PREDICTING DRUG RISK LEVEL FROM ADVERSE DRUG REACTION USING MACHINE LEARNING

MEGHAVATH DHASARATHAM
Professor Department of CSE
Teegala Krishna Reddy Engineering College, Hyderabad, Telangana, India. dasarath.m@gmail.com

B.PRAVEEN YADAV
Department of CSE Teegala Krishna Reddy
Engineering College, Hyderabad,
Telangana, India. bandapraveen107@gmail.com

K.INDUREDDY
Department of CSE Teegala Krishna Reddy
Engineering College, Hyderabad,
Telangana, India. indureddy093@gmail.com

B.PRAVEEN KUMAR
Department of CSE Teegala Krishna Reddy Engineering
College, Hyderabad, Telangana, India bunetipraveen@gmail.com

Abstract -Adverse drug reactions (ADRs) are unintended and harmful reactions caused by normal uses of drugs. Predicting and preventing ADRs in the early stage of the drug development pipeline can help to enhance drug safety and reduce financial costs. Methods: In this paper, we developed machine learning models including a deep learning framework which can simultaneously predict ADRs and identify the molecular substructures associated with those ADRs without defining the substructures a-priori. We evaluated the performance of our model with ten different state-of-the-art fingerprint models and found that neural fingerprints from the deep learning model outperformed all other methods in predicting ADRs. Via feature analysis on drug structures, we identified important molecular substructures that are associated with specific ADRs and assessed their associations via statistical analysis. The deep learning model with feature analysis, substructure identification, and statistical assessment provides a promising solution for identifying risky components within molecular structures and can potentially help to improve drug safety evaluation.

Index Terms -SMOTE and machine learning approaches

I. INTRODUCTION

Our project to work on that problem is the World Health Organization (WHO), an adverse drug reaction (ADR) is generally defined as an unintended and harmful reaction suspected to be caused by a drug taken under normal conditions [1]. It has been recognized that ADRs represent a significant public health problem all over the world. In the United States, it is estimated that over 2 million serious ADRs occur among hospitalized patients, which results in over 100,000 deaths each year [2, 3]. 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.

II. LITERATURE SURVEY

TITLE: Adverse drug reactions: definitions, diagnosis, and management.
AUTHOR: Edwards IR, Aronson JK.
YEAR:2003
DESCRIPTION:
We define an adverse drug reaction as "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." Such reactions are currently reported by use of WHO's Adverse Reaction Terminology, which will eventually become a subset of the International Classification of Diseases. Adverse drug reactions are classified into six types (with mnemonics): dose-related (Augmented), non-dose-related (Bizarre), dose-related and time-related (Chronic), time-related (Delayed), withdrawal (End of use), and failure of therapy (Failure). Timing, the pattern of illness, the results of investigations, and rechallenge can help attribute causality to a suspected adverse drug reaction. Management includes withdrawal of the drug if possible and specific treatment of its effects. Suspected adverse drug reactions should be reported. Surveillance

methods can detect reactions and prove associations.**TITLE:** Large-scale prediction and testing of drug activity on side-effect targets. Nature.
AUTHOR:Lounkine E, Keiser MJ, Whitebread S, Mikhailov D, Hamon J, Jenkins JL, Lavan P, Weber E, Doak AK,
YEAR:2012

DESCRIPTION:
Discovering the unintended 'off-targets' that predict adverse drug reactions is daunting by empirical methods alone. Drugs can act on several protein targets, some of which can be unrelated by conventional molecular metrics, and hundreds of proteins have been implicated in side effects. Here we use a computational strategy to predict the activity of 656 marketed drugs on 73 unintended 'side-effect' targets. Approximately half of the predictions were confirmed, either from proprietary databases unknown to the method or by new experimental assays. Affinities for these new off-targets ranged from 1 nM to 30 μM. To explore relevance, we developed an association metric to prioritize those new off-targets that explained side effects better than any known target of a given drug, creating a drug-target-adverse drug reaction network. Among these new associations was the prediction that the abdominal pain side effect of the synthetic oestrogen chlorotrianisene was mediated through its newly discovered inhibition of the enzyme cyclooxygenase-1. The clinical relevance of this inhibition was borne out in whole human blood platelet aggregation assays. This approach may have wide application to de-risking toxicological liabilities in drug discovery.

TITLE: Face Detection & Face Recognition Using Open Computer Vision Classifiers
AUTHOR: Lahiru Dinalankara
YEAR:2017

DESCRIPTION:
This is a model for face detection and face recognition using several classifiers available in Computer Vision Classifiers (OpenCV). These faces are detected using the Haar-cascade classifier and recognized using different applications like Eigenface, Fisherface, and Local binary pattern histogram (LBPH) algorithms. Apart from implementing this application, the project also compares the results obtained from these algorithms. Faces classifier objects are created using cv2.CascadeClassifier() and eye classifier objects are also created by using OpenCV XML files.

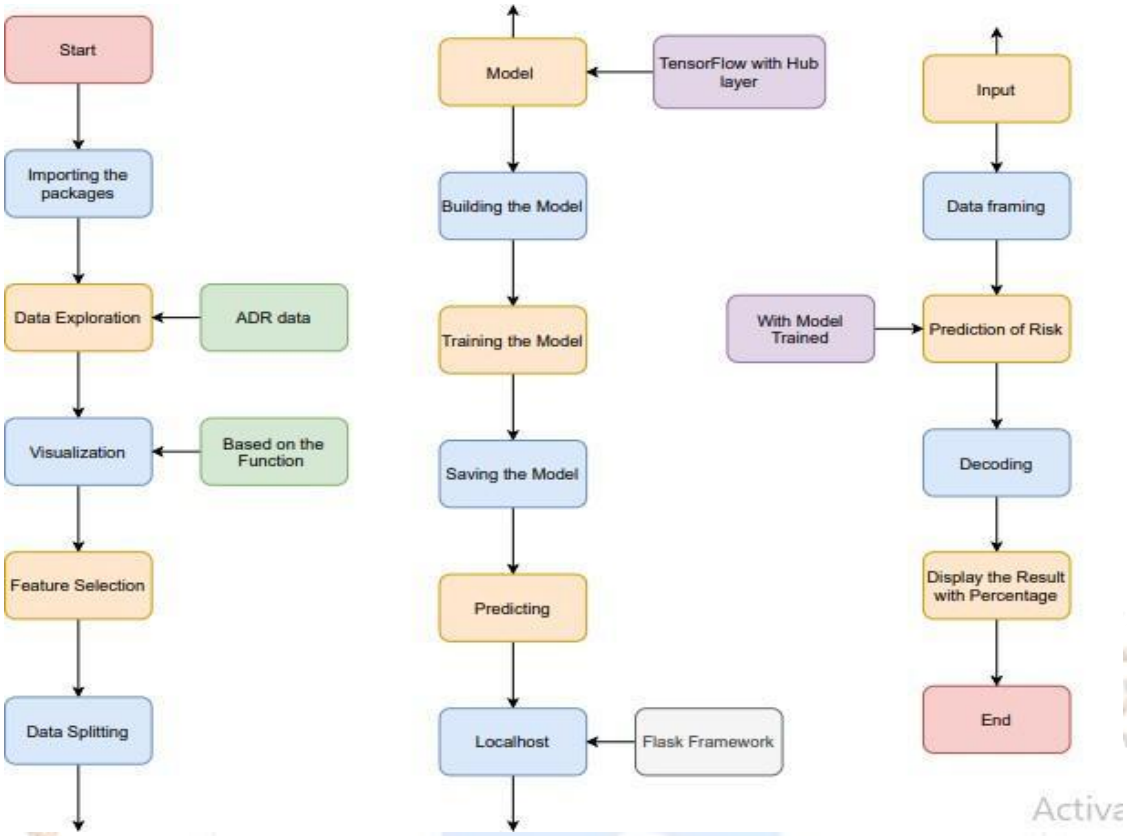
TITLE: Drug-related morbidity and mortality: updating the cost-of-illness model**.**
AUTHOR: Ernst FR, Grizzle AJ.
YEAR:2001
DESCRIPTION:
When people use medications, any number of outcomes are possible. Most commonly, the patient benefits from pharmacotherapeutic interventions; however, adverse events, ranging from minor side effects to death, may occur. Any deviation from the intended beneficial effect of a medication results in a drug related problem (DRP).1 One or more DRPs may develop in a given patient after the initial drug therapy. Researchers have shown that costs associated with DRPs exceed the expenditures for initial drug therapy; that is, the total cost of drug-related morbidity and mortality exceeds the cost of the medications themselves.2,3 DRPs are increasingly recognized as a serious and urgent—but largely preventable—medical problem.

TITLE: Predicting drug side-effect profiles: a chemical fragment-based approach. BMC Bioinformatics**.**
AUTHOR: Pauwels E, Stoven V, Yamanishi Y
YEAR:2011
DESCRIPTION:
Drug side-effects, or adverse drug reactions, have become a major public health concern. It is one of the main causes of failure in the process of drug development, and of drug withdrawal once they have reached the market. Therefore, in silico prediction of potential side-effects early in the drug discovery process, before reaching the clinical stages, is of great interest to improve this long and expensive process and to provide new efficient and safe therapies for patients.

III. PROPOSED SYSTEM

We proposed a multi-classification framework based on SMOTE and classifiers. Secondly, drugs in CSRD were taken as the samples, ADR signal values calculated by proportional reporting ratio (PRR) or information component (IC) were taken as the features. Then, we applied four classifiers: Random Forest (RF), Gradient Boost (GB), Logistic Regression (LR), AdaBoost (ADA) to the tagged data. Our paper is the first to predict drug risk levels based on ADRs. We conduct the research using SMOTE and machine learning approaches. The framework proposed is used to explore the mechanism of ADRs to determine drug risk levels, to guide and assist decision-making in the transition from Rx Drugs to OTC Drugs. More specifically, our framework aims at the problem raised by New England Journal of Medicine as early as 2001 [3], the status change of a drug from Rx to OTC concerning the quality of healthcare, patients' access to drugs, patients' autonomy, and the cost of healthcare.

Finally, as the discussion section mentioned, the optimal combination: PRR-SMOTE-RF based on the above framework was constructed and good classification prediction effect using macro-ROC curve obtained.



IV. MODULES

- Upload ADR dataset: Using this module will load our dataset
- Data Preprocessing: Using this module will process our dataset
- Feature Extraction: Using this module wIll extract important features from the dataset for processing
- Model Generation: Using this module will generate algorithms
- Build LR,RF,DT,Adaboost & GB Classifiers: Using this module will build all algorithms on dataset
- Upload Test Data: Using this module will load test data for prediction
- Predict Result: prediction result for loaded test data

V. CONCLUSION

This paper gave a diagram of utilization of information machine learning procedures in regulatory, clinical, inquire about, furthermore, instructive parts of Clinical Predictions. This paper set up that while the current down to earth utilization of information machine learning in wellbeing related issues is constrained, there exists an extraordinary potential for information mining systems to enhance different parts of Clinical Predictions. Besides, the inescapable ascent of clinical information will build the potential for information mining systems that enhances the quality and reduces cost of social insurance.

VI. REFERENCES

1. Edwards IR, Aronson JK. Adverse drug reactions: definitions, diagnosis, and management. Lancet. 2000;356(9237):1255–9.

2. Lounkine E, Keiser MJ, Whitebread S, Mikhailov D, Hamon J, Jenkins JL, Lavan P, Weber E, Doak AK, Côté S, et al. Large-scale

prediction and testing of drug activity on side-effect targets. Nature. 2012;486(7403): 361–7.

3. Ernst FR, Grizzle AJ. Drug-related morbidity and mortality: updating the cost-of-illness model. J Am Pharm Assoc (1996). 2001;41(2):192–9.

4. Pauwels E, Stoven V, Yamanishi Y. Predicting drug side-effect profiles: a chemical fragment-based approach. BMC Bioinformatics. 2011;12(1):1.

5. Zhang P, Wang F, Hu J, Sorrentino R. Exploring the relationship between drug side-effects and therapeutic indications. In: Proceedings of the 2013 AMIA Annu Symp: 16-20 Nov 2013. Washington DC: American Medical Informatics Association; 2013. p. 1568–77.

6. Wang F, Zhang P, Cao N, Hu J, Sorrentino R. Exploring the associations between drug side-effects and therapeutic indications. J Biomed Inform. 2014;51:15–23.

7. Liu M, Cai R, Hu Y, Matheny ME, Sun J, Hu J, Xu H. Determining molecular predictors of adverse drug reactions with causality analysis based on structure learning. J Am Med Inform Assoc. 2014;21(2):245–51.

8. Duran-Frigola M, Aloy P. Analysis of chemical and biological features yields mechanistic insights into drug side effects. Chem Biol. 2013;20(4): 594–603.

9.Dr.M.Dhasaratham Published a paper on "Distributed Hybrid AODV Algorithm for Path Concern in MANET Using Bio Inspired Techniques" International Journal of Adv Research in Dynamical & Control Systems, Vol. 11, No. 1, 2019, ISSN 1943-023X.**(Scopus Indexed)**

10.M.Dhasaratham, Dr.RP Singh Published a paper on "A survey on data anonymization using map reduce on cloud with scalable two-phase top-down approach" International Journal of Engineering & Technology (IJET), *7 (2.20) (2018) 254-259.* **(Scopus Indexed)**

11. Dr.M.Dhasaratham published a paper on "Criminal investigation analysis using ID3 a machine learning algorithms" The International Journal of Analytical and Experimental Model Analysis, Vol. XII, Issue V, May/2020, ISSN NO: 0886-9367

12.M.Dhasaratham, Dr. RP Singh Published a paper on "Enhanced Security for data sharing in clouds through policy and access control management" International Journal of Research in Advanced Computer Science Engineering, Vol. 3, No. 6, November 2017, ISSN 2454-423X.

13.Dr. M Dhasaratham, Dr. J Rajaram, Durasa Deksisa Geleto, Fayisa Dekebi Tusamo Published a research paper on "Low Resolution image Improvement of Aerial images using SCIKIT Tools" The International Journal of Analytical and Experimental Model Analysis (IJAEMA), ISSN no: 08886-9367

*14.*M.Dhasaratham, Mubeena Shaik, Naseema Shaik Published a paper on "Data mining concepts with customer relationship management" International Journal of EngineerResearch and Applications www.ijera.com ISSN: 2248-9622, Vol. 4, Issue7(Version 6), July 2014, pp.98-100.

15.M.Dhasaratham Published a paper on "A Report of the privacy in Data mining: Speakers Survey" International Journal of Innovative Science and Modern Engineering (IJISME) ISSN: 2319-6386, Volume-2, Issue-4, March 2014.

16.M.Dhasaratham, Dr.RP Singh Published a paper on "Cloud based multimedia resource allocation using PBRA" International Journal of Research in Advanced Computer Science Engineering, ISSN: 2454-6423X, Volume-3, Issue-6, November 2017

17.M.Dhasaratham, Dr.RP Singh Published a paper on "Evaluating the performance speed of multiple protocols for ad-hoc networks" International Journal of Engineering Development and Research© 2017 IJEDR | Volume 5, Issue 4 | ISSN: 2321-9939.

18.M.Dhasaratham, Dr.RP Singh Published a paper on "Ad-Hoc Data processing and its relation with cloud computing process using functional approach" International Journal of Engineering Development and Research| © 2017 IJEDR | Volume 5, Issue 4 | ISSN: 2321-9939.

19.M.Dhasaratham, Dr.R Vivekanandam Published a paper on "Analysis of big data in Cloud computing system based on ad-hoc data processing" Journal of Applied Science and Computations, ISSN NO: 1076-513120.