Development a Model for Drug Interaction Prediction Based on Patient State

Nashwan Ahmed Al-Majmar*
Department of CS and IT, Faculty of Science, Ibb University, Department of Computers, AlJazeera University, Yemen
E-mail: almojammer2015@gmail.com
ORCID iD: https://orcid.org/0000-0001-9757-9327
*Corresponding Author

Ayedh Abdulaziz Mohsen
Department of CS and IT, Faculty of Science, Ibb University, Department of IT, Alqalam University, Yemen
E-mail: ayedh992001@hotmail.com
ORCID iD: https://orcid.org/0000-0003-2224-4076

Mohammed Sharaf Al-Thulathi
Department of CS and IT, Faculty of Science, Ibb University, Yemen
E-mail: mohamedalthulathi@gmail.com

Received: 20 April 2022; Revised: 26 July 2022; Accepted: 24 August 2022; Published: 08 December 2022

Abstract: Drug interactions prediction is one of the health critical issues in drug producing and use. Proposing computational model for classifying and predicting interactions of drugs with high precision is a difficult problem. Medicines are classified into two classes: overlapping, non-overlapping. It was suggested an expert system for classifying and predicting interactions of drugs using various information about drugs, interference reasons and common factors between patients and active substance that causes interference, such as: effective dose of the drug, maximum dose, times of use per day and age of patients considering that only adult category selected. The proposed model can classify and predict interactions of drugs through patient's state taking into consideration that when changing one of mentioned factors, the effect of drugs will be changed and it may lead to appear new symptoms on the patients. There is a desktop application related with the mentioned model, which helps users to know drugs and drugs families and its interactions. Proposed model will be implemented in Python using following classifiers: Logistic Regression (LR), Support Vector Machine (SVM) and Neural Network (NN), which divided data according to their similarity related to the factors of occurrence of drug interference. As these techniques showed good results, NN technology is considered one of the best techniques in giving results where MLPClассifier achieved superior performance with 97.12%.

Index Terms: Prediction, Classification, Drug Interactions, Drug Family, Active Dose, Patient State.

1. Introduction

Pharmacology or clinical pharmacology is one of the areas of medical sciences that is important in our daily life in order to ensure safety, as drug interactions are among the problems of pharmaceutical sciences that cause many risks while taking some medicines, which may sometimes cause to death, where estimated mortality rate due to drug interactions is between 3% and 5% [1]. In 1999, Institute of Medicine report “It should be human to sin: building a safer health system” estimated that from 80,000 people are hospitalized there are 7,000 die every year in the United States due to medication errors [2]. Technology-based intervention has been identified as the primary means to reduce the likelihood of medication errors and improve safety [3]. Computerized systems for physician orders have been proposed as one of the most effective ways to avoid medication errors [4,5]. Particularly significant improvements are expected by pairing these systems with clinical decision support systems, which provide health professionals direct assistance when deciding prescribed medications [6].

Python is one of the most important programming languages that support the field of artificial intelligence and expert systems, due to its ability to deal with big data, and the ease of the classification and prediction process of the data. Python ranked first in 2019 and 2020 years, it helps to increase accuracy and efficiency in giving results of
classification and prediction. The process of predicting drug interactions can be done in general by the following pairs (active substance - active substance, active substance - the patient, active substance - disease state, active substance - physiological condition etc.), in this study we will apply a prediction process, depending on information related to 'active substance - the patient' only (drug interaction based on patient’s state). Drug interaction is a state that occurs when a substance (often another drug) affects the effectiveness of another drug when two drugs are taken together by the patient. This action may be synergistic (when the effect of the drug increases) or antagonistic (when the effect of the drug decreases). In general, the first thing that comes to our minds when we hear the word drug interfering is the drug with drug interfering, however, the interference may occur with food, medicinal plants or herbs [7]. To determine the cause of the interference, several important data must be provided, including the adverse event, the name of well-known drug, and the temporary relationship between drug and its adverse event [8]. In addition, there are several factors leading to appear interference, including [8,9]:

1.1. Factors Related to the Combined Drug Substances

Dosage: The interaction between two drugs does not occur until the dose is exceeded.

Method of use: Some drug interactions take place in the digestive system, such as the interference between tetracyclines. The nature of the drugs: the risk of interactions between the combined drugs increases when these drugs have similar undesirable effects or toxicity, leading to a synergy of the undesirable effects that may reach to the toxicity level.

1.2. Factors Related to the Patient

Age of the patient: the effect of drugs in new-borns and the elderly differs from that of adults, and this is due to the difference at all levels of the pharmacokinetic stage, as the drug metabolism, distribution and excretion change when age changes. Genetic factors: these factors play a role in the occurrence of drug interactions that occur during the pharmacokinetic stage. Subjective allergic reactions: Some allergic accidents may occur with the use of small amounts of the combined drugs, and genetic factors may have a role in these incidents. There are many medical errors that may cause interference, such as wrong drug administration, wrong additional dose, non-restriction (commitment) to the dose given (overdose), wrong given method, wrong time of dose administration.

Our current era has witnessed the use of technology as an aid to get rid of many problems that doctors and pharmacists face [10,11]. The process of predicting drug interactions is a difficult process in terms of studying the active substance and knowing its undesirable qualities and effects relating to the patient diagnosis. We suggested building a model to classify and predict drug interactions based on patient's state as more effective and reliable in improving appropriate decision-making systems. The remaining sections of this work begin with literature review in section 2, then we present section 3 with a methodology, section 4 includes used technology, section 5 presents a model design, experiments and results are presented in section 6 and results in discussion is presented in section 7, finally conclusion is showed in section 8.

2. Literature Review

There are many previous studies and scientific research in the field of drug interactions. Each of studies differs from others by the method used to know the cases and factors causing the occurrence of drug interference, and are often linked with patients and drugs in general, as few of them studied the active substance and the amount of interfering dose.

In [12] it was presented the evaluation of machine learning algorithms for predicting the risk of opioids overdose and identifying individuals at risk of developing opioid overdose for many patients. Programmatic overdose was defined as a food overdose as the primary diagnosis for sensitivity analysis.

In [13] it was presented a method for constructing the prediction drug-drug interaction (PDDI) minimum information model aimed at helping clinicians keep up with the evidence base, using the information elements from the PDDI. The PDDI model contains 10 basic information elements that should be used: clinical consequences, contextual information / modifying agents, drugs involved, frequency of exposure to a drug pair reaction, frequency of harm to persons exposed to a drug pair reaction, reaction mechanism, recommended actions, classification of seriousness, and operational model of the reaction. PDDI is a model that relies on clinical drug information without considering the dose of the active ingredient in which the interaction occurs.

In [14] it was identified common methods used by experts on drug interaction to find evidence about potential drug interactions. Preventing drug interactions is an important goal of drug optimization, as it summarizes potential drug interactions (PDDIS) with clinical decision support is a challenge, and there is no single repository for obtaining PDDI evidence. This study aimed to identify common methods of conducting PDDI searches used by experts who routinely evaluate this evidence. Questions about the search strategy of sex PDDI subjects were organized as follows: (1) PDDI specific information (2) critical (3) clinical consequences (4) management options (5) mechanism (6) health outcomes.

In [15] it was proposed a framework for predicting drug interactions based on four drug-related entities (gene, protein, disease and side effects) as features of each drug pair, based on a technique SVM as a basic classification technique in this work, that uses the traditional machine learning technique (positive and negative group) during data classification.
In [16] it was presented a new approach for prediction of drug interactions based on the use of efficient and useful computational methods for inferring potential DDIs. These methods require standard "bilayers" for both positive and negative interactions for training. The proposed method consists of three steps (application of self-incremental regulation maps to non-prioritized dataset results, use of the pair similarity function to determine the overlap between individual drug characteristics, and use of the SVM classifier to infer DDIs). This method is based on inferring an interaction based on the characteristics between two drugs (chemical similarity, disease, protein), but does not contain dose characteristics.

In [17] it was presented a solution that allows doctors to know drug interactions considering other influencing factors such as patient's age, weight, and physiological condition. The idea of the solution is to predict drug interactions that appear between the drugs that the patient is taking and their mutual effects with the patient's condition. This study focused on knowing the effects of the active substance without knowing the dose of the active substance that causes the interference.

In [18] it was presented a computer system that helps to reduce adverse drug side effects and interactions in intensive care unit (ICU) patients. A solution shows that doctors can reduce the dangerous events resulting from drug interactions by about a half using drug information system "Aidklinik", where the system relies on a database that occurs every 14 days. The information system currently contains about 64,000 drugs. Data were collected from 265 patients in the intensive care unit of Heidelberg University Hospital, where the study included only patients who were given eight or nine drugs at one time. With this measure, adverse events related to drug interactions such as wrong prescriptions, uninformed interactions, contraindications, restrictions on use were reduced by 43%.

In [19] it was presented the system of drug interactions of cytochrome P450. This system contains lists of the effect of enzyme analogues in drugs, and thus we can know drug interactions only for drugs whose interactions effects from competition for this enzyme, as this study was limited to medicines that are included in the treatment of diseases, whose drugs that are metabolized by this enzyme.

In [20] it was developed a system to enhance patient safety by integrating patient demographics, and recovering active drugs from pharmacy systems. The study provided an approach for developing an integrated system to alert about the negative effects in electronic drugs prescription. The proposed method is based on a static database consisting of two parts, the first part is the patient's condition and contains (name, age, sex, residence), and the second part is the part, related to drug information like (medicine, dispensing information, doctor’s prescription and dispensing medicine, quantity and duration).

In [21] it was presented a method to develop a database system for prediction of drug interactions based on four basic groups of data at Songklanagarind University Hospital. The data includes scientific name, trade name, drug group to which the drug belongs and drug-drug interaction data set. The proposed SQL algorithm was used for extracting interactions using a drug interaction dataset (MDI) file. The proposed method in this work is characterized by focusing on the study of the active substance in terms of knowing the amount of dose in which the active substance effects on the body, and also knowing the maximum amount of the patient's intake of this dose per day during regular periods. The proposed method is appropriate in the process of drug classification by simply changing the dose, which is considered one of the main reasons of drug interference.

3. Methodology

This paper presents a new solution to predict drug interference by studying the factors related to the active substance and patients, which are considered as main factors for the occurrence of drug interference. LR, SVM, and NN Supervised techniques have been used in Python to deal with digital data for classification and prediction. Drugs are classified into two categories: medicines that cause drug interaction, and drugs that do not cause drug interaction, these factors are the inputs for the proposed model, as we relied on [8] as a basic reference to obtain the effective dose amounts, the maximum dose, and the number of times of use per day for medicines. This study was limited to adults only, without taking into account the physiological condition (male - female, pregnancy - breastfeeding), and the proposed model was applied to (700 active substances) from (47 drug groups) to which these active substances belong, and thus we can classify the new drugs and deduction of potential interactions through interactions at the level of the active substance, that help to know an important part of drug interactions according to the available information. This model is based on the process of searching for drug interactions in the doctor's prescription. It also can predict prescription in the treatment of a specific disease after the patient's diagnosis process. It was used SQL server database, collected through [8], and which contains drugs and their interactions data (active drug, interaction pharmacovigilance, mechanism of intervention, measure).

4. Techniques Used

4.1. Logistic Regression

In this technique the division or classification will be done by grouping for the number of sections, where the division is based on similar characteristics and the calculation of the error rate, and it depends on the inputs. Logistic regression algorithm was used with error ratio calculation function:
Development a Model for Drug Interaction Prediction Based on Patient State

\[
f(x) = -y \log \frac{1}{1 + e^{-\theta x}} - (1 - y) \log \left(1 - \frac{1}{1 + e^{-\theta x}}\right)
\]

(1)

\[
\exists \ z = \theta x
\]

when \( y = 0 \):

\[
f(x) = \log \left(1 - \frac{1}{1 + e^{-z}}\right)
\]

(2)

when \( y = 1 \):

\[
f(x) = -y \log \frac{1}{1 + e^{-z}}
\]

(3)

\[
\exists \ h_{\theta}(x) = \log \frac{1}{1 + e^{-z}}
\]

\[
= f(x) = \min_{0} \frac{1}{m} \left[ \sum_{j=1}^{m} y^j \left(1 - \log h_{\theta}(x^j)\right) + \left(1 - y^j\right) \left(-\log \left(1 - h_{\theta}(x^j)\right)\right)\right] + \frac{\gamma}{2m} \sum_{j=1}^{n} \theta_j^2
\]

(4)

if \( y = 1 \), we want \( \theta^x \geq 0 \) (not just \( \geq 1 \))

if \( y = 0 \), we want \( \theta^x \leq 0 \) (not just \( < -1 \))

Since the function contains the constant \( \gamma \) in the second part, so we can express it as follow:

\[
A + \gamma B
\]

When \( \gamma \) increases, \( B \) decreases and \( A \) increases.

The architecture of LR algorithm shown in figure 1, first shows how the data is divided into a training set and testing set, secondly, the implementation of the LR algorithm, and thirdly, in the case of success of the LR algorithm, the classification results are analysed on the test group, otherwise the data is verified (training group), finally, drugs are classified into (0 or 1), depending on similar data.

4.2. Support Vector Machine

This technique divides and classifies data into two groups, acceptable or rejected, it is distinguished by finding

![Data processing and classification diagram](image-url)
Boundaries between the data. It separates data non-linearly (nonlinear classification), and contains a gamma variable that helps to change the default gamma setting from "automatic" to "scale". The SVM model differs from others in the data separation process, depending on the kernel, which determines the similarity and difference between two inputs, by calculating the cost and the percentage of error through the following equation:

\[ f(x) = \min_0 C \frac{1}{m} \left[ \sum_{i=1}^{m} \cos(\theta^T x_i)^2 + (1 - y_i)^2 \cos(\theta^T x_i)^2 \right] + \frac{1}{2} \sum_{j=1}^{n} \theta_j^2 \]  \hspace{1cm} (5)

\[ \exists \cos(1(z)) z \geq 1, \cos(0(z)) z \leq 1 \]

if \( y = 1 \), we want \( \theta^T x \geq 1 \) \( \text{not just } \geq 0 \)

if \( y = 0 \), we want \( \theta^T x \leq -1 \) \( \text{not just } < 0 \)

Since the function contains the constant C in the first part, then we can express it

\[ C A + B \]

when C increases, A decreases and B increases

The architecture of SVM is shown in figure 2, that illustrates algorithm work, which is similar in its work to LR algorithm, with a slight difference of SVM technology.

![Fig.2. SVM algorithm architecture](image)

![Fig.3. Neural Network algorithm architecture](image)
4.3. Neural Network

The neural network is a technique, used for analysing data with a large number of features to reach specific results. A neural network model is used in the case of deep learning of neural networks. Neural network algorithm architecture shown in figure 3, which illustrates how the algorithm works, which is similar to the LR algorithm, with a slight difference in that it relies on initial values, determined by the algorithm through the training set.

Neural network contains multiple hidden layers (huge data), that help to make correct decision in classification, as it depends on the number of hidden layers, and the amount of error rate.

5. Model Design

In this study, the proposed system was designed and developed for finding interactions in patient’s prescription that entered in the system. The system finds and shows the interactions, effects of drugs and advices or recommendations for users of the system (doctors or patients), depending on each database containing drugs.

This system was implemented as windows application, that can be worked in clinics to help doctors in making decisions when writing prescriptions and choosing good drugs and it can be worked as sub system in hospitals, clinics and pharmacies and it also can be medical application for mutual users to know drugs interactions in their prescription.

The important part of entity relationship diagram (ERD) is shown in figure 4, that contains the basic entities (pharmaceutical family, scientific formula, drug and disease) and the basic relationships between entities (pharmaceutical family interferences with pharmaceutical family, pharmaceutical family contains scientific formulas, scientific formula interferences with scientific formula, scientific formula contains drugs, the scientific formula is suitable for some diseases).

![Fig.4. ERD](image)

The proposed concept data flow diagram of the system is shown in figure 5:

![Fig.5. Concept DFD](image)

Based on the previous description of the system it was developed a windows application that can work with systems in clinics and hospitals. An electronic model for drug interference detection linked to SQL server database to select interactions and their effects and to present some appropriate measures and some guidelines and tips for this interference. Database also includes information related to drugs, such as the brand name, scientific name of the drug, the name of the group to which the drug belongs and the mechanism of interaction between drugs. In figure 6 it is shown the form of rich with drugs and their interactions database prepared for using in Python:
6. Experimental and Results

Techniques, used in this study proved their ability to classify pharmacological interventions, as the number of drugs is large, knowing the drug group of them is a difficult, in addition knowing their undesirable properties, and effects process is an expensive. A specific set of drug data (active substances) was collected, and patient’s data used as a sample to apply the proposed model in the process of classifying and predicting drugs, which was used in this experiment to test the proposed model for predicting drug interactions, which was manually entered. Due to the lack of ready-made drugs data set, which have many of characteristics, associated with the active substance and the patient at the same time. This data set contains a part of the data with classification 0, that it does not cause drug interference for the patient with known of the patient’s current condition, and a part of classification 1, that it causes drug interference for the patient known of the patient's current condition. In order to train and test the data, which was applied with the following algorithms: (LogisticRegression, SVC (kernel='rbf'), LinearSVC (c = 1.0), SVC (kernel = 'poly'), SVC (kernel = 'linear'), MLPClassifier) that could to classify pharmacological interventions.

The accuracy results of the algorithms in python, used for the proposed model are shown in table 1.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLPClassifier</td>
<td>97.12%</td>
</tr>
<tr>
<td>SVC(kernel='linear')</td>
<td>95.68%</td>
</tr>
<tr>
<td>SVC(kernel='poly', degrees=3, C=1.0)</td>
<td>95.68%</td>
</tr>
<tr>
<td>LinearSVC(C=1.0)</td>
<td>94.96%</td>
</tr>
<tr>
<td>LogisticRegression</td>
<td>93.52%</td>
</tr>
<tr>
<td>SVC(kernel='rbf', gamma=0.7, C=1.0)</td>
<td>58.99%</td>
</tr>
</tbody>
</table>

This experiment was applied by (LR, SVM and NN) techniques as in the previous table in order to understand how the algorithms used in the classification and prediction of drugs work and to know the highly efficient algorithm, by knowing the accuracy of each algorithm. We can predict new entries (active substances or drugs) by obtaining the active substance information shown in figure 7:

It was proposed a computational model for classifying and predicting interactions of drugs, using various information about drugs, interference reasons and common factors between patients and active substance that causes interference, such as effective dose of the drug, maximum dose, times of use per day and age of patients considering that only adult category selected. PC configuration used for the experiment is as follow: DELL, Elite book, 7th Gen, Core i7, Ram 8.

7. Discussion

The proposed model can classify and predict interactions of drugs through the patient's state. Developed desktop application related with the mentioned model helps users to know drugs and drugs family’s interactions, alerts doctors and pharmacists about the presence of interactions between prescribed drugs for the patient, classify most used drugs and provides other additional services such as linking scientific name with trade name and vice versa, and linking recommended drugs with diseases. The proposed model was implemented in python language using the following
Development a Model for Drug Interaction Prediction Based on Patient State

classifiers: Logistic Regression, Support Vector Machine and Neural Network. As these techniques showed good results in the process of drug classification by applying them in the proposed model, NN technology was considered the best technique in giving results through MLPClassifier. MLPClassifier achieved superior performance with 97.12%, while other techniques achieved less accurate results as SVC (kernel = 'linear') achieved performance with 95.68%, SVC(kernel='poly', degree=3, C=1.0) with 95.68%, LinearSVC(C=1.0) with 94.96, LogisticRegression with 93.52%, SVC(kernel='rbf', gamma=0.7, C=1.0) with 58.99%. The evaluations showed that a neural network is an efficient method for drug interaction prediction, determining the effective dose at which interference may occurs and depending on the patient age. The experimental results show that the proposed model is changing drug classification depending on the amount of dose. The idea of the solution is to build an electronic model to predict drug interactions, as it gives the doctor the possibilities to know drug interactions in prescription and present some of the effects resulting from these interactions and some appropriate measures. This facilitates the doctor to determine the prescription suitable to the patient's condition and this is what is superior in the proposed model to other models. In future works through some improvement in the proposed model we can predict and search for drug interactions and interference between a pair or more of drugs by studying the physiological or pathological condition of patients.

Fig.7. Algorithms accuracy

8. Conclusion and Future Works

This study provided the construction of a model for predicting drug interference by the following machine learning techniques: Neural Networks by MLPClassifier, Logistic Regression by LogisticRegression, SVM by [SVC (kernel = 'linear')), LinearSVC (c = 1.0), SVC (kernel = '), poly '), SVC (kernel = 'rbf ')], and comparative studying the accuracy and efficiency of the mentioned techniques in the classification and prediction process, depending on some factors causing interference. As these techniques showed good results in the process of drug classification, by applying them in the proposed model, NN technology is considered one of the best techniques in giving results through applying of the MLPClassifier algorithm, which refers to the use of multiple hidden layers, and multiple nonlinear transformations where it obtained the highest accuracy with 97.12%.

The presented solution is characterized by being easier and more flexible as it is possible to add cases that cause interference to the model and apply the prediction and classification process. Moreover, the proposed model assists the doctor in viewing drug interactions and reviewing them on an ongoing basis, helps to know the drugs most used by patients and can predict potential prescription for treating a disease, and vice versa. In future works through some improvement in the proposed model we can predict and search for drug interactions and interference between a pair or more of drugs by studying the physiological or pathological condition of patients.

References

of serious drug errors, JAMA, 1998, Vol. 280 (1311-6)


[19] ArcMesa Educators; Clinically Significant Drug Interaction with the Cytochrome P450 Enzyme System, Indiana University; 2009.


Authors’ Profiles

Nashwan ahmed Al-Majmar, Associate professor since 2017 with department of computer sciences and information technology, and assistant professor since 2010 with Department of Math’s and Computer Science, Faculty of Science, Ibb University. He is a dean of faculty of Science & Engineering, AlJazeera University, Yemen, since 2019. He received his Ph.D. degree in methods and systems of information protection and security from Saint-Petersburg State University of Information Technologies, Mechanics and Optics, Saint-Petersburg, Russia in 2010. His research interests include Information Security, Data Mining and Software Development. Contact him at: almojammer2015@gmail.com

Ayedh Abdulaziz Mohsen received the Ph.D degree in Technical Sciences from Saint Petersburg Electro technical University‘LET’,Saint Petersburg, Russian Federation, in 2011. He is assisted. Professor in Department of Math’s and Computer Science, Faculty of Science, Ibb University, Yemen in 2012. He is Assoc Professor in Department of IT and Computer Science, Faculty of Science, Ibb University,Yemen in 2019. Dean Vice of Information Technology and Computer Center, Ibb University (Yemen) since 2016. Department head of information technology, Faculty of Engineering and information technology, University of Al-qalam, Yemen. His research interests include Web Development Application,Computer Aided Design System(CADs), Data Mining Application and Software Development.
Mohammed Sharaf Qaid Al-Thulathi received his bachelor's degree in computer science from Department of Mathematics and Computers, Ibb University, Yemen, in 2019. He received first place with bachelor's honors. Currently he works in IT department as system administrator, Al-Kuraimi Bank, Yemen. His research interests include machine learning and systems developments.

How to cite this paper: Nashwan Ahmed Al-Majmar, Ayedh Abdulaziz Mohsen, Mohammed Sharaf Al-Thulathi, “Development a Model for Drug Interaction Prediction Based on Patient State”, International Journal of Intelligent Systems and Applications(IJISA), Vol.14, No.6, pp.28-37, 2022. DOI:10.5815/ijisa.2022.06.03