Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, Pages. 832-838
Received 15 April 2021; Accepted 05 May 2021.

A Review of Drug Target Interaction Prognostication Using Artificial Intelligence

K. Nandhini[1], Dr. G.V. Sriramakrishnan[2]
[1]Research Scholar, Department of Information Technology, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai
[2]Department of Information Technology, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai
[1]kgnandhugopi@gmail.com
[2]greatsri8@gmail.com

Abstract
Identifying new Drug-Target Interactions (DTIs) plays avital role in the drug discovery and repositioning phase as the effectiveness of currently available antibiotics declines. Drug target interaction is described as the fastening of a drug to a particular target that causes changes its actions or functions. Medicinal targets, a particle in a body, for example, some proteins and nucleic acids, which is essentially correlated with a disease process and that, could be treated by a medication in order to attain the intended curative effect. Drug Interaction is not only assisting the understanding of disease processes, but it also aids in the detection of unusual remedialaction or medication ill effects. . Drug-target relationship prognostication allows researchers to better understand drug activity, disease pathology and drug side causes. Hence, the prediction of drug-target interaction is an important arena of medicine detection and reusing. In contrast to wet-lab experimentations that cause expensive and time-consuming experimental methods, Artificial intelligence is critical for identifying DTIs in a way that is accurate, reliable, and high-throughput. Nowadays computational methods help to predict the interactions and they do so with reasonable accuracy. This paper reviews the various methodologies used in Artificial Intelligence and the computer-based methods for foreseeingthe biological targets and the datasets used to forecast the collaboration.

Keywords: Drug Target Interaction, Prediction, Datasets, DTI

1.Introduction

Drug-target interactions are interactions that occur when drugs interact with particular molecular targets (DTIs). The accurate detection of DTIs will greatly aid the drug development process. As a result, drug research has been prioritized to establish the effectual techniques for recognizing and predictingDTIs from a wide range of chemical compounds and protein targets.Artificial Intelligence plays a vital role to predict the DTI effectively and efficiently. Drug Target Interaction helps to know the collaboration between the medicine and otherelementsthat can reduce or increase the drug's effectiveness and the cause of side effects of drugs.

It is categorized intothreetypes. {a} Interactions of drugs and drugs (interaction between drugs):A drug-drug interaction occurs when two or more substances interact with one another. This medication is combined with the treatment and Over the Counter (OTC) medications. (b) Drugs with food (Drug food interactions): This happens when drugs areresponding with food and other dieteticadditions or drinks (including liquor). (c) Drug with disease (Drug-disease interactions):The existing medical conditions may get worse or aggravates this type of drug

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, Pages. 832-838
Received 15 April 2021; Accepted 05 May 2021.

interaction. For example, people should take medicine according to their existing medical conditions.

The drug therapy cycle is divided into three phases. The substance of the drug forms temporary bonds with the target unit at first. The involved drug interacts with the pharmacological activities in the second level, causing a favorable or unfavorable change. After that, it begins to shift away from the biological target. [1].

2. Artificial Intelligence – Introduction

Artificial intelligence refers to the training given to machines to behave and think like humans. The word Artificial Intelligence was coined by John McCarthy, who is widely regarded as the father of the field, described as "the science and engineering of making intelligent machines." Some subsets of Al include Machine Learning, Deep Learning, Natural Language Processing, Expert system, etc. Many authors use Machine learning, Deep learning, and Neural network methods to predict the DTI.

3. Literature Review

Kanika Sachdev etal.[1] proposed that. the prediction has been carried out using feature-based techniques such as SVM, ensemble, and miscellaneous under the title "A comprehensive review of feature based methods for drug target interaction prediction.Predictions are addressed using several classifiers such asSupport Vector Machine, Relevance Vector Machine, Rotation Forest, etc. Comparison is made among the drug and target features using several datasets like DrugBank, Pubchem, KEGG, etc.In comparison to SVM, rotation forest provides more accurate and precise outcomes. The objective of Random forest is to progress the efficacy of drug target interface prediction, and it is capable of handling large datasets while being resistant to overfitting.

Under the heading 'Computational prediction of drug–target interactions using chemogenomic approaches: an empirical survey,' Ali Ezzat et al[2] proposed a survey of the chemogenomic system under five categories, namely Neighborhood model, Bipartite Local models, Network diffusion models, Matrix factorization model, and Feature-based classification models. As compared to binary data sets, which have been used in the majority of prior DTI prediction work, continuous data sets representing drug–target binding affinities (rather than discrete 0(zero) and 1(one) values) are more useful and meaningful.

Huiqing Wang et al,[3], has suggested five datasets Nuclear Receptor, GPCR, Ion Channels, E, andDrugBank_FDA are used to estimate the evaluation of MDADTI. The outcomesdisplayed that the MDADTI technique can successfullyrecognize unknown DTIs. MDADTI's prognostic presentation was improved, especially in GPCR and NR datasets.

Shanshan Hu et al [4] proposed to use a novel convolution neural network that comes under deep learning is used to predict drug-target interactions based on drug structure and protein sequence. Three benchmark datasets were handled to investigate possible drug-target protein communications. Two datasets were created using the KEGG DRUG database, whereas the third created using the DrugBank database. These systems performed well and reach the accuracies of

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, Pages. 832-838
Received 15 April 2021; Accepted 05 May 2021.

more than ninety percent for all target families. Additional data extracted from DrugBank was used to determine the model's generalization.

F. Rayhanet al.[5] proposed a Deep Neural Net Architectures such as FRnet-Encode and FRnet-Predict where FRnet-Encode aims to abstract the convolutional structures, and FRnet-Predict which attempts to organize the interactions through the extracted features under FRnet-DTI. Investigation displays that this process works well on three out of four drug-target interaction gold standard datasets (DrugBank, KEGG, BRENDA and Super Target) to mine info around drug-target interface under the two metrics auROCand auPR. Tang et al. proposed a new method for learning drug kernel matrix and target kernel matrix by using a marginalized denoising model on heterogeneous networks to predict drug-target interactions [6]. To progress the prediction's accuracy, the above-mentioned model has been applied to a varied system. This process may be useful for recommending new drug applicants and repositioning existing medications. The benchmark datasets DrugBank, KEGG, BRENDA, and SuperTarget were used in this study. Yamanishi et al. introduced a Bi-partite network model in 2008 to predict DTIs for four target proteins: E, IC, GPCR and NR and many methods were proposed to predict the DTI accurately. .

4. Computational methods

Vitro, one of the traditional prediction strategies, has met anumber of limitations such as time consumption and monetary costs. The recently developed silico methods help to predict the potential interaction candidates in a good manner. Some bioinformatics areas, such as sickness associated with miRNA prediction, disease-genes prediction, how the protein interact with each other and subcellular position prediction, have shown promising results using computational methods. As a result, there is a constant and pressing need meant to progress DTI predictions. Three approaches of Computational Methods are available to predict DTI and it is shown in Fig.1. They are Ligand based, Docking-simulation and Chemogenomic.

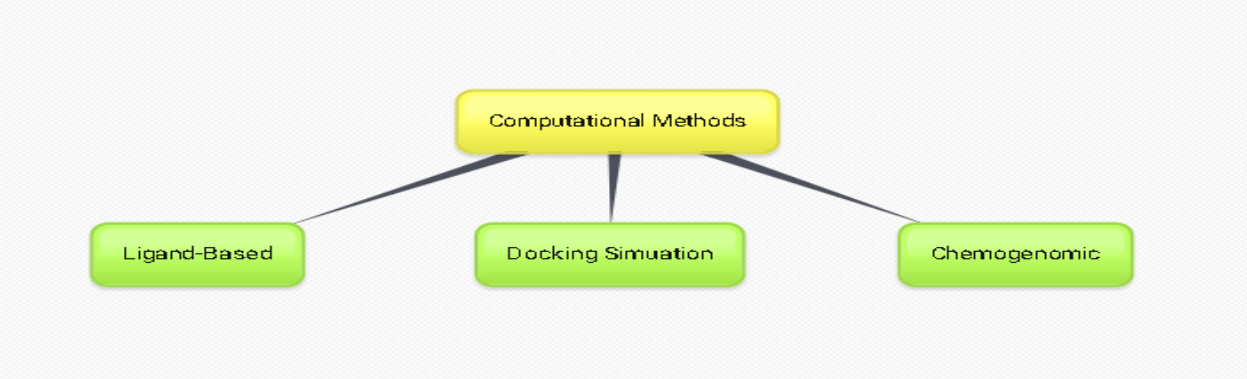

Fig. 1. Computational Methods

Ligand-based methods, is similar to QSAR (Quantitative Structure-Activity Relationship) and based on assumptions that the identical molecules bound to same type of proteins and have their alike properties. These methods search for similarities between a new ligand and existing protein ligands to predict interactions. However, this is not give good result when the number of known ligands are inadequate**.**

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, Pages. 832-838
Received 15 April 2021; Accepted 05 May 2021.

Docking simulation methods are inapplicable because there are many proteins with inaccessible three-dimensional (3D) structures since simulation involves three-dimensional (3D) structures of proteins. Moreover, it has a complex structures of membrane proteins. Docking simulations usually take a long time, making them particularly inefficient.

Chemogenomic methods, Chemogenomic methods and techniques have recently been used to overcome the difficulties of the above two methods in medicine discovery and reusing on a broad scale**.** To predict interactions, these approaches use knowledge from both drugs and proteins at the same time. The benefit of this approach is that it makes use of a vast amount of biological data that is widely accessible in public datasets. Chemogenomic approaches can be further separated into two sections. Methods are built on features and methods based on similarity [1]. Feature-based approaches use a vector of descriptors to represent the drug target pair. Feature vectors are used to represent the training data. To predict novel interactions, these vectors are feed into numerous machine learning models such as random forest, SVM, and others.The similarities between drugs or targets are measured using different similarity calculation techniques in similarity-based methods. Nearest neighbor methods, Bipartite local models, and Matrix factorization methods are all examples of kernel functions that use similarity matrices.

5. Datasets used for DTI prediction

Several factors such as the concentration of more than onemolecule and their inter-molecular interactions that influence drug target interaction. Various databases contain different amounts of information about drug molecules, target proteins, and interactions. These detailsare used to accurately envisage a wide range of novel potential interactions. The following are some of the most significant datasets:

DrugBank [8] The information in these databases are freely accessible bioinformaticsresources with extensive annotations that combine with the data on drugs with comprehensive drug-target data. DrugBankhas the details of the current drug statements, number of small molecules and biotech drugs, nutraceuticals and experimental drugs. KEGG [9]The Kyoto Encyclopedia of Genes and Genomes (KEGG) contains genetical data and organicroutes from around the world. It also includes information on a variety of disorders, medicines, and chemical compounds. PubChem[10] - This contains a variety of biochemicalmixtures as well as their associated activities.UniProt [11]. This database is exclusively created for proteins. this is a free, open-source database that contains data on protein systems as well as their biotic functions.Pfam [12] For protein families, the Pfam database was developed. It contains the proteins, as well as their annotations and sequence alignments. Hidden Markov models are used to create these sequence alignments.

SuperTarget [13] is a web based data repository for target data that incorporates medication information and its side-effects, remedial conditions, drug metabolism, trails, and Genetic factorontology.MATADOR [14] The chemical-target interactions are contained in this resource. It varies from other datasets, such as the DrugBank, in that it contains both direct and unintendedconnections. GLIDA [15] G-protein Coupled Receptors (GPCRs) are a class of proteins which have been produced specifically for GLIDA It includes both the biological and chemical knowledge about GPCRs and their different ligands. STITCH [16] STITCH stands for

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, Pages. 832-838
Received 15 April 2021; Accepted 05 May 2021.

Search Tool for Chemical Interactions. It contains information on drug target interactions, mineralconstructions, bind studies, and metabolic paths in one place. ChEMBL [17] is a database of bioactive molecules with drug properties that has been manually compiled. TDR Targets [18], an easily accessible source that makes numerous gene data and chemical databases available to aid in the detection and ranking of drugs and targets in illnessbacillus.PDTD [19] PDTD is a recordto identify targets. SIDER [20] Side Effect Resource (SIDER) combines the facts about medications, targets and drug negativecauses to provide a more complete picture of drug behavior and adverse reactions. NIST [21] is a Mass Spectral Library and it has assembled a list of numerous peer-reviewed databases.

6. Metrices

Evaluation of various database are performed by various metrics. These metrics helps to compare the evaluation and helps to find the ideal method for implementation. Some of the metrics are Positive and Negative, Precision and accuracy, Mathew's correlation coefficient, AUC (Area Under Curve), Time, Memory usage, AUPR, recall, F1-score, ROC, etc.,

7. Comparative analysis of various methods

By analyzing the various methods in the literate review, it is observed that the drug target interaction is predicted by various methods available in artificial intelligence like machine learning and deep learning. Table 1. summarizesthe comparative analysis.

Table1: Comparative Analysis of various methods

Ref. No	Methods	Datasets	Results
1	Feature-based methods such as Support vector machine, Relevance vector machines, Rotation forests were used.	DrugBank, Pubchem, KEGG	In contrast to SVM, rotation forest provides more reliable and precise results.
2	Models used for prediction include neighbourhood models, bi-partite models, network diffusion models, matrix factorization models, and feature-based classification models.	NR, GPCR, IC, E, KEGG.	The matrix factorization method is relatively a faster algorithm in predicting.
3	Multimodal Deep Autoencoder	DrugBank FDA, NR, GPCR, IC, E was utilised	The predictive performance of MDADTI was greatly enhanced, particularly in GPCR and NR datasets.
4	Deep Learning Method	DrugBank, KEGG	Accuracy levels were achieved for the target families up to 92.0 percent, 90.0 percent, 92.0 percent, and 90.7 respectively.
5	Deep convolutional neural network method, FRnet-Encode and FRnet-predict	DrugBank, KEGG, BRENDA,SuperTarget	In three-quarters of drug-target interaction datasets, FRnet-DTI performs well in the advanced process.
6	New methods Drug kernel Matrix	DrugBank, KEGG,	This tool can be used to

	and Target kernel matrix.	BRENDA SuperTarget	suggest new drug candidates as well as reposition existing drugs.
7	Bipartite network model	NR, GPCR, IC, E and DrugBank_FDA	Improves accuracy in predicting DTI.

7. Conclusion

Essential Contribution in drug-target interaction prediction is reviewed in this paper. A Literature review frames the various methodologies, datasets, and performance are discussed. It also listed the various metrices that are used to evaluate performance. The Comparative section helps to outline the optimal result under various methodologies. Many machine learning methods have been projected for drug target prediction. Machine learning approaches are highly reproducible, portable and it is easily extended for future research. As DTI prediction is one of the vital fields in research and has been continuing for almost a decade, it is important to enhance the prediction using novel frameworks and technologies to improve the efficacy.

References:

[1] Kanika Sachdev and Manoj Kumar Gupta, "A comprehensive review of feature-based methods for drug target interaction prediction", Journal of Biomedical Informatics, vol.93, 103159. May 2019 .
[2] Ezzat, Ali, et al. "Computational prediction of drug–target interactions using chemogenomic approaches: an empirical survey. "Briefings in bioinformatics vol.20.4, pp. 1337-1357, Jul2019 .
[3] Huiqing Wang et al, "A Novel Approach for Drug-Target Interactions Prediction Based on Multimodal Deep Autoencoder", Frontiers in Pharmacology, vol.10, p.1592, 2020
[4] Hu, ShanShan, et al. "Predicting drug-target interactions from drug structure and protein sequence using novel convolutional neural networks, "BMC bioinformatics vol.20.25, pp.1-12, 2019.
[5] F. Rayhan et, "FRnet-DTI: Deep convolutional neural network for drug-target interaction prediction", Heliyon 6, no. 3, p. e03444, 2020.
[6] Tang et al, "Drug-target interactions prediction using marginalized denoising model on heterogeneous networks", BMC Bioinformatics, vol. 21, no. 1, pp.1-29, 2020.
[7] Yamanishi Y, Araki M, Gutteridge A, et al. "Prediction of drug-target interaction networks from the integration of chemical and genomic spaces. Bioinformatics", vol. 24, no. 1, pp.i232-i240, 2008
[8] V. Law, et al., "DrugBank 4.0: shedding new light on drug metabolism," Nucleic Acids Research, vol. 42, no. D1, pp.D1091-D1097, 2014.
[9] M. Kanehisa, S. Goto, "KEGG: kyoto encyclopedia of genes and genomes", Nucleic Acids Research, vol.28, no.1, pp. 27–30, 2000.
[10] E.E. Bolton, et al., "PubChem: integrated platform of small molecules and biological activities", Annual Reports in Computational Chemistry, vol. 4, pp. 217-241, Elsevier, 2008,
[11] U. Consortium, UniProt: a hub for protein information, Nucleic Acids Research, vol.43, no.D1, pp. D204–D212, 2014.
[12] R.D. Finn, et al., "Pfam: the protein families database," Nucleic Acids Research, vol.42 no. D1, pp. D222–D230, 2013
[13] N. Hecker, et al., "SuperTarget goes quantitative: update on drug–target interactions", Nucleic Acids Research, vol.40, no.D1, pp. D1113–D1117, 2011.

[14] S. Günther, et al., "SuperTarget and Matador: resources for exploring drug-target relationships", Nucleic Acids Research, vol.36 (suppl_1) pp. D919–D922, 2007.
[15] Y. Okuno, et al., "GLIDA: GPCR—ligand database for chemical genomics drug discovery—database and tools update", Nucleic Acids Research, vol.36 (suppl_1), pp.D907–D912, 2007.
[16] M. Kuhn, et al., "STITCH 4: integration of protein–chemical interactions with user data", Nucleic Acids Research, vol.42 no.D1, pp. D401–D407, 2013.
[17] A. Gaulton, et al., "ChEMBL: a large-scale bioactivity database for drug discovery", Nucleic Acids Research, vol.40, no.D1, pp. D1100–D1107, 2011.
[18] M.P. Magariños, et al., "TDR Targets: a chemogenomics resource for neglected diseases", Nucleic Acids Research, vol. 40, no.D1, pp. D1118–D1127, 2011.
[19] Z. Gao, et al., "PDTD: a web-accessible protein database for drug target identification", BMC Bioinformatics, vol.9, no.1, pp.1-7, 2008
[20] M. Kuhn, et al., "A side effect resource to capture phenotypic effects of drugs", Molecular systems biology, vol.6, no.1, pp.343, 2010.
[21] P. Ausloos, et al., "The critical evaluation of a comprehensive mass spectral library", Journal of the American Society for Mass Spectrometry, vol.10, no.4, pp. 287–299, 1999.