Int. J. Sci. R. Tech., 2025 2(10)
A Multidisciplinary peer-reviewed Journal
www.ijsrtjournal.com [ISSN: 2394-7063]

Original Article

Artificial Intelligence in Drug Discovery

Aadesh Varpe, Om Ugalmugale, Sayali Pagire*, Aditya Vighne

Pharmaceutics, SND college of pharmacy, Babhulgaon

ABSTRACT

Artificial intelligence (AI) is transforming drug development by improving precision, decreasing timelines and costs, and enabling AI-powered drug design. This paper examines current advances in deep generative models (DGMs) for de novo drug creation, including various techniques and their tremendous influence. It critically examines the issues that are inextricably linked to these technologies, providing methods to realize their full potential. It includes case studies of both triumphs and failures in moving medicines to clinical trials using AI. Finally, it presents a forwardlooking strategy for optimizing, DGMs in de novo drug design, resulting in speedier and more cost-effective drug development.

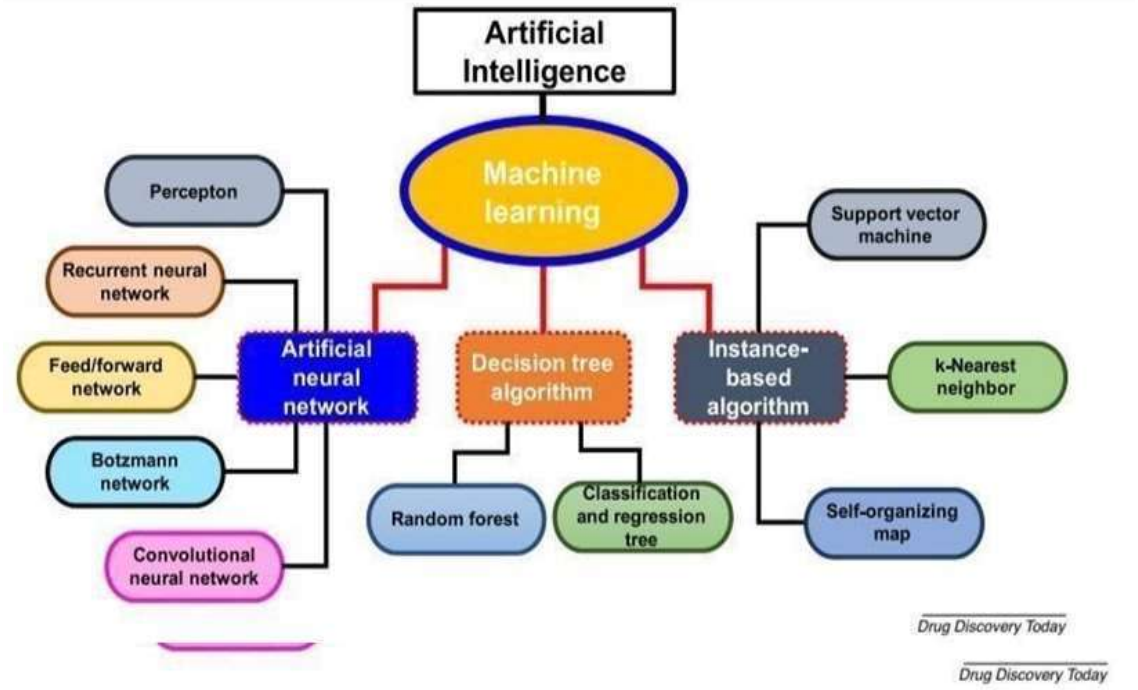
Keywords: Artificial intelligence (AI), Drug Discovery, improving precision, decreasing timelines and costs, and enabling AI-powered drug design

INTRODUCTION

Drug Discovery

In recent years, there has been a lot of interest in medicinal chemistry's application of artificial intelligence (AI) as a potential way to transform the pharmaceutical sector. [1] The process of finding and creating new drugs, or drug discovery, is a difficult and drawn-out undertaking that has historically relied on time-consuming methods like high-throughput screening and trial-and-error testing. However, by making it possible to analyze vast volumes of data more accurately and efficiently, artificial intelligence (AI) techniques like machine learning (ML) and natural language processing have the potential to speed up and enhance this process [2]. The scientists recently revealed the successful application of deep learning (DL) to accurately predict the potency of medicinal molecules. [3] . The toxicity of potential medications has also been predicted by AIbased techniques [4]. These and other studies have demonstrated AI's potential to increase the efficacy and efficiency of drug discovery procedures. But there are drawbacks and restrictions to using AI to create novel bioactive chemicals. To completely comprehend the benefits and limitations of AI in this field, more research is required, and ethical considerations must be taken into account. Notwithstanding these obstacles, it is anticipated that AI will play a major role in the creation of novel drugs and treatments during the coming years. [6].

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.



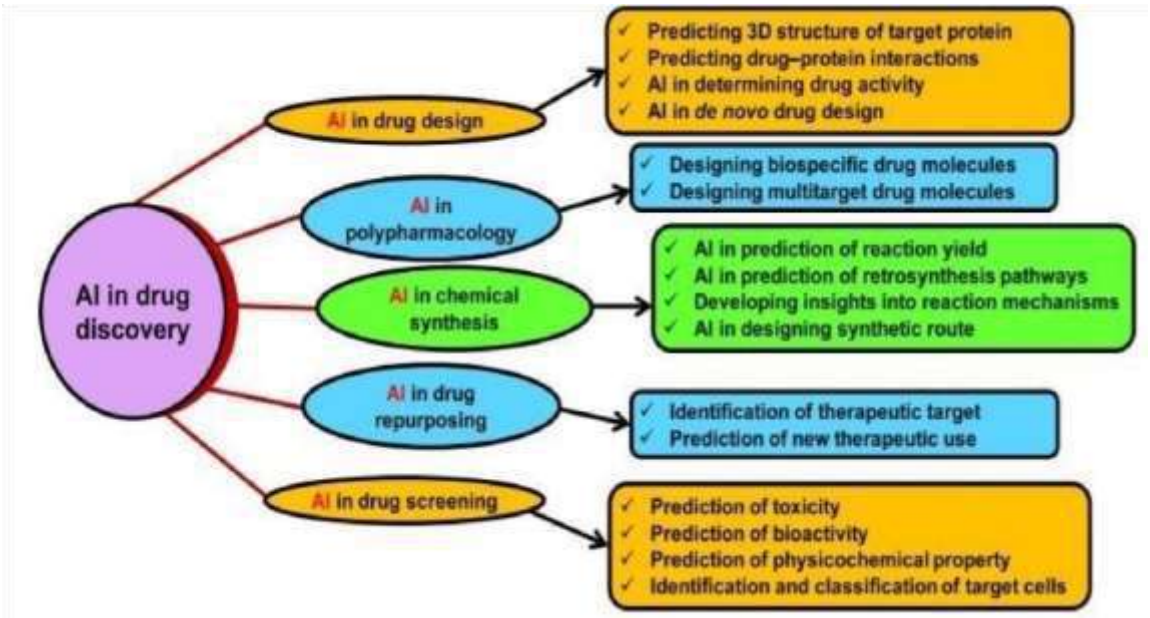
Drug Discovery in the AI Era

AI has been used extensively in the search for new drugs. Machine learning methods, such random forest (RF), have been used for VS and QSAR since the early 2000s. [7] The deep learning era began in 2012 with AlexNet41. Deep neural networks (DNN) beat the conventional RF model in predicting chemical activity shortly after in the 2012 Merck Kaggle competition. Deep learning in chemistry is a rapidly developing discipline that has been aided by the success of AI approaches in computer vision and natural language processing in recent years. 5. Researchers from InsilicoMedicine found powerful inhibitors of dis-coidin domain receptor 1 (DDR1) in 21 days in 2019[9]. In 2020, MIT researchers discovered halicin, a new antibiotic candidate that fights bacteria resistant to antibiotics [8]. 46 Keep in mind that AI can be used at various phases of drug discovery, from determining drug response to identifying and validating targets. This review focuses on lead identification, which entails two basic tasks: molecule creation and chemical property prediction. Predicting a molecule's property value based on its structure or learned representation is the foundation of molecular property prediction (VS). This can be used for a number of purposes, including toxicity prediction, druginduced liver injury (DILI) prediction, and drug-target interaction (DTI) prediction. Drug design is based on molecule generation, which entails two tiers of tasks: 1) Generating molecules in a realistic manner, that is, within the limitations set by chemical principles, and 2) goal directed molecule generation, i.e., generating chemically valid molecules with desired properties.

AI in drug discovery

More than 1060 molecules make up the enormous chemical space, which encourages the synthesis of numerous medicinal compounds. However, the medication development process is limited by a lack of sophisticated technologies, which makes it a costly and time-consuming operation that AI can help with. AI is able to identify hit and lead compounds, validate drug targets more quickly, and optimize drug structure design. Various uses of AI in drug discovery are illustrated. [10] The size, growth, diversity, and unpredictability of the data provide some serious data challenges for AI notwithstanding its benefits. Pharmaceutical businesses may have millions of molecules in their drug development data sets, which may be too large for typical machine learning methods to handle A computational model based on the quantitative structure-activity relationship (QSAR) may predict a large number of compounds or basic physicochemical characteristics, like log P or log D, in a short amount of time. These models, however, fall well short of forecasting intricate biological characteristics, such the effectiveness and side effects of substances. Small training sets, experimental data

errors in training sets, and a dearth of experimental validations are additional issues that QSAR-based models must deal with. Recent advances in AI techniques, including DL and pertinent modelling studies, can be used to address these issues by evaluating drug compounds' safety and effectiveness using big data modelling and analysis. [11In order to observe the benefits of DL in the pharmaceutical industry's drug discovery process, Merck sponsored a QSAR ML competition in 2012. For 15 drug candidate absorption, distribution, metabolism, excretion, and toxicity (ADMET) datasets, DL models demonstrated a considerable level of predictivity when compared to classic machine learning techniques. By showing the distributions of molecules and their characteristics, the vast virtual chemical space hints at a geo-graphical map of molecules. The purpose of the chemical space depiction is to gather positional data about molecules in the space in order to look for bioactive compounds; thus, virtual screening (VS) aids in the selection of suitable molecules for additional testing. A number of chemical spaces, such as PubChem, ChemBank, Drug Bank, and ChemDB, are publicly accessible. Together with structural and ligand-based approaches, a variety of in silico techniques for virtual screening compounds from virtual chemical spaces offer improved profile analysis, quicker removal of nonlead compounds, and therapeutic molecule selection at a lower cost. To choose a lead ingredient, drug design methods that take into account the physical, chemical, and toxicological profiles include coulomb matrices and molecular fingerprint identification. The intended chemical structure of a product can be predicted using a number of factors, including prediction models, molecular similarity, the molecule synthesis process, and the usage of in silico techniques. DeepVS, a novel method developed by Pereira et al. for the docking of 40 receptors and 2950 ligands, demonstrated remarkable performance when tested against 95,000 decoys. [12] Another method evaluated the form similarity, biochemical activity, and physicochemical characteristics of a cyclin-dependent kinase-2 inhibitor in order to optimize its potency profile using a multiobjective automated replacement algorithm. Potential drug candidates have been identified using QSAR modelling tools, which have developed into AI-based QSAR techniques like decision trees, random forest (RF), support vector machines (SVMs), and linear discriminant analysis (LDA), which can be used to expedite QSAR analysis. When comparing the ability of six AI systems to rank anonymous substances in terms of biological activity with that of conventional methods, King et al. discovered a small statistical difference.



Drug Design

AI greatly speeds up the drug development timeline in the field of drug design by improving the identification process of promising lead compounds. The process from concept to clinic is streamlined by AI's capacity to evaluate a broad range of molecular configurations and forecast their possible binding

affinities. [18] Finding tiny compounds that meet a number of essential requirements is the core of medication design. A favourable safety profile, appropriate chemical and biological features, pharmacological efficacy, and the innovation required to protect intellectual property rights for economic viability are some of these. [17] Traditional approaches have a number of difficulties, including lengthy input times, expensive computing costs, and inconsistent dependability, even if computational tools have transformed drug design and the approach to discovery. [13] AI stands out as a solution that can overcome these obstacles and improve the usefulness and efficiency of computational methods in drug development. [14] Because protein dysfunction is connected to many diseases, studying protein architecture is a crucial part of therapeutic creation. The goal of structural drug design is to find tiny compounds that have the ability to interact with protein targets in a specific way. Protein three-dimensional (3D) structure prediction has historically been expensive, time-consuming, and has had poor accuracy when done from scratch. The development of artificial intelligence, specifically deep learning and feature extraction technologies, has transformed this aspect of medication creation. These methods allow for the precise prediction of secondary protein structures and the mapping of protein interactions, which improves our understanding of the link between structure and sequence. [16] The ultimate goal is to use deep learning to predict 3D protein structures with greater precision, allowing for the investigation of protein-protein interactions (PPI) and furthering the science of structural drug design. [15] This incorporation of AI into drug design is a huge step forward, promising to increase the speed, cost-effectiveness, and success rate of drug development initiatives. Accurately predicting the three-dimensional (3D) structure of target proteins is an important step in structure-based drug design and discovery. [21] AI subsets such as machine learning and deep learning are crucial for tackling this dilemma. [20] AI-driven protein structure prediction relies on substantial sequence and structural data gathered from many sources. This dataset enables AI models to be trained to recognize complicated patterns that link amino acid sequences to their 3D structures. [19] AI models, particularly those based on deep learning, have demonstrated outstanding capacity to find complicated patterns in protein data by utilizing modern computational approaches. These models carefully extract information relating to amino acid properties, structural motifs, and evolutionary history, then use these insights to predict the 3D structure of proteins based on their sequences. AlphaFold, developed by Google DeepMind, is a breakthrough achievement in AI-driven protein structure prediction. To anticipate the 3D target protein structure, AlphaFold evaluates the lengths between nearby amino acids as well as the angles of peptide bonds. AlphaFold successfully predicted 25 out of 43 protein structures in a recent review, suggesting its potential in structure-based drug development. [20] Traditional approaches for determining protein structures, while accurate, are frequently resource heavy. AI provides a faster and more cost-effective alternative, generating trustworthy 3D structures from sequence data. [21] This advancement enables the design of medications that are specific to the structure of the target protein, allowing for earlier predictions of treatment efficacy and safety. Furthermore, AI techniques such as molecular dynamics (MD) simulations can use predicted 3D structures of proteins and drugs from databases such as the Protein Data Bank (PDB) and DrugBank to study the stability, dynamics, geometry, and binding efficacy of protein-drug complexes, providing valuable insights into their interactions over time. [22] AI has also demonstrated promise in modelling complicated relationships in biomedical data using graph machine learning techniques. These techniques, which portray chemical systems as graphs with atoms as fundamental units, might capture detailed patterns and interrelations between medications, diseases, PPI, and drug side effects, potentially aiding in therapeutic repurposing and response prediction.

AI in Polypharmacology

The landscape of drug discovery is experiencing a substantial upheaval, shifting away from the old "one drug, one target" paradigm and toward polypharmacology, a technique that investigates pharmacological interactions with several targets. This change is driven by the opportunity to improve therapy efficacy and more completely address the complexities of complicated diseases. AI is crucial to improvements in polypharmacology since it allows

for the study of vast biological datasets, revealing candidates with polypharmacological potential. Polypharmacology has gained popularity due to a better knowledge of disease causes and the molecular complexities involved. This evolution has been accelerated by the integration of large databases such as ZINC, PubChem, and DrugBank, among others. These resources combine massive volumes of data on molecular routes, binding affinities, and chemical characteristics, creating a rich tapestry for AI algorithms to explore and comprehend the complex linkages contained within. [28,29] The development of platforms such as DeepDDI, which aim to clarify drug-drug interactions and forecast alternate therapeutic uses with fewer side effects, demonstrates AI's impact. [27] Furthermore, AI's predictive powers extend to discovering off-target interactions, boosting our understanding of a drug's overall effects and opening the door for safer, more effective therapies. [24]. The polypharmacology paradigm shift has the potential to improve medication repurposing, predict off-target toxicity, and develop multitarget therapies rationally. Computational methodologies powered by AI have shown considerable promise in predicting polypharmacological profiles and enabling medication repurposing, which is the process of discovering new applications for previously approved pharmaceuticals. [26 Polypharmacology has been spurred by the discovery that targeting numerous nodes within complex biological networks may be more successful than targeting a single node, particularly for multifactorial disorders. This method considers elements of biological networks such as connectedness, redundancy, and pleiotropy, providing a more comprehensive view of drug discovery [25, 30]. Polypharmacology also has implications for prospective medication repurposing or re-profiling opportunities, which can drastically reduce drug development time and expense by using previously approved pharmaceuticals for new therapeutic indications. Successful examples of drug repositioning have been documented in the literature, and AI approaches can help find novel repurposing opportunities [26,30].

AI in Chemical Synthesis

The efficiency and sustainability of chemical synthesis are critical in the field of drug discovery. The introduction of AI has substantially altered this field, improving reaction times and predicting outcomes with amazing accuracy. The combination of AI technologies with chemical expertise enables the quick synthesis of complicated pharmacological compounds, extending the range of potential therapeutic discoveries [33]. The use of AI into chemical synthesis represents a significant step forward in drug research, increasing the efficiency and precision of synthetic processes. Several studies have emphasized AI's significant impact, particularly in expediting the identification of optimal reaction conditions and attaining error-free autonomous synthesis. This is accomplished by a mix of automation, real-time reaction monitoring, and artificial intelligence, which combined allow for a significant increase in the speed and reliability of the experimental workflow [31]. However, relying on automated systems and AI algorithms creates new issues, including the potential of oversimplifying the underlying intricacies of chemical interactions. Such simplifications can lead to mistakes in understanding and interpreting reaction dynamics, emphasizing the importance of carefully integrating AI tools with a solid understanding of chemical fundamentals. 7[32]

AI in Clinical Trial Design

The design of clinical trials, an important component in bringing new pharmaceuticals to market, includes establishing the number of events required to achieve statistically significant results. This stage is critical for estimating event rates within the target population, calculating patient recruitment numbers, and determining the follow-up time required to accrue the desired event count. Throughout the trial, patients are closely followed until a specific number of occurrences occur. Developing a novel medicine for the market is a time-consuming and resource-intensive procedure. To effectively navigate the drug development pipeline takes an average of 10 to 15 years and costs between USD 1.5 and 2.0 billion. [38] A considerable portion of this time and effort is committed to the clinical trial phases, which take about 6-7 years and need a significant financial investment. These clinical trials are critical for determining the safety and efficacy of a medicinal product in people for a specific illness condition. However, the success rate is frighteningly low, with only one out of every ten compounds entering clinical trials achieving regulatory clearance, resulting in a

substantial loss to the sector. 34]These failures can result from a variety of circumstances, including incorrect patient selection, a lack of technological needs, and insufficient infrastructure. [35] Preclinical activities, including as compound discovery, testing, and regulatory processes, account for 50% of R&D investment. Recruiting acceptable patients is crucial for clinical trial success, as it accounts for one-third of the trial duration. Inappropriate patient selection is responsible for around 86% of trial failures. These astonishing durations, financial burdens, and high failure rates highlight the critical need for new technologies that can streamline and improve the clinical trial process, lowering time-to-market and associated costs. With the vast digital medical data available, the implementation of AI has emerged as a promising solution, offering the potential to transform various aspects of clinical trial design and execution, ultimately accelerating the development and delivery of novel therapeutic interventions. [39] AI algorithms can quickly screen thousands of compounds by modelling interactions between drug molecules and biological targets, dramatically lowering the time and resources necessary for early-stage drug discovery. One critical part of drug discovery and biotechnology is the simulation of biomolecular structures utilizing physics-based atomic approaches such as molecular dynamics (MD). These simulations entail running MD simulations on 3D structures of proteins and medicines available from sources such as the Protein Data Bank (PDB) and DrugBank, as well as those predicted by powerful AI models like AlphaFold2[36]. This method examines the stability, dynamics, shape, and binding efficiency of protein-drug complexes, providing a time timeline of atomic movements. Advanced data analysis tools, such as deep learning, can then be used to examine these trajectories and obtain new insights into the structural changes and interactions occurring within complex biological systems. This knowledge can help to answer concerns about diseases, pathways, and drug response or resistance mechanisms. Atomwise, a company that specializes in AI-driven drug development, has used its AI platform to test a large number of tiny compounds against specific protein targets, revealing new therapeutic candidates. For example, they successfully uncovered possible Ebola remedies by virtual screening current pharmaceuticals, revealing two molecules that block the Ebola virus [37]. This approach speeds up early-stage drug discovery by expediting the identification of prospective candidates.

Challenges and Limitations of AI in Drug

Discovery

Despite AI's great promise to revolutionize the landscape of drug discovery, several serious barriers must be overcome before its full potential can be achieved. Securing data quality and accessibility presents a big problem. AI models are data-driven, and their efficacy is determined on the quantity and diversity of the data used to train them [46]. Acquiring highquality biological data is difficult due to privacy laws and data dispersion across multiple organizations. Furthermore, gathering the requisite data can be costly and time-consuming, particularly for small research teams. As a result, collaboration and data-sharing activities are critical to provide access to complete and diverse datasets. Data bias and generalizability are also major issues. When AI algorithms are trained on biased data, they may make false predictions. These biases can result from underrepresentation of various populations in clinical trials, geographical discrepancies in data sources, or differences between healthcare providers. Furthermore, overfitting, which occurs when a model performs well on training data but struggles on unseen data, can lead to the identification of inefficient medication candidates or false positives. Researchers can use bias correction approaches during the training phase of AI models to reduce the impact of biases on model outputs. For example, the SMOTE (Synthetic Minority Oversampling Technique) bias correction technique is used in an AI-powered drug discovery study to address data bias. SMOTE creates synthetic data points for underrepresented groups in the dataset, balancing it and reducing bias impact. Bias correction solutions are being investigated, however there is no general solution. Nonetheless, by using thorough dataset selection, processing, and bias correction approaches, researchers can reduce the influence of data bias in AI applications. Processing power and resource intensity are also important considerations, particularly for deep learning models. These models necessitate significant computational resources for both training and inference, which presents challenges for smaller pharmaceutical corporations and academic

research teams with restricted budgets. Cloud-based AI services and cooperation with AI technology providers are used to minimize computing costs and improve accessibility. Furthermore, regulatory approval and validation are crucial steps for AI models in drug research. Demonstrating the safety, effectiveness, and repeatability of AI-generated outcomes is critical for regulatory approval and developing trust in the pharmaceutical industry. Collaboration among regulatory bodies, pharmaceutical firms, and AI researchers is critical for developing validation protocols and standards. Cost concerns arise from the requirement for major early investments in technology, data collecting, and qualified workers. Addressing these financial issues would necessitate a long-term strategy that includes investigating government incentives, strategic alliances, and joint funding approaches.

Emergence of AI for Drug Discovery

The Knowledge Deficit

One of the most difficult issues that human investigators and AI systems encounter in drug development is managing vast amounts of diverse data of different quality. The rapid development of data and processing capacity has been cited as justification for a fourth paradigm, often known as data-driven scientific discovery. For the "why" and "what if" sorts of questions, relevant, preferably credible data must be located, inferred when absent, and connected using evidence-based reasoning, as depicted by the "connect the dots" metaphor. It is increasingly obvious that current drug discovery requires computer-based Artificial intelligence is defined as systems that can think intelligently and recognize patterns. These AI systems must be able to weigh data elements and collect examples of patterns in order to determine confidence and rationale. Automated systems that digest large sets of data using named entity recognition are an essential component of public domain databases, such as DISEASES for gene-disease associations, STRING for protein-protein interactions, and Open Targets and Pharos for complex disease-protein-drug annotations, to name a few examples. Together with AI-based protein structure prediction algorithms like Alpha Fold and RoseTTAFold, these resources have the potential to speed AI4DD.

Current e in Drug Discovery

An in-depth scientometric analysis of AI4DD revealed a significant increase in publications, from 49 in 2011 to 333 by 2020. The number of AI-powered drug discovery platforms is expected to expand in the near future. The pharmaceutical and biotech businesses, which move AI-driven drug discovery into commercial application, regularly collaborate with academic institutions, which often lead the development of algorithms and procedures. Over the last two decades, AI and machine learning have gone from being peripheral technologies to playing a major role in drug discovery. Today, we are closer than ever to accomplishing this long-awaited goal.

REFERENCE

1. Vamathevan, J., Clark, D., Czodrowski, P., et al. (2019). Applications of machine learning in drug discovery and development. Nature Reviews Drug Discovery, 18(6), 463–477. https://doi.org/10.1038/s41573-019-0024-5
2. Chen, H., Engkvist, O., Wang, Y., Olivecrona, M., & Blaschke, T. (2018). The rise of deep learning in drug discovery. Drug Discovery Today, 23(6), 1241–1250. https://doi.org/10.1016/j.drudis.2018.01.039
3. Zhavoronkov, A., Ivanenkov, Y. A., Aliper, A., et al. (2019). Deep learning enables rapid identification of potent DDR1 kinase inhibitors. Nature Biotechnology, 37(9), 1038–1040. https://doi.org/10.1038/s41587-019-0224-x
4. Mayr, A., Klambauer, G., Unterthiner, T., & Hochreiter, S. (2016). DeepTox: Toxicity prediction using deep learning. Frontiers in Environmental Science, 3, 80. https://doi.org/10.3389/fenvs.2015.00080
5. Amann, J., Blasimme, A., Vayena, E., Frey, D., & Madai, V. I. (2020). Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. BMC Medical Informatics and Decision Making, 20(1), 310. https://doi.org/10.1186/s12911-020-01332-6
6. Walters, W. P., & Murcko, M. A. (2020). Assessing the impact of generative AI on medicinal chemistry. Nature Biotechnology, 38, 143–145. https://doi.org/10.1038/s41587-019-0361-2

7. Chen, H., Engkvist, O., Wang, Y., Olivecrona, M., & Blaschke, T. (2018). The rise of deep learning in drug discovery. Drug Discovery Today, 23(6), 1241–1250. https://doi.org/10.1016/j.drudis.2018.01.039
8. Stokes, J. M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N. M., ... & Collins, J. J. (2020). A deep learning approach to antibiotic discovery. Cell, 180(4), 688–702.e13. https://doi.org/10.1016/j.cell.2020.01.021
9. Zhavoronkov, A., Ivanenkov, Y. A., Aliper, A., Veselov, M. S., Aladinskiy, V. A., Aladinskaya, A. V., ... & Aspuru-Guzik, A. (2019). Deep learning enables rapid identification of potent DDR1 kinase inhibitors. Nature Biotechnology, 37(9), 1038–1040. https://doi.org/10.1038/s41587-019-0224-x
10. Mak, K.-K., & Pichika, M. R. (2019). Artificial intelligence in drug development: Present status and future prospects. Drug Discovery Today, 24(3), 773–780. https://doi.org/10.1016/j.drudis.2018.11.014
11. Zhou, J., Wang, Q., Pan, S., & Du, X. (2020). Artificial intelligence in COVID-19 drug repurposing. The Lancet Digital Health, 2(12), e667–e676. https://doi.org/10.1016/S25897500(20)30223-4
12. Pereira, J. C., Caffarena, E. R., & dos Santos, C. N. (2016). Boosting dockingbased virtual screening with deep learning. Journal of Chemical Information and Modeling, 56(12), 2495– 2506. https://doi.org/10.1021/acs.jcim.6b00340
13. Chen, H., Engkvist, O., Wang, Y., Olivecrona, M., & Blaschke, T. (2018). The rise of deep learning in drug discovery. Drug Discovery Today, 23(6), 1241–1250. https://doi.org/10.1016/j.drudis.2018.01.039
14. Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Do, B. T., Way, G. P., ... & Greene, C. S. (2018). Opportunities and obstacles for deep learning in biology and medicine. Journal of the Royal Society Interface, 15(141), 20170387. https://doi.org/10.1098/rsif.2017.0387
15. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2
16. Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., ... & Kavukcuoglu, K. (2020). Improved protein structure prediction using potentials from deep learning. Nature, 577(7792), 706–710. https://doi.org/10.1038/s41586019-1923-7
17. Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., ... & Zhao, S. (2019). Applications of machine learning in drug discovery and development. Nature Reviews Drug Discovery, 18(6), 463–477. https://doi.org/10.1038/s41573-019-0024-5
18. Zhavoronkov, A., Ivanenkov, Y. A., Aliper, A., Veselov, M. S., Aladinskiy, V. A., Aladinskaya, A. V., ... & Aspuru-Guzik, A. (2019). Deep learning enables rapid
19. Callaway, E. (2020). 'It will change everything': DeepMind's AI makes gigantic leap in solving protein structures. Nature, 588(7837), 203–204. https://doi.org/10.1038/d41586020-03348-4
20. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2
21. Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., ... & Kavukcuoglu, K. (2020). Improved protein structure prediction using potentials from deep learning. Nature, 577(7792), 706–710. https://doi.org/10.1038/s41586019-1923-7
22. Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., ... & Hassabis, D. (2021). Highly accurate protein structure prediction for the human proteome. Nature, 596(7873), 590–596. https://doi.org/10.1038/s41586-021-03828-1
23. Anighoro, A., Bajorath, J., & Rastelli, G. (2014). Polypharmacology: Challenges and opportunities in drug discovery. Journal of Medicinal Chemistry, 57(19), 7874–7887. https://doi.org/10.1021/jm5006463
24. Bocci, G., Cassetta, L., & Gozzi, G. (2017). In silico prediction of off-targets for improved drug safety. Frontiers in Pharmacology, 8, 280. https://doi.org/10.3389/fphar.2017.00280

25. Hopkins, A. L. (2008). Network pharmacology: The next paradigm in drug discovery. Nature Chemical Biology, 4(11), 682–690. https://doi.org/10.1038/nchembio.118
26. Pushpakom, S., Iorio, F., Eyers, P. A., et al. (2019). Drug repurposing: Progress, challenges and recommendations. Nature Reviews Drug Discovery, 18(1), 41–58. https://doi.org/10.1038/nrd.2018.168
27. Ryu, J. Y., Kim, H. U., & Lee, S. Y. (2018). Deep learning improves prediction of drug– drug and drug–target interactions. Proceedings of the National Academy of Sciences, 115(18), E4304–E4311. https://doi.org/10.1073/pnas.1803294115
28. Sterling, T., & Irwin, J. J. (2015). ZINC 15 – Ligand discovery for everyone. Journal of Chemical Information and Modeling, 55(11), 2324–2337. https://doi.org/10.1021/acs.jcim.5b00559
29. Wishart, D. S., Feunang, Y. D., Guo, A. C., et al. (2018). DrugBank 5.0: A major update to the DrugBank database for 2018. Nucleic Acids Research, 46(D1), D1074–D1082. https://doi.org/10.1093/nar/gkx1037
30. Xiong, G., Luo, Y., Ji, W., et al. (2024). Artificial intelligence for polypharmacology and multi-target drug discovery. International Journal of Molecular Sciences, 25(14), 6996. https://doi.org/10.3390/ijms25146996
31. Coley, C. W., Eyke, N. S., & Jensen, K. F. (2019). Autonomous discovery in the chemical sciences part II: Outlook. Accounts of Chemical Research, 53(5), 895–906. https://doi.org/10.1021/acs.accounts.9b00640
32. Schwaller, P., Laino, T., Gaudin, T., Bolgar, P., Hunter, C. A., Bekas, C., & Lee, A. A. (2020). Machine intelligence for chemical reaction prediction. Chemical Science, 11(2), 331– 339. https://doi.org/10.1039/C9SC05704H
33. Segler, M. H. S., Preuss, M., & Waller, M. P. (2018). Planning chemical syntheses with deep neural networks and symbolic AI. Nature, 555(7698), 604–610. https://doi.org/10.1038/nature25978
34. DiMasi, J. A., Grabowski, H. G., & Hansen, R. W. (2016). Innovation in the pharmaceutical industry: New estimates of R&D costs. Journal of Health Economics, 47, 20,33. https://doi.org/10.1016/j.jhealeco.2016.01.012
35. Harrer, S., Shah, P., Antony, B., & Hu, J. (2019). Artificial intelligence for clinical trial design. Trends in Pharmacological Sciences, 40(8), 577–591. https://doi.org/10.1016/j.tips.2019.06.004
36. Jumper, J., Evans, R., Pritzel, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2
37. Wallach, I., Dzamba, M., & Heifets, A. (2015). AtomNet: A deep convolutional neural network for bioactivity prediction in structure-based drug discovery. arXiv preprint arXiv:1510.02855.
38. Wong, C. H., Siah, K. W., & Lo, A. W. (2019). Estimation of clinical trial success rates and related parameters. Biostatistics, 20(2), 273–286. https://doi.org/10.1093/biostatistics/kxx069
39. Yu, K. H., Beam, A. L., & Kohane, I. S. (2018). Artificial intelligence in healthcare. Nature Biomedical Engineering, 2(10), 719–731. https://doi.org/10.1038/s41551-018-0305-z
40. Baek, M., DiMaio, F., Anishchenko, I., et al. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. Science, 373(6557), 871–876. https://doi.org/10.1126/science.abj8754
41. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority oversampling technique. Journal of Artificial Intelligence Research, 16, 321–357. https://doi.org/10.1613/jair.953
42. Hey, T., Tansley, S., & Tolle, K. (2009). The Fourth Paradigm: Data-Intensive Scientific Discovery. Microsoft Research.
43. Mak, K. K., & Pichika, M. R. (2019). Artificial intelligence in drug development: Present status and future prospects. Drug Discovery Today, 24(3), 773–780. https://doi.org/10.1016/j.drudis.2018.11.014
44. Szklarczyk, D., Gable, A. L., Lyon, D., et al. (2021). STRING v11: Protein–protein association networks with increased coverage, supporting functional discovery in genome-wide datasets. Nucleic Acids Research, 47(D1), D607–D613. https://doi.org/10.1093/nar/gky1131
45. Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. Nature Medicine, 25(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

46. Vamathevan, J., Clark, D., Czodrowski, P., et al. (2019). Applications of machine learning in drug discovery and development. Nature Reviews Drug Discovery, 18(6), 463–477. https://doi.org/10.1038/s41573-019-0024-5
47. Zhavoronkov, A., Ivanenkov, Y. A., Aliper, A., et al. (2019). Deep learning enables rapid identification of potent DDR1 kinase inhibitors. Nature Biotechnology, 37(9), 1038–1040. https://doi.org/10.1038/s41587019-0224-x.

HOW TO CITE: Aadesh Varpe, Om Ugalmugale, Sayali Pagire*, Aditya Vighne, Artificial Intelligence in Drug Discovery, Int. J. Sci. R. Tech., 2025, 2 (10), 435-444. https://doi.org/10.5281/zenodo.17444647