INTERNATIONAL JOURNAL IN PHARMACEUTICAL SCIENCES

Journal Homepage: https://www.ijpsjournal.com

OPEN ACCESS™

Review Article

Artificial Intelligence in Predicting Drug Target Proteins- A Review

Ishu Garg*, Harish, Mohit Kumar, Vishvanshi Tyagi, Nishant

ITS college of Pharmacy, Muradnagar, Ghaziabad-201206, Uttar Pradesh, India

ARTICLE INFO

Received: 30 July 2023
Accepted: 01 Aug 2023
Published: 09 Aug 2023

Keywords:
Artificial Intelligence

DOI:
10.5281/zenodo.8229884

ABSTRACT

The world needs pharmacological medication therapy for the diagnosis and treatment of numerous ailments. Time constraints make it necessary to find novel chemical compounds for the same, nevertheless, current methods of screening medicinal molecules and their target proteins cannot be completed as quickly as with AI. Assigning the right target during drug molecule development is crucial for effective treatment. A disease involves several proteins. Designing any medication molecule for its specific target over disease is considerably aided by predicting the structure or makeup of the targeted protein. AI can help in the creation of structure-based drugs by foreseeing the effect of a molecule on the target as well as safety concerns by anticipating a target protein's 3D structure in line with its chemical environment. Using the AI tool Alpha-Fold, which is based on DNNs, the 3D protein structure was predicted by examining the distances between neighbouring amino acids and the corresponding angles of the peptide bonds. In a study, RNN was used to predict the protein structure. A recurrent geometric network (RGN) is said to be composed of three stages: computation, geometry, and assessment. Here, the torsional angles for a particular residue and a partially formed backbone obtained from the geometric unit upstream of this served as the input and output for encoding the fundamental protein sequence. The final unit produced the 3D structure. As a result, it is cutting edge to use AI approaches to screen medications based on the study of target proteins.

INTRODUCTION

The simulation of human behaviour with regard to the mental processes involved in problem-solving is known as artificial intelligence (AI). This includes markings for mechanisms from reading, observation, planning, interpretation, reasoning, correction, speech recognition, linguistics, and other areas of human cognitive science [1,2]. AI works by helping robots learn from prior experiences, connect actions and efforts to outcomes, recognise and fix errors, adapt to new and arbitrary input values, and easily carry out human-like activities through extensive scenario

***Corresponding Author:** Ishu Garg
Address: *ITS college of Pharmacy, Muradnagar, Ghaziabad-201206, Uttar Pradesh, India*
Email ✉: atharvrai10@gmail.com

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

analysis. It makes work easier by analysing, filtering, sorting, forecasting, scoping, and figuring out massive volumes of data to follow the best implementation techniques for coming up with an ideal solution [3, 4]. The discovery and development of new drug compounds, aiding pharmaceutical companies in developing therapies for difficult and rare diseases, and improved drug adherence and dosage are just a few of the significant applications of AI in the pharmaceutical business as of 2019. Finding more effective participants for clinical trials, introducing robotic pharmacy systems to fill prescriptions and dispense medication, and improving clinical data comprehension and analytical outcomes are all examples of using AI [5, 6, 7, 8].

AI uses Natural Language Processing (NLP) to coordinate with regular human speech and converts it internally to a language that a machine can understand. Deep Learning (DL) also succeeds in this task. By processing a huge quantity of data and identifying similar or novel patterns in the data, computers can do certain jobs with the least amount of human intervention when using these technologies for AI training [9, 10].

AI IN DRUG DISCOVERY

The process of developing a new drug for the benefit of human health has always been difficult, but AI has simplified it. There are five steps in the process where AI and machine learning (ML) could be used. The methods of scientific discovery are changing as a result of how machine learning and artificial intelligence are enabling scientists to enhance their scientific intuition with a deep insight into enormous research data sets [11, 12]. This approach was used by benevolent AI to identify a number of substances, including chemical A, which has the potential to treat ALS (Amyotrophic Lateral Sclerosis). A survey conducted by BenchSci and The Science Academy in December 2017 came to the following conclusions: Target identification and acceptance, safety evaluations, compound discovery, lead optimization, preclinical studies, formulation, clinical studies, including clinical trial recruitment, approval, marketing, and post-marketing surveillance are stages that use AI technology in drug discovery methods [13, 14].

AI in drug screening: The process of finding and creating a drug might begin after creating a drug library of candidate drugs. Nine out of every ten therapeutic compounds are unsuccessful in Phase II clinical trials and receiving regulatory approval using the current strategy. Deep neural networks (DNNs), Nearest-Neighbour classifiers (RF), extreme learning machines (SVMs), and Nearest-Neighbour classifiers (RF) are some of the algorithms that were developed for predicting in-vivo activity and toxicity as well as synthesis feasibility. As a result, AI narrows down the list of chemicals to those that are both inexpensive to synthesise and have medicinal promise [11, 15, 16, 17].

Prediction of the physicochemical properties: Physical and chemical characteristics of a drug, such as its solubility, partition coefficient (log P), degree of ionisation, and intrinsic permeability, can have an indirect impact on its pharmacokinetics and target receptor family. As a result, these characteristics must be taken into account when developing a new drug. It is possible to anticipate physicochemical properties using a variety of AI-based methods [18]. For instance, ML analyses sizable data sets created during the program's training phase's compound boost. In order to produce workable molecules via DNN and afterwards anticipate their attributes, drug design algorithms take into account molecular labels like SMILES strings, potential energy measurements, electron density around the molecule, and coordinates of atoms in 3D [19,20].

To track the six physicochemical parameters, Zang et al. developed a quantitative structure-property relationship (QSPR) work chart. To predict the

lipophilicity and solubility of diverse chemicals, neural networks based on the Absorption-Distribution-Metabolism-Elimination-toxicity (ADMET) predictor and ALGOPS programme have been utilised. DL techniques like undirected graph recursive neural networks and graph-based convolutional neural networks (CVNN) have been used to determine the solubility of molecules [21]. For the prediction of the acid dissociation constant of medicinal molecules, ANN-based models, graph kernels, and kernel ridge-based models were frequently developed. Similar to this, data on the cellular permeability of a variety of compounds has been generated using cell lines studied, such as Madin-Darby canine kidney cells and human colon adenocarcinoma (Caco-2) cells, and is then given to AI-assisted predictors [22].

By using 745 compounds for training, Kumar et al. developed a programme of six predictive models, including SVMs, ANNs, k-nearest neighbour algorithms, LDAs, probabilistic neural network algorithms, and partial least square (PLS), which were later applied to some compounds to predict their intestinal absorptivity based on factors that influence absorption, such as molecular surface area, molecular mass, total hydrogen count, molecular refractivity, molecular volume In silico models based on RF and DNN have been created to predict the intestinal absorption of a variety of chemical substances by humans. Thus, AI plays a significantly important role in the development of a drug, to predict both desired physicochemical properties, and desired bioactivity [24, 25].

Prediction of toxicity: To prevent any drug's harmful effects or secondary pharmacological effects, it is imperative to foresee the toxicity that it may cause. To determine a compound's toxicity, in-vitro assays based on cells are performed as an initial step, followed by pre-clinical or animal trials. There are some web-based programmes that can help lower the cost of toxicology, such as LimTox, pkCSM, admetSAR, and Toxtree. Advanced AI-based methods predict a compound's toxicity based on input features or search for similarities between compounds. The US Food and Drug Administration (FDA), the Environmental Protection Agency (EPA), and the National Institutes of Health (NIH) organised the Tox21 Data Challenge to assess various computational techniques to predict toxicity. An ML algorithm called DeepTox outperformed all approaches by identifying potential and kinetic features like molecular weight (MW) and Van der Waals volume given within the chemical descriptors of the molecules and could effectively predict MW and VDW [19, 23, 27, 33, 34, 35].

With an ML-based methodology, eToxPred approximated toxicity and the viability of synthesising tiny chemical compounds. Similar to this, toxicity forecasting also makes use of open-source programmes like TargeTox and PrOCTOR [36].

Prediction of the target protein structure: Assigning the right target for effective illness treatment is crucial for designing therapeutic drug molecules. A disease involves a number of proteins. Therefore, it is essential to predict the molecular structure of the target protein while designing the therapeutic molecule in order to treat the condition selectively. Because the design is in line with the chemical environment of the target protein site, AI helps structure-based drug discovery by anticipating the 3D protein structure. This aids in predicting the action or effect of a compound over the target protein as well as safety considerations before synthesising the compound. In order to forecast the 3D target protein structure, the AI tool AlphaFold, which is based on DNNs, was used to examine the distance between neighbouring amino acids and the corresponding angles of the peptide bonds [37, 38, 39].

RNN was used to predict the protein structure in a study by Al Qurashi. A recurrent geometric network (RGN), which the author defined as three

steps (computation, geometry, and assessment), was taken into consideration. The primary protein sequence was encoded in this case, and the input and output were the torsional angles for a specific residue and a partially finished backbone acquired from the geometric unit upstream of this. The 3D structure was the output of the final unit. The distance-based root mean square deviation (dRMSD) measure was used to evaluate how far the anticipated and experimental structures deviated from each other. The RGN parameters were tuned to minimise the dRMSD between the anticipated and experimental structures [38, 39, 40, 41].

Using MATLAB and a non-linear, three-layered NN toolkit built on a feed-forward supervised learning and back propagation error technique, a study was done to predict the 2D structure of a protein. The input and output data sets were trained in MATLAB, and the NNs served as learning algorithms and performance judges [12, 41, 42, 43, 44].

Prediction of bioactivity: Any medication molecule's efficacy is directly influenced by how well it binds to the desired protein or receptor. Drug molecules won't be able to provide the required therapeutic response if they don't interact with or have any relationship to the targeted protein. Potential interactions between created medication compounds and undesirable proteins or receptors can result in toxicity. DTBA, or drug target binding affinity, is crucial for predicting drug-target interactions, it might be concluded. AI-based approaches analyse a medicine's ability to bind to a target by looking at the traits or similarity between the drug and its target site [42, 43, 44, 45]. There are online tools for predicting their interactions, such as ChemMapper and the Similarity Ensemble Approach (SEA). To determine DTBA, many algorithms use ML and DL, including KronRLS, SimBoost, DeepDTA, and PADME. To determine DTBA, ML-based methods like Kronecker-regularized least squares (KronRLS) compare the similarity between drug compounds and target or any receptor protein molecules [46, 47].

As they use network-based techniques and don't rely on the existence of the 3D protein structure, DL methods have outperformed ML in terms of performance. Some DL approaches used to assess DTBA include DeepDTA, PADME, WideDTA, and DeepAffinity [48].

Predicting drug–protein interactions: Drug protein interactions are crucial to a therapy's success. Understanding a drug's efficacy and effectiveness, as well as how it will interact with a particular receptor protein, is crucial for drug repurposing and preventing poly pharmacy. The precise prediction of ligand-protein interactions made possible by a variety of AI techniques has improved and safeguarded medicinal efficacy. In order to find new compounds and their interactions with four important targets, Wang et al. described a model utilising the SVM approach that was constructed based on primary protein sequences and structural properties of small molecules. The model was trained on 15,000 protein ligand interactions [45, 46, 47, 48].

By combining pharmacological and chemical data with two RF models, Yu et al. were able to predict potential drug protein interactions and validate them against well-known platforms like SVM with excellent sensitivity and specificity. Additionally, these modes are capable of predicting associations between drug and target, as well as associations between target and disease. To collect the best data for the subsequent creation of iDrugTarget, Xiao et al. used the Neighbourhood Cleaning Rule and Synthetic Minority Over-Sampling Technique. The four subpredictors (iDrug-GPCR, iDrug-Chl, iDrug-Enz, and iDrug-NR) for determining interactions between a drug and G-protein-coupled receptors (GPCRs), ion channels, enzymes, and

nuclear receptors (NR) have been combined into one prediction model [48, 51, 52].

This skill was also employed to help with drug repurposing and prevent poly pharmacology. A medication that has been repurposed is eligible for Phase II clinical trials right away. Through the inhibition of the human retinoic acid receptor-related orphan receptor-gamma t (ROR-t), topotecan is also used to treat multiple sclerosis. Cellular network-based deep learning technology (deepDTnet) has been investigated to predict the therapeutic usage of topotecan. Self-organizing maps (SOMs), which fall under the unsupervised category of machine learning, are employed in drug repurposing, which employs a ligand-based strategy to find novel off-targets for a set of drug molecules. The system is trained on a predetermined number of compounds with known biological activities before being applied to the analysis of various compounds. A recent study explains, the use of DNN to repurpose existing drugs with proven activity against SARS-CoV, HIV, influenza virus, and drugs that are 3C-like protease inhibitors [46, 53, 54].

Drug protein interactions aid in the prediction of poly pharmacology, or a drug's propensity to bind with many receptors and cause side effects or secondary pharmacological effects that are not intended. In order to create safer medicinal compounds, AI can build a new molecule using the principles of polypharmacology. Several substances can be linked to a variety of targets and off-targets using an AI platform called SOM and the extensive databases that are available [47, 53, 54].

Li et al. showed how to use KinomeX, an online AI platform that uses DNNs to identify the polypharmacology of kinases based on their chemical compositions. It has practical applications in determining a drug's overall selectivity towards certain kinase subfamilies and kinases, which aids in the development of new chemical modifiers. Ligand Express, an AI platform, is used to identify receptors that can interact with a specific small chemical and result in both on- and off-target interactions. Assisting in comprehending any potential side effects of the medication [49, 50].

REFERENCES

1. Ramesh A. Artificial intelligence in medicine. Ann. R. Coll. Surg. Engl. 2004, 86:334–338.
2. Miles J., Walker A. The potential application of artificial intelligence in transport. IEE Proc.-Intell. Transport Syst. 2006, 153:183–198.
3. Yang Y., Siau K. MWAIS, 2018. A Qualitative Research on Marketing and Sales in the Artificial Intelligence Age.
4. Wirtz B.W. Artificial intelligence and the public sector—applications and challenges. Int. J. Public Adm. 2019, 42:596–615.
5. Smith R.G., Farquhar A. The road ahead for knowledge management: an AI perspective. AI Mag. 2000, 21:17–17.
6. Lamberti M.J. A study on the application and use of artificial intelligence to support drug development. Clin. Ther. 2019, 41:1414–1426.
7. Beneke F., Mackenrodt M.-O. Artificial intelligence and collusion. IIC Int. Rev. Intellectual Property Competition Law. 2019, 50:109–134.
8. Steels L., Brooks R. Routledge, 2018. The Artificial Life Route to Artificial Intelligence: Building Embodied, Situated Agents.
9. Bielecki A., Bielecki A. Foundations of artificial neural networks. In: Kacprzyk Janusz., editor. Models of Neurons and Perceptrons: Selected Problems and Challenges. Springer International Publishing, 2019. 15–28. Polish academy of sciences, Warsaw, Poland.
10. Kalyane D. Artificial intelligence in the pharmaceutical sector: current scene and future prospect. In: Tekade Rakesh K., editor. The Future of Pharmaceutical Product

Development and Research. Elsevier, 2020. 73–107.
11. Vyas M. Artificial intelligence: the beginning of a new era in pharmacy profession. Asian J. Pharm. 2018, 12:72–76.
12. Mak K.-K., Pichika M.R. Artificial intelligence in drug development: present status and future prospects. Drug Discovery Today. 2019, 24:773–780.
13. Sellwood M.A. Artificial intelligence in drug discovery. Fut. Sci. 2018, 10:2025–2028.
14. Chan H.S. Advancing drug discovery via artificial intelligence. Trends Pharmacol. Sci. 2019, 40(8):592–604.
15. Álvarez-Machancoses Ó, Fernández-Martínez J.L. Using artificial intelligence methods to speed up drug discovery. Expert Opin. Drug Discovery. 2019, 14:769–777.
16. Fleming N. How artificial intelligence is changing drug discovery. Nature. 2018, 557:S55–S55.
17. Dana D. Deep learning in drug discovery and medicine, scratching the surface. Molecules. 2018, 23:2384.
18. Zang Q. In silico prediction of physicochemical properties of environmental chemicals using molecular fingerprints and machine learning. J. Chem. Inf. Model. 2017, 57:36–49.
19. Yang X. Concepts of artificial intelligence for computer-assisted drug discovery. Chem. Rev. 2019, 119:10520–10594.
20. Hessler G., Baringhaus K.-H. Artificial intelligence in drug design. Molecules. 2018, 23:2520.
21. Lusci A. Deep architectures and deep learning in chemoinformatics: the prediction of aqueous solubility for drug-like molecules. J. Chem. Inf. Model. 2013, 53:1563–1575.
22. Kumar R. Prediction of human intestinal absorption of compounds using artificial intelligence techniques. Curr. Drug Discovery Technol. 2017, 14:244–254.
23. Rupp M. Estimation of acid dissociation constants using graph kernels. Mol. Inf. 2010, 29:731–740.
24. Chai S. A grand product design model for crystallization solvent design. Comput. Chem. Eng. 2020, 135:106764.
25. Thafar M. Comparison study of computational prediction tools for drug–target binding affinities. Frontiers Chem. 2019, 7:1–19.
26. Öztürk H. DeepDTA: deep drug–target binding affinity prediction. Bioinformatics. 2018, 34:i821–i829.
27. Lounkine E. Large-scale prediction and testing of drug activity on side-effect targets. Nature. 2012, 486:361–367.
28. Mahmud S.H. iDTi-CSsmoteB: identification of drug–target interaction based on drug chemical structure and protein sequence using XGBoost with over-sampling technique SMOTE. IEEE Access. 2019, 7:48699–48714.
29. Gao K.Y. Interpretable drug target prediction using deep neural representation. In: Lang Jérôme., editor. Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI, 2018. pp. 3371–3377.
30. Feng Q. Padme: a deep learning-based framework for drug–target interaction prediction. arXiv. 2018 arXiv: 1807.09741.
31. Karimi M. DeepAffinity: interpretable deep learning of compound–protein affinity through unified recurrent and convolutional neural networks. Bioinformatics. 2019, 35(18):3329–3338.
32. Pu L. eToxPred: a machine learning-based approach to estimate the toxicity of drug candidates. BMC Pharmacol. Toxicol. 2019, 20:2.
33. Basile A.O., Yahi A., Tatonetti N.P. Artificial intelligence for drug toxicity and safety.

Trends Pharmacol. Sci. 2019,40(September (9)):624–635. doi: 10.1016/j.tips.2019.07.005. Epub 2019 Aug 2. PMID: 31383376, PMCID: PMC6710127.

34. Lysenko A. An integrative machine learning approach for prediction of toxicity-related drug safety. Life Sci. Alliance. 2018, 1:YYY–ZZZ.
35. Gayvert K.M. A data-driven approach to predicting successes and failures of clinical trials. Cell Chem. Biolo. 2016, 23:1294–1301.
36. Jimenez-Carretero D. Tox_(R) CNN: deep learning-based nuclei profiling tool for drug toxicity screening. PLoS Comput. Biol. 2018, 14.
37. Wan F., Zeng J. Deep learning with feature embedding for compound–protein interaction prediction. bioRxiv. 2016, 2016.
38. AlQuraishi M. End-to-end differentiable learning of protein structure. Cell Syst. 2019, 8:292–301.
39. Hutson M. AI protein-folding algorithms solve structures faster than ever. Nature. 2019, XX:YYY–ZZZ.
40. Avdagic Z. Artificial intelligence in prediction of secondary protein structure using CB513 database. Summit Transl. Bioinf. 2009, 2009:1.
41. Tian K. Boosting compound-protein interaction prediction by deep learning. Methods. 2016, 110:64–72.
42. Wang F. Computational screening for active compounds targeting protein sequences: methodology and experimental validation. J. Chem. Inf. Model. 2011, 51:2821–2828.
43. Xiao X. iDrug-Target: predicting the interactions between drug compounds and target proteins in cellular networking via benchmark dataset optimization approach. J. Biomol. Struct. Dyn. 2015, 33:2221–2233.
44. Park K. A review of computational drug repurposing. Transl. Clin. Pharmacol. 2019, 27:59–63.
45. Zeng X. Target identification among known drugs by deep learning from heterogeneous networks. Chem. Sci. 2020, 11:1775–1797.
46. Achenbach J. Computational tools for polypharmacology and repurposing. Fut. Med. Chem. 2011, 3:961–968.
47. Li X. Prediction of synergistic anticancer drug combinations based on drug target network and drug induced gene expression profiles. Artif. Intell. Med. 2017, 83:35–43.
48. Reddy A.S., Zhang S. Polypharmacology: drug discovery for the future. Expert Rev. Clin. Pharmacol. 2013, 6:41–47.
49. Li Z. KinomeX: a web application for predicting kinome-wide polypharmacology effect of small molecules. Bioinformatics. 2019, 35:5354–5356.
50. Cyclica . Cyclica, 2017. Cyclica Launches Ligand Express™, a Disruptive Cloud–Based Platform to Revolutionize Drug Discovery.
51. Corey E., Wipke W.T. Computer-assisted design of complex organic syntheses. Science. 1969, 166:178–192.
52. Grzybowski B.A. Chematica: a story of computer code that started to think like a chemist. Chem. 2018, 4:390–398.
53. Klucznik T. Efficient syntheses of diverse, medicinally relevant targets planned by computer and executed in the laboratory. Chem. 2018, 4:522–532.
54. Segler M.H. Planning chemical syntheses with deep neural networks and symbolic AI. Nature. 2018, 555:604–610.

HOW TO CITE: Ishu Garg*, Harish, Mohit Kumar, Vishvanshi Tyagi, Nishant, Artificial Intelligence in Predicting Drug Target Proteins- A Review, Int. J. in Pharm. Sci., 2023, Vol 1, Issue 8, 72-78. https://doi.org/10.5281/zenodo.8229884