Artificial intelligence-driven drug interaction prediction

P SONAJI *, L SUBRAMANIAN and M RAJESH

Department of Pharmaceutics, Sankaralingam Bhuvaneshwari College of Pharmacy, Sivakasi, Tamil Nadu, India.

World Journal of Biology Pharmacy and Health Sciences, 2024, 17(02), 297–305

Publication history: Received on 16 January 2024; revised on 16 February 2024; accepted on 19 February 2024

Abstract

Artificial intelligence (AI) is developing at a rapid pace and this has led to revolutionary changes in many fields, including healthcare. Drug interaction prediction, which evaluates possible interactions between various medications to guarantee patient safety and maximize therapeutic outcomes is a crucial component of healthcare. This work investigates the use of artificial intelligence (AI) methods for predicting drug interactions, with a particular emphasis on the combination of natural language processing, knowledge graphs, and machine learning algorithms. The manual curation and experimental research that are frequently used in traditional drug interaction prediction methods limit their scalability and real-time applicability. On the other hand, artificial intelligence (AI) methods use molecular data, electronic health records, and large-scale healthcare data to improve the precision and effectiveness of drug interaction prediction. Deep neural networks and ensemble approaches are two examples of machine learning models that are essential for evaluating various datasets and spotting complex patterns related to drug interactions.

Keywords: Artificial intelligence; Drug interaction prediction; Machine learning; Natural language processing; Healthcare; Prediction models

1. Introduction

The task of gathering, evaluating, and applying the vast amount of knowledge required to resolve intricate clinical issues faces modern medicine. Artificial intelligence programs designed to assist clinicians in diagnosis formulation, treatment decision-making, and outcome prediction have been closely linked to the development of medical AI. They are intended to help healthcare professionals with daily responsibilities by supporting tasks that require the manipulation of knowledge and data. Artificial neural networks (ANNs), fuzzy expert systems, evolutionary computation, and hybrid intelligent systems are examples of these kinds of systems.[1] To create software that can automatically learn from past data to gain knowledge from experience and gradually improve its learning behavior to make predictions based on new data machine learning (ML) is a very practical area of artificial intelligence.

A family of machine learning models with a lengthy history, DL is based on deep convolutional neural networks. Because they are producing incredible results even at human performance levels, DL is very popular these days. The encouraging outcomes of diagnosing associated eye disorders and diabetic retinopathy. [2] Over the previous fifty years, AIM has undergone significant change. Applications of AIM have grown since the development of ML and DL opening doors for personalized medicine as opposed to algorithm-only-based medicine. In the future, predictive models may be utilized in preventative medicine as well as disease diagnosis and treatment response prediction. [3]

The healthcare provider has been characterized by the majority of these challenges. For instance, it has been demonstrated that standalone diagnostic applications have little clinical acceptability unless they are incorporated into integrated environments like electronic medical records.[4] AI includes the use of a computerized system (hardware or software) that can simulate intelligent behavior with little assistance from humans and may be divided into two
categories in medicine: virtual (i.e., informatics and deep learning information) and physical (i.e., robot-assisted systems)

Artificial intelligence (AI) algorithms are generally used to analyze, interpret, or manage data or complex functions with greater accuracy. AI tools make it possible to predict the pharmacokinetics of novel therapeutics, such as their skin- or blood-brain barrier permeability, appropriate dosing, and quantitative structure-property relationship (QSPR) or quantitative structure-activity relationship (QSAR). The use of in silico tools may lead to improved efficiency and lower costs in drug and research projects, given the significance of pharmacokinetic profile prediction of drug candidates.

1.1. Machine learning for improved drug interaction prediction

Drug-drug interactions (DDIs) are the phenomena where the presence of one drug affects the effects of another drug. When DDIs are dosed incorrectly, there can be serious side effects, reduced therapeutic efficacy, or even potentially fatal situations. Recently, artificial intelligence algorithms have emerged as practical instruments for DDI prediction and assessment, aiding in the identification and management of likely drug interactions.[7] The deep and complex interactions between drugs can surpass the capabilities of basic traditional machine learning algorithms, despite the large number of drugs that have entered the market over the past few decades. Because DL can handle complex relations, it is applied in DDI prediction with multiple processing layer concepts. The superior performance of DL in classification tasks over traditional methods is inspired by the architecture of human brains, which leverages its growing application in DDI prediction. In contrast to the manual features engineering used in the traditional ML method, DL carried out the data representation and prediction in a collaborative task. DL appears as a good method for resolving these stochastic problems in a complicated, ill-defined, and highly nonlinear problem as DDI prediction. Deep learning (DL) can be conceptualized as representation learning wherein the machine learns to represent its features through a series of sequential layers. This section contains a description of every top deep learning framework used for DDI extraction and prediction since the advent of deep learning. [8,9,10,11]

2. Common deep learning methods

To accomplish effective prediction and classification, DL algorithms can learn the distribution of the original data by training a deep neural network with several hidden layers to produce abstract high-level features. At present, this method can be successfully used for target detection, protein position point prediction, drug target prediction, drug interaction, and other fields. In practically all scientific and engineering sectors, the DL algorithm is considered one of the frontiers of innovation and production. Convolutional neural networks (CNNs), recurrent neural networks (RNNs), and deep neural networks (DNNs) are the three main types of artificial neural networks currently used in drug discovery.[12]

2.1. Data source integration

Large-scale access to biomedical data (i.e., medical history, medications, allergies, immunization status, blood and urine test results), physiologic signals (ECG, PPG, EEG, arterial pressure), medical images (CT, MRI, PET, X-ray, PET), histopathological images, molecular and all "omics" data will increase the understanding of the clinicopathological events, improve diagnosis accuracy, and improve diagnostic accuracy by allowing early detection of physiological changes.[11,12,13] Solutions that "provide the right treatments to the right patients at the right time" are the cornerstone of precision medicine. AI-based techniques that enable the quick and thorough characterization of the patient will help determine the primary biomarkers that are essential for choosing the best treatment plan promptly. When AI techniques are used with large datasets, patterns or trends that are not normally visible in smaller data sets can be directly identified and their meaning inferred from the data itself. Therefore, there is a greater likelihood that the hypotheses tested in individual scientific studies won't apply to patients who are entirely different from each other if they were based on a cohort of patients with particular characteristics. This knowledge will contribute to the development of novel target therapies, cutting-edge biological understanding, and ultimately, personalized medicine.[14]

Data sharing, which ensures privacy and security for patients, will enable huge and fast progress in healthcare. Currently, automatic methods, such as computer-aided diagnosis (CAD), have promising results and would lead to improvements in patient care workflow and reduce costs related to examinations, the rate of medical interactions, and, as a result, the costs.[15] When AI is used for medical diagnosis, it has been demonstrated to outperform clinicians in some clinical condition detection. However, there remains a great deal of room for improvement in terms of CAD performance when dealing with large data sets across multiple applications.[16]
2.2. Feature engineering

2.2.1. Neural network

Datasets We used the 5,134 drugs we used for predicting drug-target prediction before. Drug feature representation, including chemical properties, such as constitutional properties, topological properties as well as geometrical properties, are often referred to as drug descriptors. All these can be considered as the features of the drugs. We downloaded drug structure information from Drugbank. To get drug descriptors from the drug structure information, we used an online server named the Online Chemical database with a modeling environment [14]. Finally, we retrieved 2,216 drug descriptors for each of the 5,134 drugs. We downloaded 192,284 drug-drug interactions with known DDI types from DrugBank. There are a total of 86 DDI types for the drugs described in DrugBank, which belong to four major groups: antagonism, synergism, potentiation, and interaction with metabolism.[17]

Methods DDI extraction is recognized as a multiclass classification problem for all possible interacting pairs of drugs in the same sentence. Each pair of drugs is classified into one of the predefined types of DDIs or classified as a noninteracting pair. Given a sentence with \( n \) drugs, a total of \( C_n^2 = n(n-1)/2 \) DDI candidates need to be classified. The preprocessing module first blinds drugs tokenizes sentences, normalizes tokens, and filters out noninteracting pairs from DDI candidates.

The CNN module is used for DDI extraction. In the training phase, DDI candidates that are annotated in the training set are positive samples with different types, and the other candidates are negative samples. The task of training is to obtain a CNN model on these samples. In the test phase, all DDI candidates are classified into different types of DDIs or non-DDI.[18]

2.2.2. Predictive models

Drug safety evaluation has been revolutionized by artificial intelligence (AI) and machine learning (ML)-based computational techniques, which are now key enablers in the field. These tools ensure the efficacy and safety of new treatment options while speeding up the drug development process by simulating and predicting ADR and toxicity. Risks associated with medication development, such as unexpected side effects, expensive failures in late-stage clinical trials, and post-market withdrawals, could be reduced. The capabilities and applications of the currently available AI and ML tools are reviewed in Table 1, with an emphasis on open-access, web-based, and commercial options.[19]

Table 1 AI and ML-based tools and software for the prediction and modeling of a drug's ADR and toxicity

<table>
<thead>
<tr>
<th>Tools</th>
<th>Types</th>
<th>Usage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ProTox-II</td>
<td>Free access to a web server</td>
<td>Toxicity prediction</td>
<td>A total of 33 models for toxicity endpoint prediction</td>
</tr>
<tr>
<td>pkCSM</td>
<td>Free access to a web server</td>
<td>ADMET Prediction</td>
<td>14 quantitative regression models and 16 predictive classification models for predicting</td>
</tr>
</tbody>
</table>

2.2.3. Network-based prediction models

The dual-layer integrated cell line-drug network, which connects a drug similarity network and a cell line similarity network by known drug-cell line responses, is the network structure most frequently used for drug response prediction. Chemical structure fingerprints can be used to determine drug similarity, while gene expression and genetic variants can be used to determine cell line similarity. A random walk with restart (RWR) algorithm or a linear weighted model based on the known responses of neighboring nodes was used to make predictions because it was assumed that similar cell lines might respond similarly to similar drugs.[20]

2.3. Data modalities

We concentrate on research using two main data modalities in this review:

2.3.1. Medical imaging modality

In medical imaging research, artificial intelligence (AI) is the most talked about topic at the moment, both for therapeutic and diagnostic applications. Researchers have used artificial intelligence (AI) to provide quantitative evaluations of radiographic features and automatically identify complex patterns in imaging data. This refers to N-dimensional
imaging data obtained from clinical practice, including ultrasound, computed tomography (CT), positron emission tomography (PET), functional MRI (fMRI), structural MRI (sMRI), X-ray, and magnetic resonance imaging (MRI).[21]

2.3.2. Electronic health record

Clinical data from electronic health records (EHRs) is still underutilized, even though artificial intelligence (AI) and its subfield, machine learning (ML), are increasingly applied to imaging data to improve early detection of pancreatic and other cancers. In this area, machine learning has only been applied in a small number of studies to create prediction models using imaging data.[23,24,25]

This comprises free-text data that is both organized and unstructured. Coded data, such as procedure and diagnosis codes, numerical data, such as test results from laboratories, and categorical data, such as demographics, family history, vital signs, and prescriptions, are examples of structured data. Clinical notes and medical reports are examples of unstructured data. We take into account studies that integrate the two modalities imaging and EHR. Nonetheless, there are instances in which the data may only include multiple imaging modalities (such as PET and MRI) or multiple EHR modalities (both structured and unstructured). We view these data as belonging to a single modality, such as imaging or the EHR modality.[22]

With the development of AI and ML models, it is possible to combine multimodal data with high dimensionality[21], different statistical properties, and distinct missing value patterns in a useful way.[26,27] The field of multimodal machine learning can combine various data modalities. Multimodal data fusion has drawn a lot of interest lately as a means of automating diagnosis and clinical outcome prediction. This is evident in the diagnosis and prognosis of Alzheimer’s disease.[28,29,30]

2.3.3. Drug repurposing

A thorough understanding of a drug’s intended and unintended or so-called off-target, effects is crucial for modeling the drug’s toxicity or efficacy in different tissue and cancer types, as well as for comprehending the drug’s underlying mechanism of action (MoA). The axitinib study serves as a compelling example of how drug-target activity profiles are extremely useful for drug repurposing. We encourage the use of publicly available drug-target activity resources here, as opposed to proprietary resources, which were utilized, for example, in Drug Repurposing Hub. These resources can be helpful in training supervised machine learning models for in-silico off-target predictions and drug repurposing.[31]

The number of substances, objectives and interactions addressed; additionally, whether or not an API is offered for programmatic data access in support of AI-based investigations.

To make things simpler, we have separated the different types of compound-target activity data into three categories: binary interactions (which include both active and inactive drug-target pairs), unary interactions (which only include active drug-target pairs), and quantitative bioactivity data (such as from multi-dose Kd, Ki, or IC50 assays). These categories establish whether one has true positive and true negative examples for training the supervised prediction models and whether regression or classification algorithms are appropriate for the target activity predictions. One of the resources listed is the foundation for the majority of in-silico DTI prediction studies.

The most widely used target activity resource for regression modeling (i.e., quantitative drug-target binding affinities prediction) to date is ChEMBL. Algorithms for classification attempt to forecast if a medication is sufficiently potent against the specified target. [32] To prevent reporting overly optimistic drug-target activity prediction results, we have argued that, in addition to the problem formulation (regression vs. classification), at least the following factors should be taken into account in in-silico target prediction studies: The prediction model’s application domain can be assessed using (i) multiple evaluation datasets tailored to specific drug and target families; (ii) an evaluation procedure where nested cross-validation is preferred over standard cross-validation; and (iii) the prediction problem setting, which determines whether the training and test sets of compound-target pairs share common drugs and targets, only drugs and targets, or neither—the latter being frequently the most difficult case. Naturally, the prediction algorithm will have better coverage the more comprehensive the data is in the databases, for example, regarding drug classes and target families. Before recommending a medication be repurposed, the anticipated target activities should also be confirmed through experimentation. That’s why we recently held an IDG-DREAM Challenge, in which teams predicted quantitative target activities using bioactivity data from ChEMBL, DTC, and BindingDB.[33]

2.3.4. NBI network-based inference in drug repositioning

NBI might work well as a tool for repositioning medications. Because NBI relied only on known DTI information, it was unable to predict targets for a new drug for which there was no known target information in the training set. There is a
methodological error here. However, potential targets for a novel drug can be found by combining DBSI, TBSI, and NBI. Diffusion theory is being used to integrate drugs, proteins, and phenotypic features in an effort to develop a novel approach to network inference. Our methods could also be used for the prediction of other biological networks, like drug-gene, drug-disease, and gene-disease networks, by adding more similarity measures among diseases, genes, and drugs.[34]

Table 2 Drug–target interaction resources for target activity predictions

<table>
<thead>
<tr>
<th>Resource</th>
<th>Drug Bank</th>
<th>PubChem</th>
<th>ChEMBL</th>
<th>Drug Target Profiler (DTP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brief description</td>
<td>Combines drug information (i.e. chemical, pharmacological, and pharmaceutical) with drug target information (i.e. sequence, structure, and pathway)</td>
<td>Provides information on chemical structures, identifiers, chemical and physical properties, biological activities, patents, health, safety, and toxicity data.</td>
<td>Most comprehensive manually-curated bioactivity data from HTS of compound activities.</td>
<td>Contains drug target bioactivity data and implements network visualizations. DTP also contains cell-based response profiles of the drugs and their clinical phase information.</td>
</tr>
<tr>
<td>Data type</td>
<td>A, C</td>
<td>B, C</td>
<td>A, B</td>
<td></td>
</tr>
<tr>
<td>Compound s</td>
<td>≥12 K</td>
<td>≥95 M</td>
<td>≥1.9 M</td>
<td>0.9 M</td>
</tr>
<tr>
<td>Targets</td>
<td>≥5K</td>
<td>≥58 K</td>
<td>≥12 K</td>
<td>6 K</td>
</tr>
<tr>
<td>Interaction s</td>
<td>≥18.9 K</td>
<td>≥264.8 M</td>
<td>≥15.5 M</td>
<td>4.4 M</td>
</tr>
<tr>
<td>Mut</td>
<td>√</td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Vis</td>
<td></td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Api</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Ref</td>
<td>35</td>
<td>36</td>
<td>37</td>
<td>38</td>
</tr>
</tbody>
</table>

2.4. Improvement strategies in clinical decision support ddi alerts:

DDI alerts are customized for each patient, taking into account variables like age, gender, current medical conditions, and prescribed drugs. Although CDSS has been shown to reduce medication errors, equally important is the recognition of its inherent limitations so strategies can be developed to fully optimize this technology. Several modifications leading to potential improvements in the effectiveness of CDSS alerts in preventing prescribing errors have been attempted.[39] These improvements have focused on ways to overcome alert fatigue, which is a major barrier to the utility of CDSS. Striking the balance between an all-comprising and a customized DDI database to generate only clinically relevant as well as potential high-severity alerts has been a dilemma for many institutions.[40] It is important to clearly outline any potential adverse clinical outcome(s) for the patient taking interacting drugs (e.g., hyperkalemia, QT prolongation, reduced HIV drug efficacy) so that the clinician can use clinical judgment and pertinent patient information to weigh the benefits against the risks. When the DDI’s mechanism is known, the alert should be explained so the clinician can comprehend the issue and determine a possible course of action. When selecting therapeutic alternatives, for instance, clinicians may find it helpful to know that the mechanism of the interaction involves inhibition of a specific metabolic pathway (e.g., selecting an alternative in the same pharmacologic or therapeutic class that uses a different metabolic pathway for elimination).[41]

2.5. Drug-drug interaction networks

DDI networks simulate how different drugs interact with one another. In a DDI network, drugs are supplied as nodes and their interactions are displayed as edges. Unlike the previous networks, a DDI network does not represent a
biological process. Still, because DDI networks offer a helpful representation of drug interaction data, more researchers are paying attention to them these days. There is a lot of research done on DDI networks in polypharmacy. As we’ve seen, biological networks are a rich representation of biological data because they record both the relationship between entities and the entity itself. DDI networks mimic the interactions between various drugs.[42] AI and ML-based systems bridge various domains for heterogeneous healthcare data analysis and visualization. Predicting potential drug-drug interactions by integrating chemical, biological, phenotypic, and network data.[43]

2.6. Ethical legal consideration
Although using AI in clinical settings has great potential to enhance healthcare, there are now ethical concerns that need to be addressed. Four main ethical concerns need to be resolved for AI in healthcare to reach its full potential: Important considerations include: (1) informed consent to use data; (2) safety and transparency; (3) algorithmic fairness and biases; and (4) data privacy. Not only is the question of whether AI systems are legal, but it is also a politically sensitive one. The goal is to assist legislators in making sure that the ethically challenging circumstances brought about by implementing AI in healthcare settings are addressed early on. The majority of legal discussions surrounding artificial intelligence have focused on the issue of algorithmic transparency limitations. The use of AI in high-risk scenarios has raised the need for transparent, fair, and responsible AI design and governance. The two most crucial components of transparency are the information’s comprehensibility and accessibility. Deliberately making information about algorithmic functionality hard to find is a common practice.[44]

It is said that machines with the ability to follow arbitrary rules and pick up new behavioral patterns pose a threat to our ability to assign blame to their creators or operators. There is reason for concern regarding the alleged "ever-widening" divide, as it poses a threat to "both the moral framework of society and the foundation of the liability idea in law." AI use could leave us without anyone to hold responsible for any harm caused.[45]

3. Pharmacogenomics
Pharmacogenomics is the study of how a person’s genome, or their particular genetic composition, affects how they respond to drugs that are prescribed to them. Protein-coding genes may affect how a drug is treated by facilitating its absorption, breaking it down, or moving it to desired or undesirable sites. It will be easier to observe how individuals with similar genetic variations may respond to a given treatment when they are grouped. Thiopurine methyltransferase testing, which is used to identify candidates for thiopurine drug therapy which is used to treat autoimmune diseases like Crohn’s disease and rheumatoid arthritis is an example of pharmacogenomics. Pharmacogenomics offers numerous applications in precision medicine and related domains, facilitating drug enhancement throughout the entire pharmaceutical life cycle.[46]

Integration into healthcare workflow in AI

A new era in healthcare has begun with the introduction of Artificial Intelligence (AI), which presents revolutionary opportunities for streamlining operational procedures and improving patient care delivery. Chatbots and virtual assistants powered by AI can be incorporated into healthcare processes to interact with patients, give them information, and provide support. This improves communication, teaches patients, and motivates them to take an active role in their healthcare. (0) The ability of AI to evaluate enormous datasets, offer predictive insights and enhance clinical Healthcare stakeholders’ attention has been drawn to double-blind peer-reviewed journal decision-making, which has created an urgent need for structured frameworks that enable its seamless integration.[47]

- Operational Efficiency: As indicated by shorter wait times and more efficient use of resources, there was a discernible 30% average increase in operational efficiency.
- Diagnostic Error Reduction: A comparative analysis conducted before and after the framework’s adoption revealed a noteworthy average reduction in diagnostic errors of 25%.
- Improvement in Patient Satisfaction: As a result of more individualized care and streamlined procedures made possible by the integrated AI framework, patient satisfaction scores increased by an average of 20%.[48]

3.1. Continuous improvement
Based on the real-time data, use predictive models to evaluate the possibility of drug interactions while taking a variety of variables into account, such as patient demographics, medical histories, and prescription schedules.[49] Install an alert system that, upon detection of possible drug interactions, instantly sends out notifications or cautions. To help determine the best course of action, categorize alerts according to the clinical importance and level of severity of the
anticipated interactions. Integrate the alert system seamlessly with the current clinical workflows to guarantee that medical personnel respond promptly. Give medical professionals decision support resources, such as information on substitute drugs and practical suggestions. [50] Automate the process of creating reports that summarize known drug interactions, possible outcomes, and recommended actions. Provide reporting systems that let the members of the healthcare team communicate with one another and make joint decisions easier. Create a feedback loop to let real-world data and results drive ongoing updates and improvements to predictive models. [51] Give the system the flexibility to adjust to new drug interactions, patient conditions, and growing medical knowledge. Use user-friendly interfaces to help patients comprehend possible interactions and give them the tools they need to interact with healthcare providers efficiently. To preserve patient privacy and adhere to laws governing healthcare data, make sure that strong security measures are in place. Respect the laws and guidelines controlling the use of medical data in AI applications. [52]

4. Conclusion

In summary, the application of artificial intelligence (AI) to the prediction of drug interactions is a ground-breaking development in healthcare that has great potential to improve patient safety and treatment effectiveness. The integration of natural language processing, knowledge graphs, and machine learning algorithms has produced a multimodal solution to the challenges involved in anticipating and comprehending drug interactions. Prediction accuracy is increased and real-time analysis and adaptation are made possible, resulting in more prompt and individualized interventions in clinical settings. Knowledge graphs have made it easier to comprehend the complex interactions between medications, target proteins, and biological pathways comprehensively. By providing insights into the underlying mechanisms and aiding in the prediction of direct drug-drug interactions, this comprehensive approach enables healthcare professionals to make more informed decisions regarding treatment strategies. By extracting useful information from scientific literature natural language processing has expanded our knowledge base even further and updated our understanding of drug interactions as new research is conducted making it possible to analyze enormous datasets, such as molecular data and electronic health records, which allows for a more thorough evaluation of possible drug interactions.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

Reference


[17] Predicting Drug-Drug Interactions Using Deep Neural Network Xinyu Hou


[19] Application of artificial intelligence and machine learning in early detection of adverse drug reactions (ADRs) and drug-induced toxicity Siyuan Yang, Supratik Kar *


[21] Artificial intelligence-based methods for fusion of electronic health records and imaging data Farida Mohsen1, HazratAli1, Nady El Hajj1,2 & Zubair Shah1*scientific reports


[23] The role of artificial intelligence in medical imaging research Xiaoli Tang https://doi.org/10.1259/bjro.20190031

[24] Bihan Shen , Xiaoqin Mou , Yiuxue Li , Hong Li https://doi.org/10.1016/j.jgg.2021.03.007


Clinical decision support systems: great promises for better management of patients drug therapy Turgeon,a,b,c,d...