International Journal of Research in Engineering, Science and Management
Volume-3, Issue-6, June-2020
www.ijresm.com | ISSN (Online): 2581-5792

Adverse Drug Reaction Prediction Using Machine Learning

Nebisathul Irfana[1*], Sowmya[2], Pooja Subray Tandel[3]
[1,3]*Student, Dept. of Information Science and Engineering, Srinivas Institute of Technology, Mangalore, India*
[2]*Assistant Professor, Department of Information Science and Engineering, Srinivas Institute of Technology, Mangalore, India*
*Corresponding author: irfanasnidpalli@gmail.com

***Abstract*: Pharmacogenomics is progressively being utilized to boost Drug Safety Profiles. It's public knowledge that genetic influences area unit typically vital in predicting however patients can reply to medicine. As a result, many corporation's area units currently commences to integrate the genetic element into drug discovery and development to boost safety profiles. Lexi-Interact severity, that is predicated on clinical trials and literature reviews, doesn't embrace all drug interactions although their area unit several prescribed drug mixtures that haven't been erstwhile coated by literature. This paper is enforced by training a model that has comparatively high accuracy and recall with best-known Lexi-Interact severity values, the goal would be to check it on drug interactions with best-known severity. Specifically, a drug would have a foreseen severity and so a panel of clinical pharmacists, people conversant in clinical outcomes of drug interactions, would rate the validity of that foreseen severity. The proposed system has a tendency to realize such reactive medicine and report them to the doctors.**

***Keywords*: Decision Tree, ADR, Lexi-interact.**

1. Introduction

Discovering unknown ADRs as early as doable is very fascinating as a result of they have an effect on sizable amount of individuals and might facilitate in raising early warning against adverse effects of medication and facilitate medical examiners in creating treatment effective and timely. In today's digital era a large quantity of knowledge correlative to adverse effects of medication is being collected at hospitals, drug retail stores and by drug producers. All medication has tiny low probability of aspect effects, taking completely different medications combined inevitably will increase the risk of adverse drug reactions (ADRs) and introduces the extra danger of interactions between medicines. Medical aspect reactions have a key role to play once it involves prescribing them. Unknowingness of the chemical reactions, such medication could cause aspect effects which might be severe. Thus it's necessary to possess information on the adverse effects of a drug interaction. We tend to will use Machine Learning to accumulate and analyze numerous information to forestall such health hazards. We tend to propose a machine learning methodology to predict ADRs of combined medication from medical specialty databases. To effectively apply machine learning techniques to the current prediction drawback, formulate it into a binary classification task, wherever inputs area unit vectors of drug pairs and labels area unit ADRs.

2. Related Work

Recently, numerous network-based strategies are projected. Network has become an efficient tool to predict underlying drug–target associations rule et al. developed a procedure formula to infer potential drug targets by consistently analyzing the transformation between the illness state and also the desired state during an illness network. The aim of MTOI is to seek out multiple target best intervention (MTOI) solutions that provide the simplest illness state transformation. Therefore, the output of MTOI includes not solely many potential drug–target interactions, however conjointly best combinatorial intervention solutions. The strategy was applied to Associate in Nursing inflammation-related network—the arachidonic acid (AA) metabolic network (AAnetwork) for the identification of best multi-target anti-inflammatory drug intervention solutions.

Network-based and machine learning-based models have their blessings and drawbacks. The key advantage of most of those approaches is that they're applicable to compounds while not acknowledged 3D structures. What is more, most the models will effectively predict novel drug–target interactions for medicine, that have a minimum of one acknowledged associated target proteins. a lot of significantly, some models may be any applied to the new compounds with none acknowledged associated target proteins by group action drug similarity, target similarity and acknowledged drug–target interactions.

As for the supervised machine learning methodology, it's vital limitations as follows. Firstly, there are not any experimental valid non-drug–target interactions so it's tough to pick negative samples. Most of supervised learning methodology regard the unknown drug–target interactions or every which way choose the unconfirmed drug–target pairs as negative samples, which might for the most part influence the prognosticative accuracy of the strategy. Secondly, 2 totally different classifiers from drug and target area unit created in some strategies, like BLM; thence, the ultimate result's the

typical of those 2 predictions, which can lead to biases. Though semi-supervised learning methodology NetLapRLS makes use of the unlabeled info and overcome the problem of choosing negative samples, it conjointly has an equivalent limitation of classifier combination [1]. Later piece of paper [2] during this Paper, they need developed machine learning models as well as a deep learning framework which might at the same time predict ADRs and establish the molecular substructures related to those ADRs while not shaping the substructures a-priori. They evaluated the performance of their model with 10 totally different progressive fingerprint models and located that neural fingerprints from the deep learning model outperformed all different strategies in predicting ADRs. Via feature analysis on drug structures, they known vital molecular substructures that area unit related to specific ADRs and assessed their associations via applied mathematics analysis.

The deep learning model with feature analysis, substructure identification, and applied mathematics assessment provides a promising answer for characteristic risky parts among molecular structures and might probably facilitate to enhance drug safety analysis. They harvested drug-ADR associations from the SIDER information, and generated 10 differing kinds of chemical fingerprints from molecular structures. They developed L2 norm regular provision regression models for all fingerprints to predict ADRs, and conjointly leveraged convolutional deep learning framework to integrate neural fingerprint generation and model development. They evaluated the performance of all eleven models and located that the neural fingerprints achieved the simplest overall performance. Supported the outputs from the neural fingerprints, they extracted the chemical substructures of the medicine which may be related to specific ADRs, evaluated their associations exploitation applied mathematics analysis and located proof in 2 case studies. The projected structurebased models can't solely acquire smart performance in ADR prediction; however conjointly establish the potential connections between substructures and ADRs. This study provides a helpful work flow for drug developers to spot risky substructures and will probably facilitate to enhance the security analysis of pipeline drugs [2].

3. System Implementation

The proposed system is designed with structural model with python programming and Django interface being embedded. The libraries being used are Numpy, Pandas, SKlearn, Seaborn, Matplotlib. The first stage is data collection for training. The process of gathering collecting the data depends on the kind of project we want to create. In the proposed system we used a custom dataset with 30,000 datasets that has the different severity levels mentioned for various mixtures of fifteen medical medications in several dose. The dose worth ranges from zero to three. The severity is classified based on these three completely different levels low, medium, and high. The second step is data processing. Data pre-processing helps in building machine learning models a lot of accurately. This is the step which cleaning the raw data i.e. the data is collected in the real world and is converted to a clean data set.

Therefore, certain steps are executed to convert the data into a small clean data set. In each information someone ought to pay eightieth time for information pre-processing and two hundredth time to truly perform the analysis. Therefore, these steps are executed to convert the data into a small clean data set which is used for training. Next stage is classification of drag severity level by using decision tree algorithm. While training the model initially split the model into training dataset and testing dataset.

Model is trained on training dataset and its performance is measured on the testing dataset. This step is necessary since we need to test it on unseen data. In this project we have used Decision Tree algorithm in order to classify the dataset classes of drug severity level. A built in library sklearn is used to implement these algorithms. The input to the model is using the dosage of 15 medical drugs and output is three levels of severity (low, medium, high). Both the models were been instantiated and trained on the training dataset. Later the trained model is saved as a .sav formatted file using pickle library. These saved models can be used for future predictions of drug severity level.

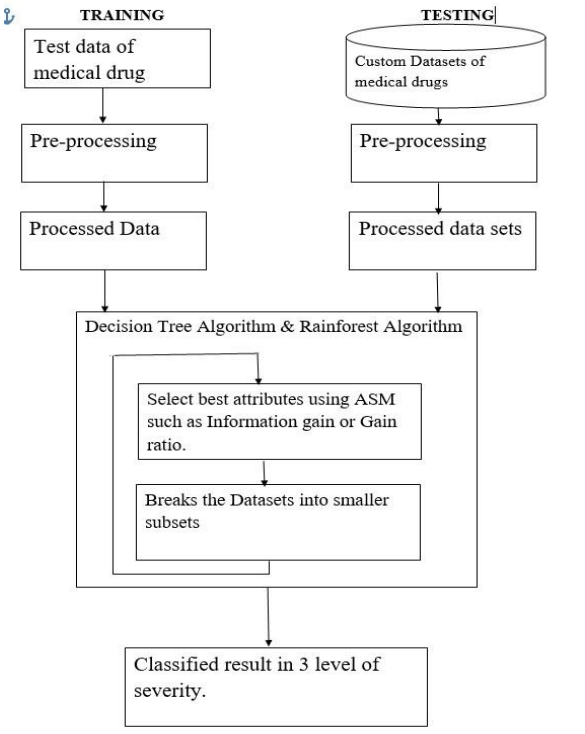

Fig. 1. Block diagram of proposed system

The figure 1 shows the block diagram of proposed system. Initially the user enters the drug names to the system using machine learning algorithm. The severity categories are assigned to each input while training the data set using decision tree algorithm. Finally, the severity class will be displayed to the user.

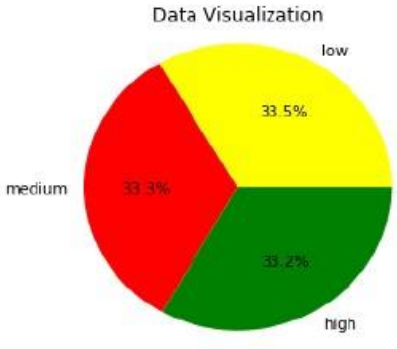

Fig. 2. Data visualization of the proposed system model

The above figure 2 showed the data visualization of the proposed system.

4. Experimental Results

The system is being implemented and following results have been obtained. With Decision tree classifier, the model is 75.4% accurate and with Random forest classifier comprising of 50 estimators and maximum depth 25. Hence, it is proved to be 75.6% is accurate. So we trained the model by using Decision Tree algorithm.

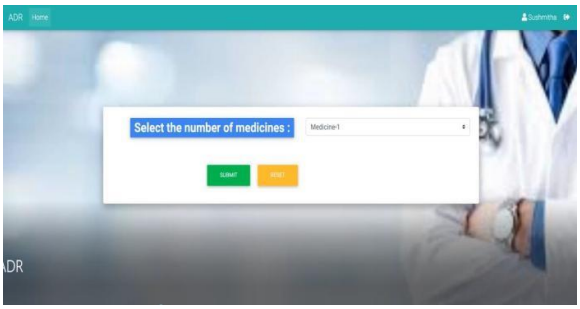

Fig. 3. Input option of the drug for the model designed

Figure 3 is the main page of the project where the user enters the number of drugs taken by the patient.

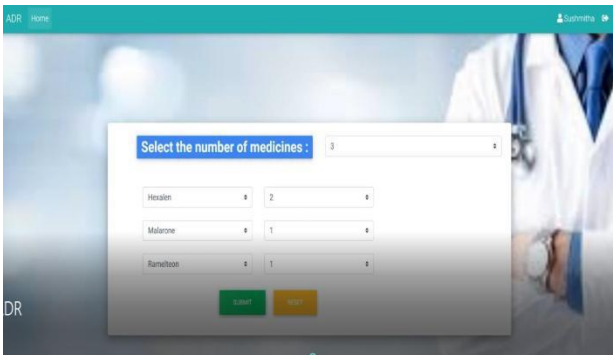

Fig. 4. Drug dosage selection in the interface

Figure 4 is the selection page where the user selects the medicine name and dosage of each medicine.

Fig. 5. Output interface of the proposed system

Figure 5 is the output of the project which shows the severity level of selected drugs.

5. Conclusion

The aim of this paper is to find adverse reactions when two or more medicines prescribed in different dosages by physicians. It is designed for doctors to see which combination of drugs may cause problems in patients. This project is implemented by training a model that has relatively high accuracy and recall with known Lexi-Interact severity values, the goal would be to test it on drug interactions with known severity. Specifically, a drug pair would have a predicted severity and then a panel of clinical pharmacists, individuals familiar with clinical outcomes of drug interactions, would rate the validity of that predicted severity. We intend to find such reactive drugs and report them to the doctors. The doctors can refer the dataset which helps them to treat patients in better way. This reduces the death rate and other risks.

References

[1] Sanjoy Dey, Heng Luo, Achille Fokoue, Jianying Hu and Ping Zhang "Predicting adverse drug reactions through interpretable deep learning framework" The International Conference on Intelligent Biology and Medicine (ICIBM) 2018, Los Angeles, CA, USA, 10-12 June 2018.

[2] Liliya Akhtyamova, John Cardiff, Mikhail Alexandrov, "Adverse Drug Extraction in Twitter Data using Convolutional Neural Network."