\partial \hat{y}}{\partial z} = \frac{\partial \sigma(z)}{\partial z} \quad (11)$$

$$= \sigma(z)(1 - \sigma(z)) \quad (12)$$

$$= \hat{y}(1 - \hat{y}) \quad (13)$$

In addition with the regularized goal, the added approaches are utilised to avoid over fitting still more. Shrinkage measures the recently appended weights by a factor w_j^*following everyphase of tree boosting.

In analogous to a learning degree in optimization, contraction decreases the effect of every individual tree and permits space for forthcoming trees so as to enhance the model.

For a permanent structure q(x), calculate the optimal weight $w_j^*$of leaf j by

DOI: 10.37200/IJPR/V24I4/PR201596
Received: 17 Jan 2020 | Revised: 19 Feb 2020 | Accepted: 28 Feb 2020

$$w_j^* = -\frac{\sum_{i\in I_j} g_i}{\sum_{i\in I_j} h_i+\lambda}(14)$$

Here ∈ is afactor of approximation. Instinctively, this indicates that there are approximate selective points. Every data point is weighted by h_i that correspond to the weight, and Eq (10) could be rewritten as

$$\sum_{i=1}^{n} \frac{1}{2} h_i (f_t(x_i) - g_i/h_i)^2 + \Omega(f_t)(15)$$

Define $I_j = \{i|q(x_i) = j\}$ as the example group of leaf j. And it can be modified as Eq (17) by elaborating Ω as below

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} [g_i f_t(x_i) + \frac{1}{2} h_i {f_t}^2(x_i)] + \Omega(f_t)(16)$$

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} [g_i f_t(x_i) + \frac{1}{2} h_i {f_t}^2(x_i)] + \frac{1}{2}\lambda \sum_{j=1}^{T} {w_j}^2 \ (17)$$

which is precisely weighted squared loss having labels gi/hi and weights hi. It is significant to uncover candidate splits that assure the conditions for bigger datasets.

IV. RESULTS AND DISCUSSION

To analyze the implementation of the proposed approach for prediction of drug side effect, datasets were gathered from Drug Bank, having the chemical substructures outlined in PubChem and the subsequent drug side effects obtained from SIDER.

In the initial sequence of experiments, this dataset was utilised which encompassed the operations of wide-ranging try-outs to validate the proposed technique which is followed by an in-depth examination to analyse the proposed method from various viewpoints.

Prior to accomplishing the experiments, a short investigation is performed on the dataset which were utilised for the initial sequence of experiments. It comprises 1002 various drugs and 3903 side effects, and an aggregate of 7257 attributes were employed for the drug characterization.

4.1. Evaluation Metrics

In the experiments, various conditions which are often applied in binary classification are used to analyze the function of various approaches in the side effect prediction. Initial operation is to determine the true positive (TP), false positive (FP), true negative (TN), and false negative (FN) values which is then utilised to compute different performance measures. The foremost performance measure was precision which represents the quantity of recovered instances that were appropriate.

The next performance metric recall is defined as the quantity of appropriate instances that were recovered. In spite of the frequent contradictory nature, the precision and recall measures are both significant in determining the functioning of a prediction technique. And hence, these two measures can be merged with equal weights to acquire a single metric, termed the F-measure. The concluding performance metric accuracy is described as the amount of appropriately predicted examples comparative to all the instances that are predicted.

DOI: 10.37200/IJPR/V24I4/PR201596
Received: 17 Jan 2020 | Revised: 19 Feb 2020 | Accepted: 28 Feb 2020

Table 1: Performance Metrics Along with Different Methods

Metrics	RF classification	PCA+Similarity	MFVR-SVM	MSXGB
Accuracy	81.400	87.300	93.300	95.900
Precision	80.595	87.000	93.161	94.246
Recall	81.649	87.617	93.124	95.331
F-Measure	81.119	87.310	93.142	94.785

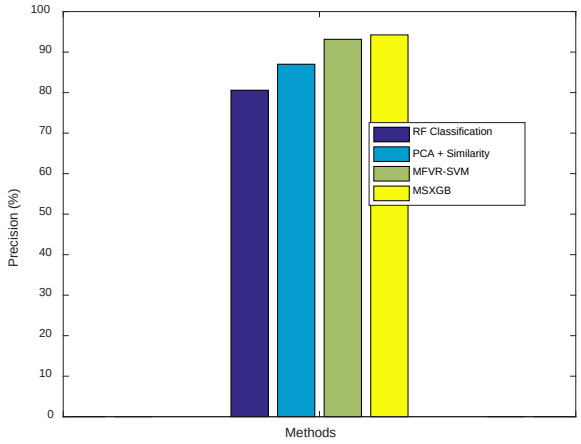

Figure 3: Precision Comparison Results between the Different Methods for Predicting the Drug Side Effects

The fig.3 illustrates that the precision comparison results between the different methods for predicting the drug side effects using the proposed Modified Shrinkage function based Extreme Gradient Boost (MSXGB) classification.

From the achieved outcomes it is concluded that the proposed technique has more accurate results compared to the existing predicting techniques.

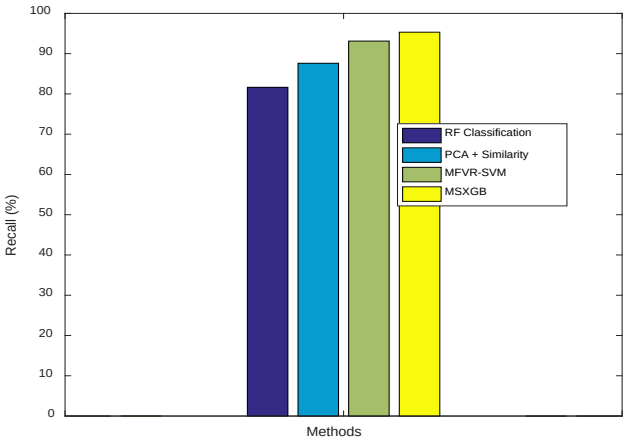

Figure 4: Recall Comparison Results between the Different Methods for Predicting the Drug Side Effects

DOI: 10.37200/IJPR/V24I4/PR201596
Received: 17 Jan 2020 | Revised: 19 Feb 2020 | Accepted: 28 Feb 2020

The fig.4 illustrates that the recall comparison results between the different methods for predicting the drug side effects using the Modified Shrinkage function based Extreme Gradient Boost (MSXGB).From the attained results it concludes that the proposed technique has high values of recall compared to the existing predicting techniques.

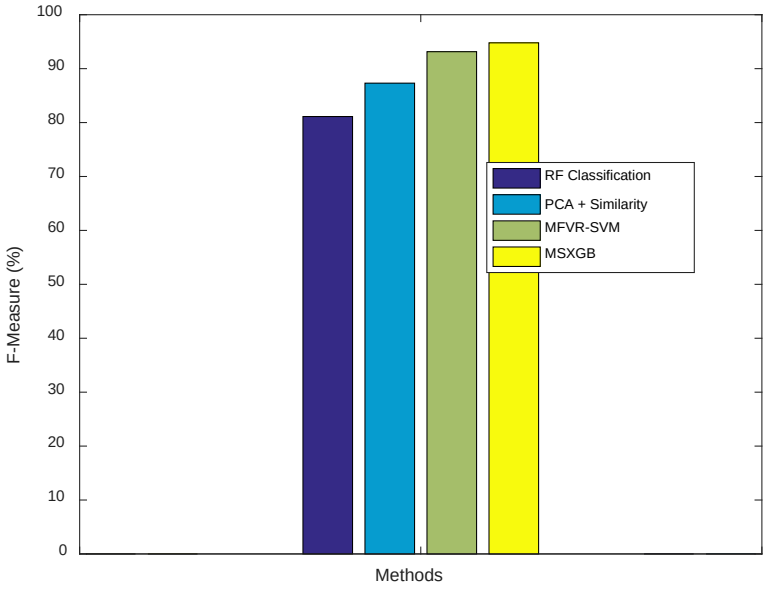

Figure 5: Comparison Results of F-Measure between the Various Methods for Predicting the Drug Side Effects

The fig.5 illustrates that the comparison results of F-measure between the various approaches for predicting the drug side effects using the Modified Shrinkage function based Extreme Gradient Boost (MSXGB). From the obtained results it concludes that the proposed technique has high F-measure results compared to the existing predicting techniques.

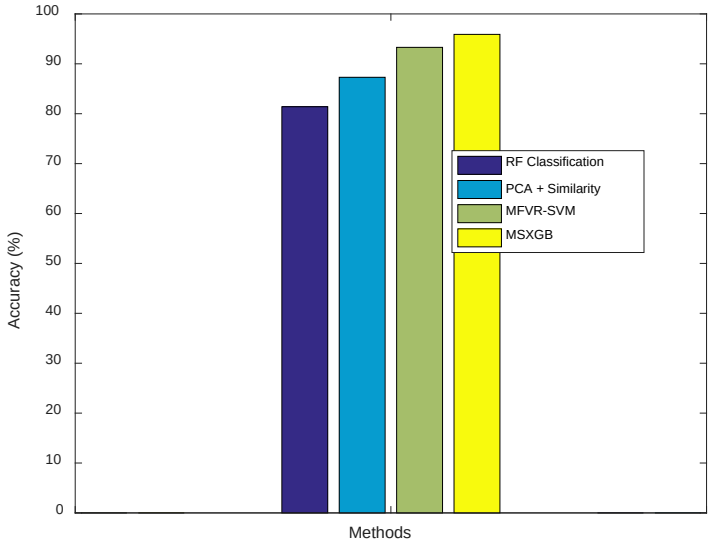

Figure 6: Accuracy Comparison Results between the Different Methods for Predicting the Drug Side Effects

DOI: 10.37200/IJPR/V24I4/PR201596
Received: 17 Jan 2020 | Revised: 19 Feb 2020 | Accepted: 28 Feb 2020

The fig.6 illustrates that the accuracy comparison results between the different methods for predicting the drug side effects using the Modified Shrinkage function based Extreme Gradient Boost (MSXGB). From the experimental results it concludes that the proposed technique has high precision results compared to the existing predicting techniques.

V. CONCLUSION

In the proposed work, the evolution of automated techniques utilizes a computational method employing drug data sets that are presented openly for the drug side effects prediction. This research work exhibits the application of a machine learning methodology to create classifiers of side effects making use of a suitable data features set. This approach utilizes the outlook of data analytics to explore the influence of drug distribution in the feature space, classify side effects into various spaces, embrace appropriate tactics for every interval, and consequently build the data models. The proposed research work investigate a new approach to measure the features (attributes) in drug prediction dataset using the Modified Orthogonal Opposition Learning based Cuckoo Search Algorithm (MOOLCSA). Finally in this work, Modified Extreme Gradient Boost (MXGB) was used as a classifier to predict the ADRs. The sequences of experiments were carried out to validate the appropriateness of the proposed approach in prediction of side effect. The experimental results reveal that this method was competent enough consider the features of various kinds of side effects and thus achieving improved prediction. Together with the prediction performance, this work is extended to investigate the interpretability of the model which is a further significant concern from the clinicians stand point.

REFERENCES

[1] Liu, M., Wu, Y., Chen, Y., Sun, J., Zhao, Z., Chen, X. W., ...&Xu, H. (2012). Large-scale prediction of adverse drug reactions using chemical, biological, and phenotypic properties of drugs. *Journal of the American Medical Informatics Association*, *19*(e1), e28-e35.

[2] Bresso, E., Grisoni, R., Marchetti, G., Karaboga, A. S., Souchet, M., Devignes, M. D., &Smaïl-Tabbone, M. (2013). Integrative relational machine-learning for understanding drug side-effect profiles. *BMC bioinformatics*, *14*(1), 207.

[3] Zhang, W., Liu, F., Luo, L., & Zhang, J. (2015). Predicting drug side effects by multi-label learning and ensemble learning. *BMC bioinformatics*, *16*(1), 365.

[4] Cheng, F., Liu, C., Jiang, J., Lu, W., Li, W., Liu, G., ...& Tang, Y. (2012). Prediction of drug-target interactions and drug repositioning via network-based inference. *PLoS computational biology*, *8*(5), e1002503.

[5] Wang, K., Sun, J., Zhou, S., Wan, C., Qin, S., Li, C., ...& Yang, L. (2013). Prediction of drug-target interactions for drug repositioning only based on genomic expression similarity. *PLoS computational biology*, *9*(11), e1003315.

[6] Montanari, F., &Ecker, G. F. (2015). Prediction of drug–ABC-transporter interaction—Recent advances and future challenges. *Advanced drug delivery reviews*, *86*, 17-26.

[7] Tatonetti, N. P., Liu, T., & Altman, R. B. (2009). Predicting drug side-effects by chemical systems biology. *Genome biology*, *10*(9), 238.

[8] Shaked, Itay, Matthew A. Oberhardt, NirAtias, RodedSharan, and EytanRuppin. "Metabolic network prediction of drug side effects." *Cell systems* 2, no. 3 (2016): 209-213.

[9] Cheng, T., Hao, M., Takeda, T., Bryant, S. H., & Wang, Y. (2017). Large-scale prediction of drug-target interaction: a data-centric review. *The AAPS journal*, *19*(5), 1264-1275.

[10] Bugrim, A., Nikolskaya, T., &Nikolsky, Y. (2004). Early prediction of drug metabolism and toxicity: systems biology approach and modeling. *Drug discovery today*, *9*(3), 127-135.

[11] Kuhn, M., Campillos, M., González, P., Jensen, L. J., & Bork, P. (2008). Large-scale prediction of drug–target relationships. *FEBS letters*, *582*(8), 1283-1290.

DOI: 10.37200/IJPR/V24I4/PR201596
Received: 17 Jan 2020 | Revised: 19 Feb 2020 | Accepted: 28 Feb 2020

[12] Tatonetti, N. P., Patrick, P. Y., Daneshjou, R., & Altman, R. B. (2012). Data-driven prediction of drug effects and interactions. *Science translational medicine*, *4*(125), 125ra31-125ra31.
[13] Yamanishi, Y., Pauwels, E., &Kotera, M. (2012). Drug side-effect prediction based on the integration of chemical and biological spaces. *Journal of chemical information and modeling*, *52*(12), 3284-3292.
[14] Zhang, P., Wang, F., Hu, J., &Sorrentino, R. (2015). Label propagation prediction of drug-drug interactions based on clinical side effects. *Scientific reports*, *5*, 12339.
[15] Gottlieb, A., Stein, G. Y., Ruppin, E., &Sharan, R. (2011). PREDICT: a method for inferring novel drug indications with application to personalized medicine. *Molecular systems biology*, *7*(1).
[16] Cheng, F., Li, W., Wu, Z., Wang, X., Zhang, C., Li, J., ...& Tang, Y. (2013). Prediction of polypharmacological profiles of drugs by the integration of chemical, side effect, and therapeutic space. *Journal of chemical information and modeling*, *53*(4), 753-762.
[17] Zhao, X. M., Iskar, M., Zeller, G., Kuhn, M., Van Noort, V., & Bork, P. (2011). Prediction of drug combinations by integrating molecular and pharmacological data. *PLoS computational biology*, *7*(12), e1002323.
[18] Shaked, I., Oberhardt, M. A., Atias, N., Sharan, R., &Ruppin, E. (2016). Metabolic network prediction of drug side effects. *Cell systems*, *2*(3), 209-213.
[19] Pauwels, E., Stoven, V., &Yamanishi, Y. (2011). Predicting drug side-effect profiles: a chemical fragment-based approach. *BMC bioinformatics*, *12*(1), 169.
[20] Mizutani, S., Pauwels, E., Stoven, V., Goto, S., &Yamanishi, Y. (2012). Relating drug–protein interaction network with drug side effects. *Bioinformatics*, *28*(18), i522-i528.
[21] Huang, L. C., Wu, X., & Chen, J. Y. (2013). Predicting adverse drug reaction profiles by integrating protein interaction networks with drug structures. *Proteomics*, *13*(2), 313-324.
[22] Takarabe, M., Kotera, M., Nishimura, Y., Goto, S., &Yamanishi, Y. (2012). Drug target prediction using adverse event report systems: a pharmacogenomic approach. *Bioinformatics*, *28*(18), i611-i618.
[23] Atias, N., &Sharan, R. (2011). An algorithmic framework for predicting side effects of drugs. *Journal of Computational Biology*, *18*(3), 207-218.
[24] Lee, W. P., Huang, J. Y., Chang, H. H., Lee, K. T., & Lai, C. T. (2017). Predicting drug side effects using data analytics and the integration of multiple data sources. *IEEE Access*, *5*, 20449-20462.
[25] Yang, X. S., & Deb, S. (2009, December). Cuckoo search via Lévy flights. In *2009 World Congress on Nature & Biologically Inspired Computing (NaBIC)* (pp. 210-214). IEEE.
[26] Yang, X. S., & Deb, S. (2013). Multiobjective cuckoo search for design optimization. *Computers & Operations Research*, *40*(6), 1616-1624.
[27] Bansal, A., &Kaur, S. (2018, April). Extreme Gradient Boosting Based Tuning for Classification in Intrusion Detection Systems. In *International Conference on Advances in Computing and Data Sciences* (pp. 372-380). Springer, Singapore.
[28] Shi, H., Wang, H., Huang, Y., Zhao, L., Qin, C., & Liu, C. (2019). A hierarchical method based on weighted extreme gradient boosting in ECG heartbeat classification. *Computer methods and programs in biomedicine*, *171*, 1-10.

DOI: 10.37200/IJPR/V24I4/PR201596
Received: 17 Jan 2020 | Revised: 19 Feb 2020 | Accepted: 28 Feb 2020