JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Machine Learning Based Drug Interactions Prediction

Mr. Paluru Chakali Pavan Kumar (M.C.A). Rajeev Gandhi Memorial college Of Engineering and Technology, Nandyal

***Mr. Dr R. Kaviarasan MTech, Ph.D. Rajeev Gandhi Memorial college Of Engineering and Technology, Nandyal**

Abstract

An relationship between two medications in which one drug's pharmacological actions are altered by another is known as a drug-drug interaction (DDI). Positive DDIs typically enhance patients' therapeutic outcomes, however negative DDIs are the primary culprits behind unfavorable medication responses, which may possibly lead to the drug's recall and the patient's demise. As a result, DDI identification has become crucial to the development of new drugs and the management of disease. In this paper, we offer an innovative technique called DDI-IS-SL, which is based on integrated similarity and semi-supervised learning. The cosine similarity method is used by DDI-IS-SL to combine the drug's chemical, biological, and phenotypic data in order to determine how similar the medications' features are. Additionally, the kernel similarity of medicines' Gaussian Interaction Profiles is determined using known DDIs. The Regularized Least Squares classifier is used to determine the interaction probability scores of drug-drug pairings using a semi-supervised learning approach. By way of In comparison to previous approaches, the DDI-IS-SL can obtain better prediction performance using the 5-fold, 10-fold, and de novo drug validation. Additionally, DDI-IS-SL takes less time on average to compute than other comparable approaches. Finally, case studies provide more evidence of how well DDI-IS-SL performs in real-world settings.

1. INTRODUCTION

1.1 Introduction

The world is experiencing a doctor shortage due to the exponential increase in coronavirus cases, particularly in rural areas where there are fewer experts than in urban areas. It takes a doctor between six and twelve years to acquire the relevant credentials. Therefore, it is impossible to increase the number of doctors in a short period of time. In this challenging moment, a Telemedicine framework needs to be powered up as much as feasible. Clinical errors occur often today. Every year, medication errors have an impact on over 200 000 people in China and 100,000 people in the USA. Over 40% of the time when prescribing medicine, doctors err since they only have a limited amount of knowledge to base their decisions on. selecting the best medicine is important for patients who require doctors who have a broad knowledge of patients, antibacterial drugs, and microscopic organisms. Daily, a new study is released that providing daily access to more medications, diagnostics, and clinical staff As a result, selecting a treatment or medication for a patient based on indications and past clinical history proves to be ever more difficult for clinicians. With the rapid growth of the internet and online commerce item reviews have developed into a crucial and essential aspect for purchasing goods everywhere. People all across the world have gotten used to reading reviews and browsing websites before making a purchase decision. While the majority of prior research focused on evaluating expectations and proposals for the E-Commerce industry, the area of healthcare or therapeutic medicines has

only sometimes been covered. The number of people seeking an online diagnosis because they are concerned about their health has increased. As evidenced by a Pew According to a 2013 American Research Center survey, about 35% of users looked for diagnosing health disorders online, while about 60% of adults searched for health-related topics online. an antibiotic With the aim of assisting professionals and assisting people in building their understanding about medications on particular health issues, recommender framework is actually essential.

2. Literature Survey

• The goal of recommender systems is to give customers individualized recommendations and offer solutions to reduce the growing problem of information overload online. Since the mid-1990s, various recommender framework techniques are predicted, and countless Recently, many types of recommender framework code were developed for a variety of applications[1]. The majority of the recommended technological advancements are related to areas like e-government [2], e-business [3], e-commerce/e-shopping [4], e-learning [5], and e-tourism[6], among others. This research focuses on organizing the medication recommender framework and data mining using therapeutic case information. The medication space includes novel recommender innovations that are rare. Through online social networking, communication is enormously advanced and completely unique interested of information is offered on the internet primarily at a rapid pace. The full range of facts must be shared in order to emphasize highlights. of potential edges and accessibility to useful information, objects, people's behaviors, and other things [7]. The therapeutic and health sciences are among the most important fields, therefore ponder about that one. online discussions, blogs, audits, online overviews, etc. about social perspectives [9]. The health-related information supplied through online surveys or comments includes veiled assumptions that derive from completely different medical sources and assist the pharmaceutical industry[8]. In addition, internet shopping, the purchase of a variety of goods through various websites, and online drug delivery are all extremely popular in recent years. Many websites and blogs allow users to assess products based on their satisfaction, product quality, logistics, administrations, and customer feedback.of potential edges and accessibility to useful information, objects, people's behaviors, and other things. The therapeutic and health sciences are among the most important fields, therefore ponder about that one. online discussions, blogs, audits, online overviews, etc. about social perspectives. The health-related information supplied through online surveys or comments includes veiled assumptions that derive from completely different medical sources and assist the pharmaceutical industry. In addition, internet shopping, the purchase of a variety of goods through various websites, and online drug delivery are all extremely popular in recent years. Many websites and blogs allow users to assess products based on their satisfaction, product quality, logistics, administrations, and customer feedback. Which the clients examine for a particular medicine or on quality of administrations.

3. OVERVIEW OF THESYSTEM

3.1 Existing System

Numerous computer methods to predict probable DDIs have recently been developed based on machine learning models. The key characteristics of the medications employed in the signal finding approach developed by Tavonatti et al. to infer DDIs are adverse event profiles for drugs. An effective medicine can be created by combining similarities in its chemical makeup, side effects, protein-protein interactions, and target sequences. Two categories of medication interactions—potential CYP (Cytochrome P450)-related DDIs and non-CYP-related DDIs (NCRDs)—were employed in the development of the INDI (Inferring medication Interactions) framework, which was designed to predict DDIs.

3.1.1 Disadvantages of Existing System

- The system is not implemented Regularized least squares classifier.
- In conjunction with security threats, an emerging concern on ML-based solutions is not suitable for drugs test, namely the non ml classifies are very weak of information from the ML models to the adversaries.

3.2 Proposed System

In this study, we build a computational method (named DDI-IS-SL) to predict DDIs by integrating the chemical, biological, and phenotypic information of medications. These data about medications include their chemical compositions, interactions with their targets, and dosages. drug indications, drug side effects, drug off side effects, drug transporters, drug routes, and known DDIs. In order to determine the feature similarity of pharmaceuticals using the cosine similarity method, we first build a high dimensional binary vector based on this drug information. Additionally, using known DDIs, we calculate the Gaussian Interaction Profile (GIP) kernel similarity of medicines. Their feature similarity and GIP similarity

combine to create the final drug similarity. Then, DDIs are predicted using a Regularized Least Squares (RLS) classifier. Through the use of the node-based drug network, we also determine the relational starting scores for novel medications without drug interactions. diffusion strategy. Therefore, not only for known medications but also for new drugs, our method can forecast probable DDIs. The 5-fold cross validation, 10-fold cross validation, and de novo validation are all used to systematically evaluate the prediction performance of our method compared to other competing methods. The performance of computational approaches is assessed using the AUC (area under the ROC curve) metric. Our approach outperforms other alternatives in terms of AUC.

3.3 Methodology

Database System Module: It contains a drug review dataset with attributes like unique Id, drug name, condition (disease of patient), date, useful count, reviews, and ratings given by the patients on the drugs. Data Preparation Module: It comprises of information investigation and information preprocessing. The real-world information is crude information which can be fragmented, boisterous and messy. Thus, information arrangement is utilized to clean information. it comprises of missing value processing, correlation analysis and removing data redundancy

Data Exploration: a) Find unique number of patient ids to check if a patient has written multiple reviews. b) Analyze number of drugs per condition by considering condition and number of drugs. Data preprocessing:

• Find out the number of missing values for all the attributes.

• Find out the set of co-occurring words (gram) from the reviews starting from uni-gram.

• 1-gram: Analyze the text with a single corpus. But it does not classify the emotion well. Visualization module: It primarily gives the visualization innovation to show a few important information behind the determination case information Classifiers

Classification Algorithm Distinct machine-learning classification algorithms were used to build a classifier to predict the sentiment. Logistic Regression, Multinomial Naive Bayes, Stochastic gradient descent, Linear support vector classifier, Perceptron, and Ridge classifier experimented with the Bow, TF-IDF model since they are very sparse matrix and applying tree-based classifiers would be very time-consuming.

Applied Decision tree, RandomForest, LGBM, and CatBoost classifier on Word2Vec and manual features model. A significant problem with this dataset is around 210K reviews, which takes substantial computational power. We selected those machine learning classification algorithms only that reduces the training time and give faster predictions

4. Architecture

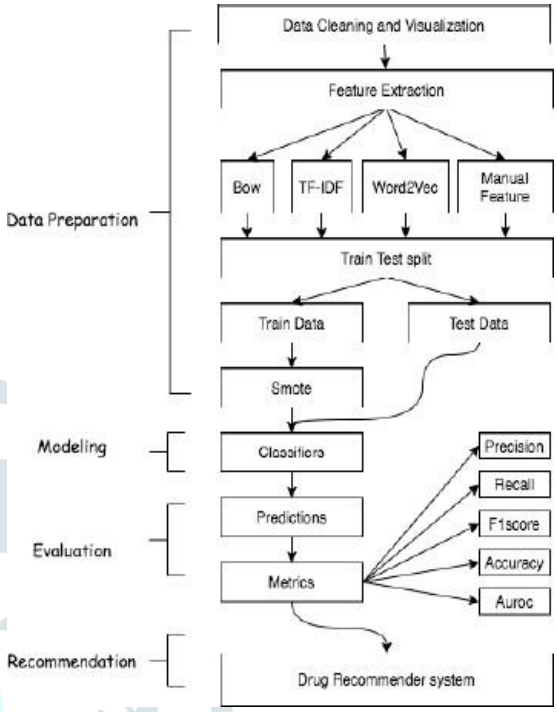


Fig 1: Frame Work of Proposed Method

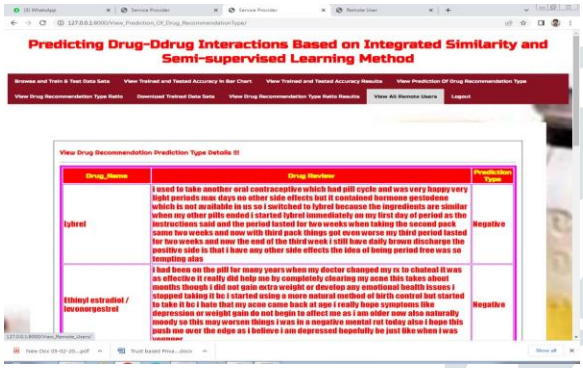
5. RESULTS AND SCREEN SHOTS

Home Page:

Upload Data:

Analysis:

Predict Result:



6. CONCLUSION

✓ Reviews are becoming an essential part of our daily lives; before going shopping, making an online purchase, or visiting a restaurant, we first read reviews to help us make the best choices. Due to this, my motivation for this research was Using various machine learning classifiers, including Logistic Regression, Perceptron, Multinomial Naive Bayes, Ridge classifier, Stochastic gradient descent, LinearSVC, applied on Bow, TF-IDF, and classifiers like Decision Tree, Random Forest, Lgbm, and Catboost, applied on Word2Vec and Manual features method, a study of drug reviews was conducted in order to build a recommender system. Our evaluation of them using five metrics—precision, recall, f1score, accuracy, and AUC score—shows that the Linear SVC on TF-IDF performs best, with an accuracy rate of 93%, eclipsing all other models.

7. Future Enhancement

✓ The Word2Vec Decision Tree Classifier, on the other hand, performed the poorest, reaching only 78% accuracy. We combined the best-predicted emotion values from each method: Linear SVC on TF-IDF (93%), Perceptron on Bow (91%), and LGBM. To create a recommender system, utilize Word2Vec (91%), Random Forest (88%), and multiply the results by the normalized useful Count to obtain the overall score of the medicine by condition.

8. References

[1] S. Fox and M. Duggan. Health online 2013. PewInternet and American Life Project. http://pewinternet.org/Reports/2013/Healthonline.aspx, 2013.

[2] X. Guo, J. Lu, Intelligent e-government services with personalized recommendation techniques, International Journal of Intelligence Systems,2007,401–417

[3] T. Lee, J. Chun, J. Shim, S.-g. Lee, An ontology-based product recommender system for B2B marketplaces, International Journal of Electronic Commerce, 2006,125–155.

[4] J.B. Schafer, J. Konstan, J. Riedl, E-commerce recommendation applications, Applications of Data Mining to Electronic Commerce,Springer, US 2001, 115– 153.

[5] O.R. Zaiane, Building a recommender agent for elearning systems, Proceedings of 2002 International Conference on Computers in Education, 2002, 55–59

[6] T.Hung-Wen,S.VonWun,A personalized restaurant recommender agent for mobile e-service, 2004 IEEE International Conference on e-Technology, eCommerce and e-Service. EEE, 2004, 259–262.

[7] Popescu AM, Etzioni O (2005)Extracting product features and opinions from reviews. Proceedings of the conference on human language technology and empirical methods in natural language processing. Association for Computational Linguistics, pp: 339-346

[8] Mei Q, Ling X, Wondra M, Su H, Zhai C (2007) Topic sentiment mixture: modeling facets and opinions in weblogs. Predictive Modeling of Web Users, pp: 171- 180.

[9] https://www.livewell.pk/ [Accessed on: August 2017].