International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 6 Volume No.6 October - November – 2022
DOI:10.46647/ijetms.2022.v06i06.083 ISSN: 2581-4621

A Machine Learning Framework to Predict Adverse Drug Reactions from Electronic Health Records

Edwin Ponraj T[1], J. Charles[2]
[1] *Research Scholar - Computer Applications, Noorul Islam Centre for Higher Education, Kumaracoil, 629180, Tamilnadu*
[2] *Associate Professor - Software Engineering, Noorul Islam Centre for Higher Education, Kumaracoil, 629180, Tamilnadu*

ABSTRACT
The advent of medical industry is immense in the recent years, with increasing epidemics, pandemics and endemics, thereby ensuring proper treatment to the patients with affordable medications. Right from the chemical compositions, the manufacturing involves a huge procedural overhead along with the clinical trials and pre-marketing studies. Such medications have to clear stringent norms and policies in order to reach the market, during which, the medication should exhibit compliance to lesser side effects, and increased control over the diseases. Adverse Drug Reactions (ADR) are typically reactions which are unintended and may result in harmful effects over the patients. Multiple models understand the data collected from various sources to monitor the effects of medications in regular intervals, in order to prepare for counter actions. The risk of side effects and adverse drug reactions can be reduced with timely detection of drug to protein interaction and drug to drug interactions. It is a practice of co-prescriptions for addressing multiple medical conditions in elderly people and patients with multiple ailments. Compared to the previous manual techniques, prediction of adverse drug reactions was carried out using machine learning techniques lately. The proposed technique introduces a novel mechanism using regularized logistic regression technique to effectively trace the drug-to-drug interactions. The datasets are considered from openly available sources, and electronically stored information are fed into regression models for finding relevant patterns. Empirical studies applied with necessary cross validation checks and numerous failproof tests deliver promising outcomes in form of drug-ADR targeted profiles for signifying the results of the supposed study. From the investigative results, it is evident that the proposed technique ensures utmost quality and interesting insights for making appropriate biological and protein-drug based decisions.
Keywords— A Machine Learning, drug-drug interaction, ADR, logistic regression, cross validation.

1. Introduction

Every medication is associated with a proper response to the disease affected in a patient, and in case of some undesired events or unwanted phenotypic conditions, it results in adverse drug reactions. The alterations typically occur in the biological routes between the different organs and affect the interactions between drugs and proteins. According to the medical world, an adverse drug reaction is termed to be “an appreciably harmful or unpleasant reaction, resulting from an intervention related to the use of a medicinal product, which predicts hazard from future administration and warrants prevention or specific treatment, or alteration of the dosage regimen, or withdrawal of the product.” [1] The mortality or morbidity rate of adverse drug reactions are considerably high and hence the medical industry is taking necessary steps to mitigate the increasing effects of adverse drug reactions and events. For any disease, the number of medications consumed is relatively high, incrementing the chances of multiple side effects. The changes registered as side effects or undesirable reactions are also dependent on age, gender, and demographics [2]. Medications are considered to be entities which are supposed to work independently with relevant protein structures, thereby signifying the recovery of human diseases.

DOI:10.46647/ijetms.2022.v06i06.083 ISSN: 2581-4621

Medical experts determine the effects of ADRs by determining the drug risk level and consider it to be a valuable auxiliary judgement. Furthermore, the process of determination is considered to be vital in one application scenario where a pharmaceutical company evaluates a new medicine before it is released to the market and used as a remedial clinical practise for a length of time. The next significant purpose is to generate ADR data, gather the collective information and to establish the drug's risk level by estimating the severity of after-effects. Based on these summaries of information and effects, the medicine is re-evaluated to determine if the same can be prescribed to patients or to prohibit the circulation. The Food and Drug Administration (FDA) divides medications into two primary categories namely the drugs which are available through a prescription (Rx) [3] from pharmaceuticals and over-the-counter (OTC) drugs. According to the definition, Rx drugs are those that are recommended by a doctor. Since the Rx drugs are said to possess powerful pharmacological effects, the application technique and duration of prescription have unique restrictions and can be defined only by the experts in the specific field of expertise. Apart from being the powerful variant of the two categories, the Rx drugs can be very poisonous, highly reliant, or freshly marketed. On the other hand, OTC drugs need not be prescribed by an experienced doctor. The patient can acquire OTC drugs selectively without consulting a doctor to treat moderate short-term sickness and discomfort [4]. These drugs do not have any harmful effects in line with the required specification, they are reasonably safe to use. Another standard organisation namely the Chinese FDA (CFDA) divides OTC drugs into two categories for further clarity and named them as OTC-A drugs and OTC-B drugs. OTC-B medications have more risk controllability than OTC-A medications.

Depending on the dosage and type of medication, the changes to respective statuses of medications from Rx to OTC-A or OTC-B tend to improve the quality of treatments. An interesting point to be noted is by eliminating prescription requirements saves time for both healthcare providers, medical institutions and patients. The entire process is being tracked for monitoring the signal detection, based on which, ADR reports are often collected from spontaneously updated databases. The intention of spontaneity is to prevent the harmful medications from further prescriptions in case of ADR or events. The spontaneous reporting system is an approach of pharmacovigilance tools and practices used by various nations to detect and report ADR signals. In order to facilitate the universal database, doctors, patients, and medical institutions voluntarily submit to drug regulatory organisations via the network. The original databases maintained internally, certainly lack a clear probability structure due to human error, improper translations, redundancy and various other factors, making it impossible to determine incidence rates or severity rates [5]. The reports also lack the detail on how severe the ADRs are and hence the report becomes incomplete. However, the target drug risk level is influenced not only by the frequency of ADR reports, but also by the ADR severity. ADRs from a high-risk medicine are uncommon, yet they can cost the lives of patients if left unnoticed or unattended. Simply looking at the frequency of ADRs would ignore the important information included in the ADR reports database. As a result, many signals identification approaches, such as Proportional Reporting Ratio (PRR), Reporting Odds Ratio (ROR), Medicines and Healthcare products Regulatory Agency (MHRA), and Information Component (IC), have been developed to overcome this problem. These estimates, often known as disproportionality studies, are not the same [6][7][8]. The PRR, MHRA, and ROR are non-Bayesian frequentist models, but the IC is Bayesian. We may build a multi-classification model in supervised machine learning using Rx Drugs, OTC-A Drugs, and OTC-B Drugs as labels and signal detection values of drug-ADR pairings as the dataset.

From the investigations carried out in this research, the severity of drugs are uneven and cannot be bifurcated correctly with incomplete datasets. The spontaneous datasets are thereby required for deriving the metrics for standard classification [9]. Traditional categorization procedures are dominated by the dominant class and ignore the minority class. However, all of the OTC medications were mistakenly categorised in the majority Rx Drugs class, demonstrating that the accuracy rate is insufficient for unbalanced categorization. The goal of this project is to use Synthetic Minority Oversampling Technique (SMOTE) and machine learning to predict drug risk level from ADRs in order to investigate the mechanism of ADRs determining drug risk level. We present a framework

DOI:10.46647/ijetms.2022.v06i06.083 ISSN: 2581-4621

for a spontaneous reporting database that incorporates signal detection, the enhanced SMOTE algorithm, and the multi-classifier [10]. Simultaneously, we provide the optimum structure combination and evaluation metrics that perform well on unbalanced multi-classification.

2. Literature Survey

From the survey of articles, it has been identified that the prominent approaches to detect ADR signals are based on non-Bayesian and Bayesian approaches. Frequentist approaches are defined by the techniques which implement PRR, ROR, and MHRA, and on the contrary, Bayesian methods are represented by IC. The advantages and negatives of various techniques were classified in a survey that broadly acknowledged future investigations in the study. Multiple techniques have incorporated PRR since the technique was highly explanatory and extensively applicable [11]. The only drawback of the said approach was its standard error which cannot be estimated during every iteration. Comparatively, ROR is simple to use and can be applied with a simple logistic regression model, but its drawbacks are obvious. The common drawbacks were the difficultly to comprehend, unreliability to detect the lesser number of events or reactions and impossible to compute if the denominator is zero. IC, when compared to the other approaches, are considered to be always relevant, appropriate for enormous amounts of computations, and used for high-dimensional pattern recognition. There is no chance of defining the universal norms that can be used to predict or represent every scenario. Given that these two approaches are standard, we use the same parameters as the representative PRR and IC for comparison targets in the proposed prediction framework to investigate the optimum signal detection strategy [12]. The research on predicting medication labels based on ADR signal detection is still open and need various improvements with a huge scope for betterments. ADR signals from various datasets were manually matched to medication label modifications that offered the stages of ADRs.

For any investigation, the first step is to decide an enhancement technique for detecting the signals better than conventional and manual approaches. Statistical foundations were required for these classic quantitative approaches available for detecting the signals that indicate the adverse drug reactions. The approach has offered examples of how they might be used to uncover novel medication interactions [13]. The next technique investigated the process of automating the collection of information from social media and use them for monitoring. The approach may offer warnings for dangerous and unusual incidents earlier than traditional signal detection from spontaneous reporting databases. The technique has corroborated, to some extent, the faster pace of social media. A new prediction model-based strategy was introduced for improving the efficiency of entire database screening with EHRs, clinical reporting systems and spontaneous reporting systems. In another study, it is observed that the AUC increased from 0.649 to 0.740 when ROR approach is applied for the comparisons of adverse drug reaction signals. The model concluded that the fraction of possible ADRs and their signals increased from 12.3% to 19.4%. In a novel framework suggested recognised ADR signals from multiple data sources using two steps namely Monte Carlo Expectation Maximization and signal combining [14].

The entire concept of detecting the ADR and categorization based on their symptoms has grown in popularity in recent years. A systematic classification of ADRs in 2016 included the type A (known, high morbidity and low mortality), type B idiosyncratic (low morbidity and high mortality), type C (continuous), type D (delayed usage), and type E (termination of use). In another investigation, a deep neural network model was introduced with the natural language processing domain that incorporated several ADR corpora to address entity-level ADR classification problems [15]. A neural network that blended sentiment analysis with transfer learning approaches to enhance ADR categorization in social media postings was tested for monitoring the symptoms of ADRs.

Finally, a model for predicting the drug-drug and drug-protein interaction was applied based on ADRs in a popular study topic. Since the concept was focused on humongous volumes of data, a novel data-driven technique for checking the different combinations of medications was considered for prediction of adverse drug reactions using data from a spontaneous reporting system in 2016. The

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 6 Volume No.6 October - November – 2022
DOI:10.46647/ijetms.2022.v06i06.083 ISSN: 2581-4621

model exhibited an intriguing connection in their research, in which one medicine might lessen the ADRs of the other. The application of machine learning to address classification issues are widely employed in the domain of medical industries [16]. The common problem when numerous data is handled and data imbalance has to be addressed before any prediction model is tested with real life datasets. This situation is also a prevalent phenomenon in classification. The imbalance imposes a number of challenges for classifiers and classification assessment metrics. The standard technique implied to handle the missing information and imbalances is known as SMOTE's suggestion. The approach dramatically increases categorization accuracy. SMOTE [17] and classifiers have been widely employed in the medical area in recent years. A model employed a Random Forest classifier with SMOTE and feature reduction approaches to aid cervical cancer detection in 2017, while another technique applied SMOTE-Random Forest to build a prediction tool called "WarfarinSeer" in 2018.

3. Proposed Methodology

The dataset is a composition of nearly 994665 adverse drug reaction reports collected from various spontaneous reporting systems, especially from FAERS, SIDER and FDA. Otherwise known as a tuple, an ADR report had the following fields in the spontaneous reporting database. The commonly included parameters are report ID, report address from the report generation framework, patient details such as age, gender of the patient, medication name, and ADR name. These values are consistently maintained for all the patients and numerous reports constitute the datasets [18]. From the reports, a pattern can be identified in form of a relationship and ADR reports were classified as one-to-one relationships. The association between the medication and the ADR are identified for measuring the severity and frequency of ADRs. As a result, drug-ADR pairs with their corresponding frequencies have been determined [19]. Then the Frequency DATA dataset is collected for monitoring the frequency of occurrences of ADRs, which contains 3262 medications and 3163 ADRs. The dataset has to be utilized in modelling and prediction only after standardising the medication and drug names, ADRs and removing drug-ADR pairings with lower frequencies [20]. After finalizing all the details, the final Normalized Frequency DATA dataset is obtained including 1047 medicines and 751 adverse drug reactions. The proposed architecture is illustrated in Figure 1.

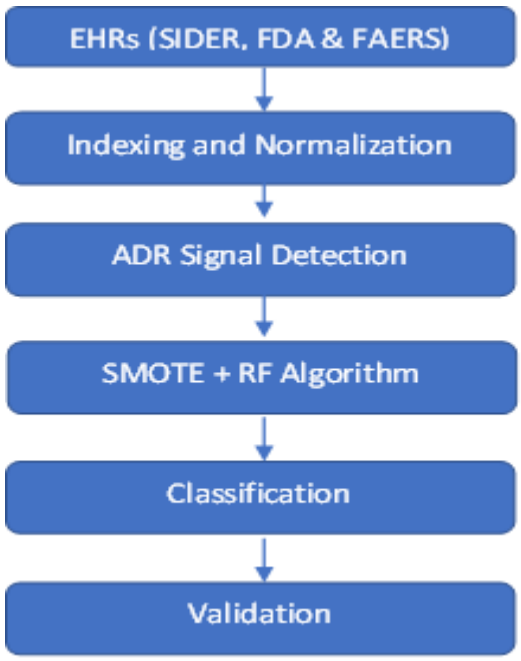

Figure 1: Proposed Architecture

The target medication risk level can be derived and established by determining the absolute number of ADR reports, along with the measurement of severity of ADRs. The severity of medications can be identified by measuring the intensity of side effects, either with an advice of stopping the

DOI:10.46647/ijetms.2022.v06i06.083 ISSN: 2581-4621

medication or dosage or switching to another medication. However, major ADRs are not listed owing to the fact that they are uncommon in the database. In other words, while a high-risk medicine may showcase fewer ADRs [21], but the effect of the side effects shall be too harsh as they may be severe or even fatal. From these input levels, it is observed that both the parameters exhibit the drug danger level and the impact on human bodies, we established two signal detection approaches. The first approach is based on erroneous reporting, known as frequentist and the second approach is based on Bayesian, respectively. Following that, Feature selection is a vital element in assuring the accuracy where the signal detection findings were used as feature values throughout the classification procedure.

Feature scaling is an approach used to speed up the process of finding the best features and highlight them for careful classification results. Better the feature selection, better the accuracy of a classification model when dealing with a large dataset. On the other hand, the Proportional Reporting Ratio or Chi-Square values [22] ensure that the reports or symptoms are reported on the databases without any miss. The various ADR characteristics have a wide range among all general characteristics of adverse drug reactions or events. In order to improve the efficiency, we combined the MaxAbsScaler with our approaches to reduce the impact of incomplete or sparse datasets. The general problems produced by sparseness without influencing the sparse dataset are huge and hence need to be addressed before the dataset is considered as complete. The Proportional Reporting Ratio (PRR) [23] is a renowned approach for detecting ADR signals. Evans was the first to use it in prediction modelling for monitoring and reporting adverse drug reactions. PRR values obtained are defined as the indicators of association strength between drugs and their drug reactions that behave similarly to relative risks. In other words, the higher the PRR, the stronger the drug-ADR pair signal. The purpose of the proposed framework is to predict drug risk level from the obtained ADRs by using SMOTE and various machine learning methodologies. SMOTE is applied to investigate the mechanism of ADRs depending on the affecting drug risk level. There are four primary processes in the proposed technique, commencing from pre-processing, classification, validation, and application [24]. More specifically, SMOTE algorithm has a significant impact on the classification results, and it is evident from the investigative results where the signal detection methods (PRR and IC) are better suited to our model structure. From the results, when all the four parameters are compared, the pros and cons of the four classification methods (RF, GB, LR, and ADA) combined with SMOTE, were tested for evaluations and hence derive two classification evaluation curves to use for classification purposes. The processes following the pre-processing include the modelling of Random Forests with SMOTE applied with bagging classification technique. The process is continuously applied with logistic regression for validating the information from balanced and imbalanced information. Once the data is validated from the class imbalance removal techniques, the cross-validation technique is implied for ensuring the accuracy of classification. Tenfold cross validation is used for testing and training the model. According to the designed model, the maximum depth of the random forest is considered to be 150, and maximum iteration is applied to be 200. Decision trees are utilized as base estimators, and grid search models was applied for training the model.

4. Results and Discussions

L-regulated logistic regression is best known for training and testing the model with higher datasets, and when the classes are imbalanced. The datasets are known for their excessive noise and overtraining issues and the proposed technique has implemented the logistic regression as the base training parameter. Given that the decision tree is used for classification, Receiver Operating Characteristic Area Under the Curve is considered as the performance indicating parameter along with sensitivity, precision and correlation coefficient. The drug to drug interaction is also measured in the proposed model where the targeted protein and gene is mapped to its respective association. When the association between two genes or proteins are higher, the relationship indicates that the interaction intensity is extremely higher. When two medications concentrate on the same protein, the connectivity or association between the drugs are signified using their relationship on the graph.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 6 Volume No.6 October - November – 2022
DOI:10.46647/ijetms.2022.v06i06.083 ISSN: 2581-4621

When the distance between two drugs is longer, the model deems that the association is lesser between the drugs or the proteins.

SMOTE is applied in the proposed technique assuming that the information is imbalanced and hence the input datasets have to be cleaned and balanced. The given dataset contains nearly 2668 drugs which are divided into training and testing data based on a 80:20 ratio. The model is applied with SMOTE for removing the class imbalances and PRR used for highlighting the adverse drug reaction signals. According to Figure 2 and 3, logistic regression with PRR has the higher levels of classification accuracy owing to SMOTE and associating the drugs and adverse drug reaction signals. The macro and micro curve for the proposed classifier represents the highlighted difference between the evaluation metrics. The laws of curve changing and AUC values also signified the difference between the sub graphs and the values converge at (0.98, 0.98) respectively.

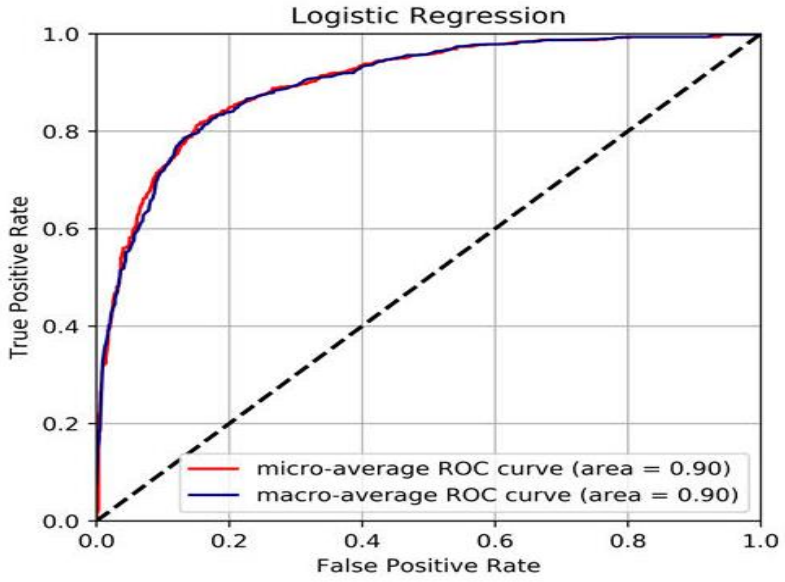

Figure 2: AUC Curve using Logistic Regression

Irrespective of SMOTE being implemented or not, the accuracy rates of the proposed technique has exhibited better accuracy which is comfortably around 87%. The consistency of the accuracy is promising after SMOTE being implemented. Classification of imbalanced dataset is more challenging than that of models without SMOTE. The classification will be based on the majority of features and their classes will be decided based on major similarities. The classification has no guarantee of being the right one and hence the imbalance is returned to the datasets to ensure the right bifurcation. The following table lists out the metrics of random forest during the application stages.

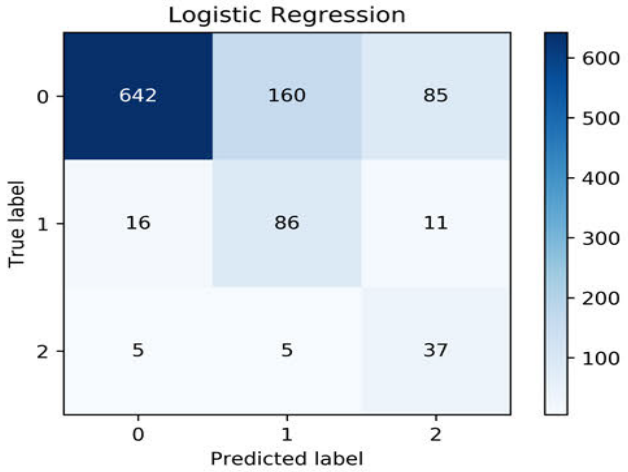

Figure 3: Confusion Matrix using Logistic Regression

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 6 Volume No.6 October - November – 2022
DOI:10.46647/ijetms.2022.v06i06.083 ISSN: 2581-4621

According to the Table 1, the signal detection mechanism plays a vital role in determining the classification accuracy and before SMOTE is applied, the technique suffers poor classification accuracy. Comparatively, PRR has better accuracy over the other techniques, when compared to IC, upon which the classifiers are applied. The classification of logistic regression technique is better in terms of PRR and lesser when IC is processed. The reason behind the poor classification when IC processing is carried out is since the scalability of information across the huge dataset. PRR processing is technically the best approach to address large scale information and the level of interpretability is higher. The overall presentation and classification accuracy has improved since the testing is performed at the application and validation stages

Table 1: Evaluation Results with Random Forest and SMOTE

Parameter		**Random Forest with SMOTE**
Precision	Label = 0	0.95
	Label = 1	0.98
	Label = 2	0.82
Recall	Label = 0	0.96
	Label = 1	0.96
	Label = 2	0.94
F1	Label = 0	0.87
	Label = 1	0.89
	Label = 2	0.84

CONCLUSION

The intention of the proposed technique is to derive the risk levels of adverse drug reactions and how the same affects the human bodies. The technique has incorporated SMOTE with logistic regression after the class imbalance is removed in a random forest model with decision trees as base classifier. The drug association with other drugs and proteins are monitored and used to predict the risk levels. With the right risk levels derived through the proposed model, the approach assists in proper decision making to further support the prescription of the same medication. Classification accuracy was measured after SMOTE and Random Forest has been derived as the prediction model and investigative results have shown promising results. The approach is a valid model for efficient classification and to assist the bifurcation of harmful drugs and adverse drug reactions from FDA, SIDER and FAERS datasets. The future directions are to derive more deep learning and machine learning models for better classification accuracy and to reduce the complexity of computations.

References

[1]Steyn, S. J. & Varma, M. V. S. Cytochrome-P450-mediated drug–drug interactions of substrate drugs: Assessing clinical risk based on molecular properties and an extended clearance classification system. Mol. Pharm. 17(8), 3024–3032 (2020).
[2]C. E. Pierce, K. Bouri, C. Pamer, S. Proestel, H. W. Rodriguez, H. Van Le, C. C. Freifeld, J. S. Brownstein, M.Walderhaug, I. R. Edwards, and N. Dasgupta, ``Evaluation of facebook and Twitter monitoring to detect safety signals for medical products: An analysis of recent FDA safety alerts," Drug Saf., vol. 40, no. 4, pp. 317-331, Apr. 2017.
[3]Watanabe, J.H., McInnis, T., and Hirsch, J.D. (2018). Cost of prescription drug-related morbidity and mortality. Ann. Pharmacother.
[4]Deodhar, M. et al. Mechanisms of CYP450 inhibition: Understanding drug–drug interactions due to mechanism-based inhibition in clinical practice. Pharmaceutics 12(9), 846 (2020).

DOI:10.46647/ijetms.2022.v06i06.083 ISSN: 2581-4621

[5]C. Xiao, Y. Li, I. M. Baytas, J. Zhou, and F.Wang, ``An MCEM framework for drug safety signal detection and combination from heterogeneous real world evidence," Sci. Rep., vol. 8, no. 1, pp. 1-10, Dec. 2018.
[6]Tannenbaum, C., and Day, D. (2017). Age and sex in drug development and testing for adults. Pharmacol. Res. 121, 83–93.
[7]Medina-Franco, J. L. et al. Rationality over fashion and hype in drug design [version 1; peer review: 2 approved]. F1000Research 10(Chem Inf Sci), 397 (2021).
[8]I. Alimova and E. Tutubalina, ``Entity-level classification of adverse drug reactions: A comparison of neural network models," in Proc. Workshop Widening NLP, 2019, pp. 132-134.
[9]Watson, S., Caster, O., Rochon, P.A., and Ruijter, H.d. (2019). Reported adverse drug reactions in women and men: aggregated evidence from globally collected individual case reports during half a century. EClinicalMedicine 17, https://doi.org/10.1016/j.eclinm.2019.10.001.
[10] Ferdousi, R., Safdari, R. & Omidi, Y. Computational prediction of drug–drug interactions based on drugs functional similarities. J. Biomed. Inform. 70, 54–64 (2017).
[11] H. Alhuzali and S. Ananiadou, ``Improving classification of adverse drug reactions through using sentiment analysis and transfer learning," in Proc. 18th BioNLP Workshop Shared Task, 2019, pp. 339-347.
[12] Zhou, L., and Rupa, A.P. (2018). Categorization and association analysis of risk factors for adverse drug events. Eur. J. Clin. Pharmacol. 74, 389–404.
[13] Zhang, W., Chen, Y., Li, D. & Yue, X. Manifold regularized matrix factorization for drug–drug interaction prediction. J. Biomed. Inform. 88, 90–97 (2018).
[14] S. F. Abdoh, M. Abo Rizka, and F. A. Maghraby, ``Cervical cancer diagnosis using random forest classier with SMOTE and feature reduction techniques," IEEE Access, vol. 6, pp. 5947559485, 2018, doi:10.1109/ACCESS.2018.2874063.
[15] Y. Tao and Y. Zhang, ```WarfarinSeer': A predictive tool based on SMOTE-random forest to improve warfarin dose prediction in Chinese patients," in Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM), Dec. 2018, pp. 1022-1026.
[16] Paavola, A. (2019). 10 most prescribed drugs in the US in Q1. https://www.beckershospitalreview.com/pharmacy/10-most-prescribed-drugs-in-theu-s-in-q1.html.
[17] Shtar, G., Rokach, L. & Shapira, B. Detecting drug–drug interactions using artificial neural networks and classic graph similarity measures. PLoS ONE 14, e0219796 (2019).
[18] X. Lin, X. Zhang, and X. Xu, ``Efficient classification of hot spots and hub protein interfaces by recursive feature elimination and gradient boosting," IEEE/ACM Trans. Comput. Biol. Bioinf., early access, Jul. 30, 2019, doi: 10.1109/TCBB.2019.2931717.
[19] Razzaque, M.S. (2018). Can adverse effects of excessive vitamin D supplementation occur without developing hypervitaminosis D? J. Steroid Biochem. Mol. Biol. 180, 81–86.
[20] Dhami, D. S., Kunapuli, G., Das, M., Page, D. & Natarajan, S. Drug–drug interaction discovery: Kernel learning from heterogeneous similarities. Smart Health (Amst.) 9–10, 88–100 (2018).
[21] S. M. H. Mahmud, W. Chen, H. Jahan, Y. Liu, N. I. Sujan, and S. Ahmed, ``IDTi-CSsmoteB: Identification of drug-target interaction based on drug chemical structure and protein sequence using XGBoost with over-sampling technique SMOTE," IEEE Access, vol. 7, pp. 48699-48714, 2019.
[22] Schatzberg, A.F., and Nemeroff, C.B. (2017). The American Psychiatric Association Publishing Textbook of Psychopharmacology (American Psychiatric Association Publishing), OCLC: 971615789, ISBN: 978-1-61537-122-8 978-1-61537-162-4.
[23] Lee, G., Park, C. & Ahn, J. Novel deep learning model for more accurate prediction of drug–drug interaction effects. BMC Bioinform.20, 415 (2019).
[24] Fabregat, A. et al. The Reactome pathway knowledgebase. Nucleic Acids Res. 46(Database issue), D649–D655 (2018).