



INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X

Impact Factor: 6.078

(Volume 7, Issue 4 - V7I4-1472)

Available online at: <https://www.ijariit.com>

Application of AI in drug discovery

Mohit Singh Negi

mohitnegi165@gmail.com

Uttaranchal Institute of Management, Dehradun,
Uttarakhand

Anosh Singh

anoshsingh007@gmail.com

Uttaranchal Institute of Management, Dehradun,
Uttarakhand

ABSTRACT

Target-driven drug discovery is a process in which a known target is used to search for small molecules that interfere with it or influence its role in cells. These methods function well for easily druggable targets with a well-defined structure and well-understood interactions within the cell. However, due to the complexity of cellular interactions and a lack of understanding of intricate cellular pathways, these methods are severely limited. By detecting novel associations and inferring the functional significance of various components of a cellular pathway, AI will conquer these obstacles. It extracts useful knowledge from a broad dataset using complex algorithms and machine learning techniques. QSAR modeling based on structure, de-novo drug design, automated synthesis planning is just a few of the cutting-edge applications available. Screening of the compounds along with optimizing the lead compound, goal validation and selection, nonclinical research and studies, and clinical drug trials are all places where AI is used extensively.

Keywords: Artificial Intelligence, Deep Learning, Drug Discovery, Machine Learning, QSAR

1. INTRODUCTION

There are over 10^{60} molecules in the chemical space that encourages the creation of many drug molecules^[9]. The lack of advanced technology, on the other hand, restricts drug production, leads to a costly and time-consuming challenge that could be solved with AI^[10]. Hit and lead molecules can be distinguished by AI, allowing for faster drug target validation and structure design optimization^{[11][11]}. The widespread use of ML, particularly deep learning with a variety of scientific fields, as well as advances in computing hardware and software, are all contributing to this development. Most of the initial skepticism about AI's applications in pharmaceutical discovery is dissipating, which is good news for medicinal chemistry.

For computer-assisted drug discovery, machine learning algorithms have been commonly used. Due to the ability to automatically extract features from training data and fetch the

non-linear relationship between input and output data, the methods of deep learning (consisting of internal artificial layers of processing neural networks^[5]) has regained popularity. Deep learning strategies have properties that complement conventional machine learning methods that use human-crafted molecular descriptors. Deep learning has undergone a late revival of interest in drug discovery, resulting in an unparalleled proliferation of novel modelling methods and applications. The ever-improving advances in deep learning have already helped several areas of the chemical sciences. Quantitative structure-based (QSAR structure activity relationship) modelling, de-novo drug design, and automated synthesis planning are all discussed. This opinion piece highlights a few characteristics which has helped deep learning methods to thrive in chemo-informatics, outperforming established approaches in some cases.

2. CUTTING-EDGE DRUG DISCOVERY APPLICATIONS OF AI

2.1. Structure based QSAR modeling

QSAR stands for Quantitative Structure-activity Relationship. The models are a long way from being able to predict complex biological characteristics like compound effectiveness and side effects, but its computational modeling can easily deduce compounds in huge quantities or basic physicochemical parameters like log D or log P^[13]. Shortage in observational validations, error in training data set, and limited training groups are all issues that QSAR-based models face. To address these issues, latest AI techniques has been developed, like deep learning and relevant modelling studies, big-data analysis and modeling can be used to evaluate the efficacy and safety of drug design. Merck sponsored a 'QSAR ML challenge' in 2012 for the investigation of the benefits of deep learning in the pharmaceutical industry's process of developing drugs. As for drug candidates dataset ADMET (absorption, distribution, metabolism, excretion, and toxicity), DL models outperformed conventional ML methods in terms of predictability^{[6][7]}.

2.2. De-novo Drug Design

De-novo Drug Design (DNDD) refers to the use of computational growth algorithms to create new compounds

that meet a collection of chemical constraints. The term "de-novo" translates to "from the beginning", implying about the approach can be used to create new molecular entities without the use of a starting template. De-novo drug design has many benefits, including the discovery of a larger chemical space, the creation of new critical and advanced property potential for therapies that are enhanced and novel, and time/cost-effective manner of candidates for the drug production. In de-novo drug design, one of the biggest issue is the created molecular structure's accessibility synthesis^[14]. A chemical property or any desired characteristics, such as bounds for solubility predefined, specific groups of chemicals constituted in the design, and toxicity falling under a limit, are examples of design constraints. Toxicogenomics integration and potential in the development of vaccine are the next boundaries for de-novo ML-enabled drug design, that are viewed as another front for this critical field^[2].

2.3. Automated synthesis planning

In recent years, computer-assisted synthesis has gotten a lot of coverage. Due to the massive dimensionality of reaction and chemical space, it is a difficult subject in itself. Whenever the objective is to propose synthesis which can be done in constant flow, this becomes even more difficult. While constant flow has several potential advantages, not all reactions are suitable for constant operation. Three machine learning models were developed in this study to determine whether the constant operation leads to the benefit of a given reaction, what success probability is there in constant flow for a specific series of components for reaction like reagents, reactants, catalysts, products and solvents, and what alternate reaction components should be used when the probability of success is small. Without depending on potentially ambiguously specified reaction models, a reaction template's abstract version is used by the first model obtained for measuring its relative increment in the publishing frequency through gaussian mixture modelling in constant flow. The second model is a 75 percent effective artificial-neural-network which classifies feasibility and in-feasibility of reaction elements. If a set of reaction components is explicitly referenced in the database as being used in continuous synthesis, it is considered feasible; otherwise, it is considered infeasible. While the neural network classifies some scenarios as feasible which by this criterion are infeasible. Further investigation reveals that several of these cases are at least possible – they have simply not been tested to disprove this. With a top-1 accuracy of 95%, the final model proposes alternative constant flow components. They provide a black-box estimation that if a reaction and a series of reaction components are suitable for consistent syntheses when used together^[7].

3. EXTENSIVE APPLICATIONS OF AI IN DRUG DISCOVERY

3.1. Selection of drug target and its validation

Identifying target is the process for determining the purpose for a potential drug target (molecule's proteins and genes) and the position in an illness in order to find a drug's efficacy target. Genomics such as structural and functional, testing cell based (in-vitro) and animal (in-vivo) assays, proteomics are all required for this. To predict therapeutic potential, the Drug Information Bank (which includes interaction of protein-protein, records of clinical data, expressions of genes and appropriate candidate for drug) is being analyzed out of a library with the help of AI. For instance, on 'genome-wide drugs, protein interaction network and their targets details,' implementing binary classification and relief algorithm from

Xgboost algorithm, or feature engineering by deep auto-encoder to generate results towards possible targets to allow goal prioritization. Discrete chemicals can be coded as consistent latent vector space for drug target site identification, allowing for gradient-dependent tuning in chemical space and estimates using graph convolutional networks based upon binding affinity as well as other parameters. AI applications trains ML algorithms and computer vision using cryo-EM microscope dataset (2D-structure) in order to explain the accurate spatial 3D-structure for molecular and protein complexes. A number of desirable properties, including pharmacokinetics, pharmacology, and toxicity profile, must be considered when selecting drug candidates. Algorithm for AI drug development for 'Simplified molecular input line-entry scheme' string, which is a specification in the form of line notation for representing chemical species using short ASCII strings, can be effectively implemented using reinforcement learning. Molecular diagrams of different atom or bond weights, potential energy measurements, coulomb matrices, atomic coordinates in three dimensions, molecular bonds or fragments, and so on are all included.

3.2. Lead optimization and compound screening

Lead optimization and compound screening steps includes the identification of Leads following Hits, where candidates for drug are chosen using virtual screening, high throughput screening, and combinatorial chemistry. A compound database created by extracting large amounts of data from freely accessible chemogenomics databases, that includes abundance of chemical catalogs with structure details is known as Virtual Screening^[8]. To help medicinal chemists in identifying possible lead molecules quickly within thousands of compounds and speeding up the early stages of drug production, it utilizes algorithms such as k-Nearest Neighbors, Naive-Bayesian Classifiers, Random Forests, Artificial Neural Networks, and Support Vector Machines. To utilize AI for simplification of chemical syntheses with limited manual intervention, AI Retrosynthesis Pathway Prediction in planning chemical synthesis are ambitious proposals. In this situation, AI synthesized robots can be used. The 3N-MCTS AI platform, that integrates Monte-Carlo Tree search with 3 different deep neural-networks for organic computer-aided compound syntheses, will choose only reactions that are well-identified for the synthesis of target compounds by screening out the building blocks that are most effective. For cell target classification, the AI model must be trained to identify various cell types quickly and automatically. To minimize the dimensions of the extracted features, analysis of the principal component can be utilized^[12]. To distinguish different cell types, 'least-square support vector machine' which is an AI-driven technique can be prepared. 'Intelligent Image-Activated Cell Sorting' machines, that calculates mechanical, electrical, and optical cell attributes using AI-driven extremely complex algorithms of deep neural network, are useful during the process of cell sorting for accurately separating various types of cells in the test samples.

3.3. Nonclinical research and studies

Preclinical trials, also known as non-clinical studies, are laboratory experiments performed in-vitro and in-vivo to determine the effectiveness and safety profile of a new drug product. Unsupervised approach of clustering-based machine learning tools evaluates RNA sequencing technologies for determining 'molecular mechanism of action', which shortens the time to capture specific significant quantities of biological knowledge. Also, it contributes to the discovery of hundreds of formerly unknown connections among various cytokines and

stimuli influenced by them. Pharmacokinetics and pharmacodynamics are the two ML Modelling methods, used as predicting concentration of the dose (exposure) response relationship in in-vitro and preclinical PK studies. To find successful combinations of multi-drug with a limited test, models in machine learning are now getting utilized for forecasting the intended response of drug-doses. Deeptox Algorithm has already been tested towards environmental drugs and chemicals in the toxicology profile evaluation of a substance for more than twelve distinct toxicity effects in specifically configured tests which are very slow and tedious and costly task. As a result, it may add significant importance towards medication production by properly anticipating the toxic effects of the compound. 'In-Silico'^[6] methods which uses DL algorithms, uses transcriptomic data from different biological systems and conditions to predict pharmacological properties.

3.4. Clinical studies and drug trials

An AI method design for drug trials could be perfect in gene targets definition, detecting patient disorders, determining on and off targets, and anticipating the impact of molecular designs. In Phase II clinical trials, one such mobile application AI platform improved prescription adherence by 25% as compared to standard 'modified directly observed therapy'^[4]. A monitoring approach for clinical drug trials that satisfies procedural criteria while completely moving away from monitoring the source data in AI risk-based monitoring will dramatically enhance clinical trial behaviour in all phases. AI can be used in Phase II and III clinical trials to classify and anticipate disease biomarkers relevant to humans in order to pick and enroll particular target populations, increasing the success rate of clinical trials.

4. CONCLUSION

AI applications appear to be becoming more common in the discovery of drug and its design. With notable advancements in De-novo drug design, QSAR modelling, and automated synthesis planning, amongst other things, these methods gradually live up to certain expectations of the community. Although, if these methods will be helpful in quickly designing and synthesizing appropriate candidate for drug to the researchers is yet to be seen. As for the current situation, with all of the existing molecules of drug, testing a new compound of lead is thoroughly difficult. To investigate known and unknown side effects and complications, plenty of research is needed. That being said, after one such algorithm technique in AI is developed, it would play major role in hastening out efforts applied in the development of a drug.

ACKNOWLEDGEMENT

I would like to thank all the authors for their time and support in conducting this research.

REFERENCES

- [1] Paul D, Sanap G, Shenoy S, Kalyane D, Kalia K, Tekade RK. Artificial intelligence in drug discovery and development. *Drug Discov Today*. 2021;26(1):80-93.
- [2] Mouchlis VD, Afantitis A, Serra A, et al. Advances in de Novo Drug Design: From Conventional to Machine Learning Methods. *Int J Mol Sci*. 2021;22(4):1676.
- [3] Mayr, A. et al. (2016) DeepTox: toxicity prediction using deep learning. *Front. Environ. Sci.* 3, 80.
- [4] Bain, E.E. et al. (2017) Use of a novel artificial intelligence platform on mobile devices to assess dosing compliance in a Phase 2 clinical trial in subjects with schizophrenia. *JMIR Mhealth Uhealth* 5, e18 Aicure Clinical trials.
- [5] Yang X, Wang Y, Byrne R, et al. Concepts of artificial intelligence for computer-assisted drug discovery. *Chem. Rev.* 2019;119(18):10520–10594.
- [6] Winiwarter S, Ahlberg E, Watson E, et al. In silico ADME in drug design – enhancing the impact. *ADMET DMPK*. 2018;6(1):15–33.
- [7] Pieter P., Connor W., Gao H., Vermeire H., Maarten R., Stevens V., Kevin M., William H, Artificial Intelligence for Computer-Aided Synthesis In Flow: Analysis and Selection of Reaction Components, *Frontiers in Chemical Engineering*, 2673-2718.
- [8] Schneider, G. Virtual Screening: An Endless Staircase?. *Nat. Rev. Drug Discovery* 2010, 9, 273–276.
- [9] Mak K.-K., Pichika M.R. Artificial intelligence in drug development: present status and future prospects. *Drug Discovery Today*. 2019;24:773–780.
- [10] Vyas M. Artificial intelligence: the beginning of a new era in pharmacy profession. *Asian J. Pharm.* 2018;12:72–76.
- [11] Sellwood M.A. Artificial intelligence in drug discovery. *Fut. Sci.* 2018;10:2025–2028.
- [12] H.C. Stephen Chan, Hanbin Shan, Thamani Dahoun, Horst Vogel, Shuguang Yuan. "Advancing Drug Discovery via Artificial Intelligence", *Trends in Pharmacological Sciences*, 2019.
- [13] Debleena Paul, Gaurav Sanap, Snehal Shenoy, Dnyaneshwar Kalyane, Kiran Kalia, Rakesh K. Tekade. "Artificial intelligence in drug discovery and development", *Drug Discovery Today*, 2021.
- [14] Varnavas D. Mouchlis, Antreas Afantitis, Angela Serra, Michele Fratello et al. "Advances in de Novo Drug Design: From Conventional to Machine Learning Methods", *International Journal of Molecular Sciences*, 2021.