STM JOURNALS

Emerging Trends in
Chemical Engineering

ISSN: 2349-4786
Volume 12, Issue 3, 2025
September–December
DOI (Journal): 10.37591/ETCE

https://journals.stmjournals.com/etce

Review ETCE

Application of Artificial Intelligence in Drug Discovery

Gaurang Rai*

Abstract

The application of artificial intelligence (AI) in medicine, especially through machine learning (ML), is revolutionizing new-age drug discovery research. AI is found as an efficient and powerful tool to narrow the gap between disease detection and developing and identifying potential therapeutic agents for a cure. This review provides a summary of the latest developments in AI and its potential application in drug discovery for untreatable diseases. The review also examines the various basic stages of the drug discovery process, like (a) disease identification, (b) disease proper diagnosis, (c) target identification, (d) screening, and (e) drug discovery. AI's application will greatly help to analyze huge datasets and discern patterns, which is essential in these early stages of drug discovery. Thus, enhancing accurate predictions and efficiencies in disease identification and most suitable drug discovery, followed by clinical trial management. In this review, the role of AI in expediting drug development work is highlighted, and its great potential to analyze huge historical data volumes, which will significantly reduce the time and costs required for new drug development and its market introduction. The need for data quality, algorithm training, and ethical considerations, especially in patient data handling during clinical trials during drug discovery development work using AI, is emphasized. Considering above-mentioned factors, AI seems to be capable of transforming drug development and assures significant benefits to patients and society globally, which is a great support to humanity.

Keywords: Artificial intelligence, machine learning, disease identification, drug design, drug discovery

INTRODUCTION

Historically, drug discovery has been found to be a long, complex, and costly process, and generally requires a decade or more of research and development and a large amount of CAPEX and OPEX in investment before developing and commercially launching a newly discovered drug in the market [1, 2]. This lengthy and circuitous cycle involves identifying several promising target molecules, screening thousands of compounds, preclinical and clinical trials, and lastly but most importantly, regulatory approvals. The high attrition and failure rate during drug discovery development, where many candidate molecules fail during trials, further makes the task more challenging.

Recently, artificial intelligence (AI) has emerged as a transformative and promising technology in this domain [3, 4]. By applying powerful mathematical and computational theoretical models, massive and diverse datasets compilation, and predictive analysis, AI has the potential to accelerate drug discovery work manifold, which will substantially reduce costs and improve success rates of drug development, and to help in launching personalized medicine as never before possible.

***Author for Correspondence**
Gaurang Rai
E-mail: gaurangrai29@gmail.com

Student, Department of Computer Science and Engineering (CSE), Vellore Institute of Technology, Bhopal University, Bhopal-Indore Highway, Kothrikalan, Sehore, Madhya Pradesh, India

Received Date: July 20, 2025
Accepted Date: August 31, 2025
Published Date: September 17, 2025

Citation: Gaurang Rai. Application of Artificial Intelligence in Drug Discovery. Emerging Trends in Chemical Engineering. 2025; 12(3): 22–29p.

© STM Journals 2025. All Rights Reserved

Typically, drug discovery work is a very complicated and challenging task that requires several years and around $1 to 2 billion for the commercial launch of an approved new drug [5]. The main reason for such a long time is mainly due to rising failure rates during R&D and also extended clinical trials duration [6]. It is concluded that despite considerable investment in resources, almost 90% of potential drug molecules fail even after they have passed through to the phase-I clinical trial [7].

With the help of AI, it is possible to perform computational screening of large sets of potential candidates and enhance the success rate of the identified compounds during clinical trials [8]. However, these methods may have limitations like inefficiency and inaccuracy [9]. To overcome these challenges, deep learning (DL) and ML algorithms can be deployed, which are a sub-branch of AI, to help identify the most suitable solutions [10]. These AI tools are able to predict macrosystem characteristics with high precision and accuracy, and in turn, low computational costs. Due to these reasons, there has been an increased application of AI algorithms in the drug discovery process by chemical and biological industries.

ML is employing algorithms such as DL, Bayesian network (BN), random forest (RF), clustering, and support vector machine (SVM) in the drug discovery process. The broad categorization of ML can be seen in Figure 1. DL models are used to process and analyze large sets of data in tasks, especially in clinical imaging [11, 12], virtual screening (VS) [13, 14], and bioactivity predictions [15]. BNs are used to precisely predict toxicity or bioactivity and subsequent patients' response to treatments [16]. RF models are helpful for the identification of target molecules and feature appropriate selection, and clustering identifies patterns or co-relationships within data [17]. SVM is a supervised learning algorithm used to segregate data into diverse categories, and applications like predicting pharmacokinetic properties, VS, and toxicity assessment [18].

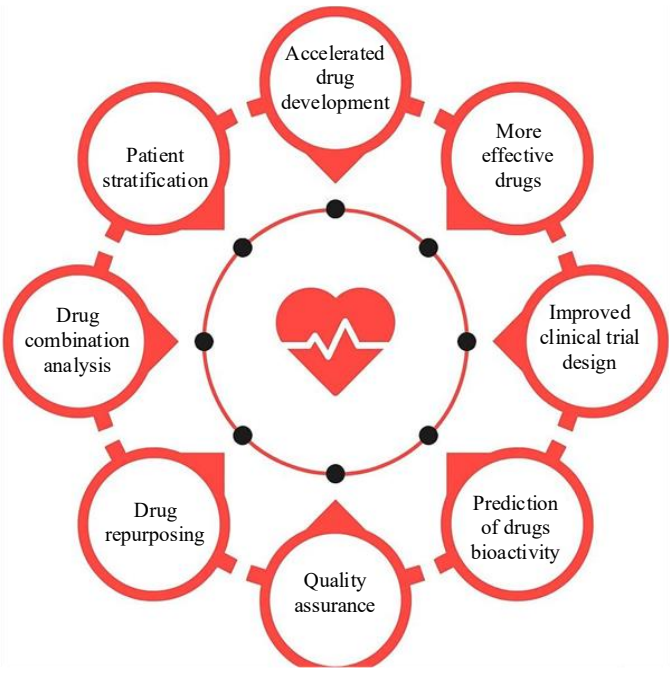

Figure 1. Possible application area of AI in drug discovery.

OVERVIEW OF AI IN DRUG DISCOVERY

The application of artificial intelligence (AI) in medicinal chemistry, i.e., drug discovery, is gaining significant momentum and recognition in recent times as a promising means of revolutionizing the pharmaceutical industry [19]. Drug discovery is defined as the process of identifying and developing new molecules. It is a complex and time-consuming process that traditionally depends on labor-intensive tasks, such as trial-and-error experiments and high-throughput screening. However, recently

© STM Journals 2025. All Rights Reserved

developed AI Algorithm techniques like machine learning (ML) and natural language processing can accelerate and improve this tedious process due to more efficient and accurate analysis of large sets of data [20]. Several reported studies describe the successful use of deep learning (DL) to predict the efficacy of drug compounds with very high accuracy [21]. AI-based models have also been successfully able to accurately predict the toxicity of drug candidates [22]. These and other research efforts have concluded the capability of AI to greatly improve the efficiency and effectiveness of time-consuming drug discovery processes. It may also be mentioned that the application of AI in discovering and developing new bioactive molecules is also full of challenges and limitations. Moreover, the following ethical considerations are a must in drug discovery work, and further research is required to completely understand the advantages and disadvantages of applying AI in these areas [23]. However, despite various challenges, AI is projected to significantly contribute to the rapid development of new drugs and therapies in the coming few years.

AI means the simulation of human intelligence in machines that are capable of learning, reasoning, problem-solving, and grasping. In the process of drug discovery, AI integrates machine learning (ML), deep learning, natural language processing (NLP), and predictive modeling to discover and analyze vast biological and chemical datasets.

Key Capabilities of AI in Drug Discovery

1. *Target identification:* The goal of this task is to clearly identify a biological target that is responsible for the disease. Traditionally, this process comprises of determination of biochemical and cell-based assays, and relevant animal studies.
2. *Lead discovery:* This involves identifying a molecule that has been found to have the ability to interact with the target and modify or alter its activity. Traditionally, this approach involves screening of large compound molecule libraries that already exist to find suitable active molecules, and this is a very time-consuming and laborious process.
3. *Lead optimization:* This process involves optimization of the pharmacological aspects of a lead molecule, like its potency, selectivity, and bioavailability. This task generally includes a combined approach based on medicinal chemistry, computational modelling, and *in vitro* testing.
4. *Preclinical testing:* involves testing the efficiency and safety of the lead molecules in animal models. This process is also very lengthy and often requires a huge capital investment.
5. *Clinical trials:* Clinical trials involve systematic examination of the safety and efficiency of lead molecules in human beings. This is the most important, expensive, and time-consuming stage of the drug discovery process. It requires extensive clinical trials over large samples to analyze and conclude the safety and efficacy of the drug in patients prior to its launch for regular use.

AI can significantly reduce the time of stages mentioned above and thus can improve the drug discovery process with respect to reducing costs, time, and increasing the probability of success at each stage of drug discovery.

Snapshot of Stages of New Drug Discovery

- Data mining from existing literature and biological databases.
- Predictive modeling (simulation) to assess molecular behavior and interactions.
- *De novo* drug design using AI generative models.
- Optimization of pharmacokinetics and toxicity of new molecules.
- Patient stratification and development of precision medicine applications.

APPLICATIONS OF AI ACROSS THE DRUG DISCOVERY PIPELINE

AI can significantly improve drug target prediction by studying diverse and vast biological datasets. AI models can identify potential targets more efficiently than traditional practices by combining genomics, proteomics, and other sources [24]. AI can analyze genomic data to find genetic variations

© STM Journals 2025. All Rights Reserved

related to diseases and identify genes and their encoded proteins as promising targets with high accuracy and precision [25]. Also, AI can assess proteomic data, such as protein structures and interactions, to identify proteins responsible for disease pathways and assess their druggability [26, 27]. Moreover, AI can combine multiple data sources like DrugBank [28], PubChem [29], Antibiotic Combination DataBase (ACDB) [30], Antibiotic Adjuvant DataBase (AADB) [31], as well as clinical examination data and electronic health datasets, to identify potential targets and predict their therapeutic capacity [32, 33]. ML algorithms, like DL and NLP, are very important for studying and examining complex datasets and finding patterns and relationships that may not be readily possible for human researchers [34]. ML models have proven to be powerful for elucidating gene-disease associations and searching for and identifying novel biomarkers. These models can assess complicated datasets, such as gene expression profiles, single-nucleotide polymorphisms (SNPs), and protein-protein interference networks, to find patterns and relationships that traditional statistical methods might not be able to find. Supervised learning algorithms, such as SVMs and random forests, can be trained efficiently on labelled gene expression and disease status data to predict disease risk and find genes associated with disease susceptibility [35, 36]. Modern Statistical learning models, such as clustering and dimensionality reduction techniques, can be used to identify groups of genes with similar expression patterns and uncover novel disease subtypes [37]. Furthermore, DL models such as recurrent neural networks (RNNs) and CNNs can examine and analyze complex genomic and proteomic data to identify intricate patterns and predict disease outcomes with high accuracy [38]. For instance, datasets comprising 10,000–15,000 entries have been used for target proteins such as Mpro (the main protease of SARS-CoV-2) in antiviral drug development and hERG (human Ether-à-go-go-Related Gene) in assessing cardiotoxic effects precisely and accurately [39, 40].

SUPPORTING TECHNOLOGIES AND TECHNIQUES

AI-Based Nanorobots for Drug Delivery

Nanorobots mainly consist of integrated circuits, sensors, a power supply, and secure backup of data, which are appropriately maintained with advanced computational technologies, such as AI [41, 42]. They are programmed and conditioned to avoid collisions, target identification, detect and attach, and lastly exit from the body. Advances in nano/microrobots have enabled the capability to navigate towards the targeted site assessed by physiological aspects, such as pH, therefore improving the accuracy and efficiency, as well as reducing adverse effects [42]. Development of implantable nanorobots for controlled delivery of drugs and genes requires certain crucial parameters, such as dose adjustment, sustained release, and controlled release. The release of the drugs also requires automation, which can now be done with the help of several AI available tools, i.e., NNs, fuzzy logic, and integrators [43]. Microchip implants are used for programmed release and to identify the exact location of the implant in the body.

AI in Combination with Drug Delivery and Synergism/Antagonism Prediction

Currently, many different combinations of drugs are approved and marketed to fight complex diseases, such as TB and cancer, because of the synergistic effectiveness of such combinations for quick recovery [44, 45]. For precise selection and potential most suitable drugs for combination often requires high-throughput screening of a sizable number of drugs, which makes the process very tedious; like cancer therapy requires six or seven drugs for the therapy. ANNs, logistic regression, and network-based modelling can be very effective for screening several drug combinations and improving and optimizing dose regimen [44, 46]. A quadratic phenotype optimization process for the determination of optimal combination therapy for the treatment of bortezomib-resistant multiple myeloma with 114 FDA-approved drugs. Recommendations based on the studies conducted using this model concluded that decitabine (Dec) and mitomycin C (MitoC) as the best two-drug combination and Dec, MitoC, and mechlorethamine as the superior three-drug combination for effective treatment [45].

Combination drug delivery can be proven to be more effective if backed up by AI data generated for the synergism or antagonism of drugs administered together. It is possible that the Master Regulator Inference Algorithm used ‘Mater regulator genes’ data to predict 56% synergism. Other AI-

© STM Journals 2025. All Rights Reserved

based methods, like Network-based Laplacian regularized least square synergistic drug combination and RF, can also be effectively employed for the same [46].

Case Studies

- *Insilico Medicine:* AI tools were applied to accelerate the traditional drug discovery pipeline, leading to the design of an anti-fibrosis drug in under 18 months [47]. Normally, this process takes several years, but Insilico's generative models and machine learning platforms optimized molecular design and target selection, drastically reducing development time and cost.
- *AlphaFold:* DeepMind's AlphaFold system solved the long-standing protein folding problem using advanced deep learning architectures [48]. By predicting protein 3D structures with near-laboratory accuracy in a matter of hours, AlphaFold transformed structural biology, enabling breakthroughs in understanding disease mechanisms and accelerating the development of new therapies.
- *Benevolent AI:* Leveraging natural language processing and machine learning, BenevolentAI identified new uses for existing drugs, including repurposing them for COVID-19 treatment [49]. This demonstrated how AI can rapidly sift through vast biomedical data to generate actionable hypotheses for urgent global health challenges.

Benefits of AI in Drug Discovery

- *Speed:* Significantly cuts down years of development time for new drugs.
- *Cost-efficiency:* Minimize the number of lab-based screening and failed trials.
- *Precision:* Enables precision medicine development through patient stratification techniques.
- *Scalability:* Can analyze millions of compounds or the health data of patients.
- *Innovation:* Enables breakthroughs in the discovery and application of non-obvious targets and novel molecules [50].

Challenges and Limitations

- *Data quality:* Incomplete or biased datasets may mislead models and lead to wrong results.
- *Interpretability:* Many models are still black boxes, as the interpretation of the data generated is difficult.
- *Regulatory barriers:* Validation step through actual trials is still required and mandatory.
- *Integration:* Application of AI tools requires multidisciplinary collaboration for effective and proper usage.
- *Ethical concerns:* It is also mandatory to ensure responsible AI use, mainly in patient data analysis.

Future Prospects

- *AI and quantum computing:* Could enhance molecular simulations.
- *AI in personalized medicine:* Tailoring treatments using comprehensive patient data.
- *Federated learning:* Enabling collaborative drug discovery while maintaining data privacy [51].

CONCLUSION

It can be concluded that the integration of AI into drug discovery can revolutionize biomedical research and pharmaceutical innovation. It can accelerate the pace of development work in target identification, designing novel compounds, and optimizing clinical trials. AI can transform every step of the pipeline in the development of drug discovery. How few challenges still remain and need to be sorted out by sustained R&D efforts, especially around data and transparency, as well as regulation. As new models are developed that are more accurate, datasets are more comprehensive, and technologies are more integrated, AI is poised to drastically reduce the time, cost, and risk associated with new drug development, which ultimately will lead to faster cures and more personalized treatments for human beings.

© STM Journals 2025. All Rights Reserved

REFERENCES

1. Vamathevan J, Clark D, Czodrowski P, et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. 2019; 18(6): 463–477.
2. Zhavoronkov A, Ivanenkov YA, Aliper A, et al. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. Nat Biotechnol. 2019; 37(9): 1038–1040.
3. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021; 596(7873): 583–589.
4. Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T. The rise of deep learning in drug discovery. Drug Discovery Today. 2018; 23(6): 1241–1250.
5. Hinkson IV, Madej B, Stahlberg EA. Accelerating therapeutics for opportunities in medicine: A paradigm shift in drug discovery Front Pharmacol. 2020; 11: 770.
6. Barker R. A flexible blueprint for the future of drug development. Lancet. 2010; 375(9712): 357–359.
7. Sun D, Gao W, Hu H, et al. Why 90% of clinical drug development fails and how to improve it? Acta Pharm Sin B. 2022; 12(7): 3049–3062.
8. Hassanzadeh P, Atyabi F, Dinarvand R. The significance of artificial intelligence in drug delivery system design. Adv Drug Deliv Rev. 2019; 151–152: 169–190.
9. Bolcer JD, Hermann RB. The development of computational chemistry in the United States. Rev Comput Chem. Vol. 5. Wiley; 1994; 1–63.
10. Bianconi F, Filippucci M. Digital Wood Design: Innovative Techniques of Representation in Architectural Design. Vol. 24. Cham: Springer; 2019.
11. Ding Y, Sohn JH, Kawczynski MG, et al. A deep learning model to predict a diagnosis of Alzheimer disease by using 18F-FDG PET of the brain. Radiology. 2019; 290(2): 456–464.
12. Narin A, Kaya C, Pamuk Z. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. Pattern Anal Appl. 2021; 24(3): 1207–1220.
13. Gentile F, Yaacoub JC, Gleave J, et al. Artificial intelligence–enabled virtual screening of ultra-large chemical libraries with deep docking. Nat Protoc. 2022; 17(3): 672–697.
14. Carpenter KA, Cohen DS, Jarrell JT, et al. Deep learning and virtual drug screening. Future Med Chem. 2018; 10(21): 2557–2567.
15. Bule M, Jalalimanesh N, Bayrami Z, et al. The rise of deep learning and transformations in bioactivity prediction power of molecular modeling tools. Chem Biol Drug Des. 2021; 98(5): 954–967.
16. Berzuini C, Bellazzi R, Quaglini S, et al. Bayesian networks for patient monitoring. Artif Intell Med. 1992; 4(3): 243–260.
17. Kapsiani S, Howlin BJ. Random forest classification for predicting lifespan-extending chemical compounds. Sci Rep. 2021; 11(1): 13812(13p).
18. Heikamp K, Bajorath J. Support vector machines for drug discovery. Expert Opin Drug Discov. 2014; 9(1): 93–104.
19. 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. doi: 10.1016/j.drudis.2020.10.010
20. Xu Y, Liu X, Cao X, Huang C, Liu E, Qian S, Liu X, Wu Y, Dong F, Qiu CW, et al. Artificial intelligence: A powerful paradigm for scientific research. Innovation. 2021; 2(4): 100179. doi: 10.1016/j.xinn.2021.100179.
21. Zhuang D, Ibrahim AK. Deep learning for drug discovery: A study of identifying high efficacy drug compounds using a cascade transfer learning approach. Appl Sci. 2021; 11(17): 7772. doi: 10.3390/app11177772
22. Pu L, Naderi M, Liu T, Wu HC, Mukhopadhyay S, Brylinski M. EToxPred: A machine learning-based approach to estimate the toxicity of drug candidates. BMC Pharmacol Toxicol. 2019; 20(1): 2. doi: 10.1186/s40360-018-0282-6.
23. Rees C. The ethics of artificial intelligence. In: Strous L, Johnson R, Grier DA, Swade D, editors. Unimagined Futures – ICT Opportunities and Challenges. Cham: Springer; 2020. (IFIP Advances

© STM Journals 2025. All Rights Reserved

in Information and Communication Technology; vol. 555). doi:10.1007/978-3-030-64246-4_5.
24. Abou Hajal A, Al Meslamani AZ. Insights into artificial intelligence utilisation in drug discovery. J Med Econ. 2024; 27(1): 304–8.
25. Hopkins AL, Groom CR. The druggable genome. Nat Rev Drug Discov. 2002; 1(9): 727–30.
26. Lin Q, Tam PK-H, Tang CS-M. Artificial intelligence-based approaches for the detection and prioritization of genomic mutations in congenital surgical diseases. Front Pediatr. 2023; 11: 1203289.
27. Overington JP, Al-Lazikani B, Hopkins AL. How many drug targets are there? Nat Rev Drug Discov. 2006; 5(12): 993–6.
28. Knox C, et al. DrugBank 6.0: the DrugBank Knowledgebase for 2024. Nucleic Acids Res. 2024; 52(D1): 1265–d1275.
29. Kim S, et al. PubChem 2025 update. Nucleic Acids Res. 2024; 53(D1): D1516–25.
30. Lv J, et al. ACDB: an antibiotic combination database. Front Pharmacol. 2022; 13: 869983.
31. Lv J, et al. AADB: a manually collected database for combinations of antibiotics with adjuvants. IEEE/ACM Trans Comput Biol Bioinf. 2023; 20(5): 2827–36.
32. Chen W, et al. Artificial intelligence for drug discovery: resources, methods, and applications. Mol Ther Nucleic Acids. 2023; 31: 691–702.
33. Qiu X, et al. Advances in AI for protein structure prediction: implications for cancer drug discovery and development. Biomolecules. 2024; 14(3): 339.
34. Serrano DR, et al. Artificial intelligence (AI) applications in drug discovery and drug delivery: revolutionizing personalized medicine. Pharmaceutics. 2024; 16(10): 1328.
35. Osama S, Shaban H, Ali AA. Gene reduction and machine learning algorithms for cancer classification based on microarray gene expression data: a comprehensive review. Expert Syst Appl. 2023; 213: 118946.
36. Statnikov A, et al. A comprehensive evaluation of multicategory classification methods for microarray gene expression cancer diagnosis. Bioinformatics. 2005; 21(5): 631–43.
37. Friedman J. The elements of statistical learning: data mining, inference, and prediction. New York, NY: Springer; 2009.
38. Ching T, et al. Opportunities and obstacles for deep learning in biology and medicine. J R Soc Interface. 2018; 15(141): 20170387.
39. Boby ML, et al. Open science discovery of potent noncovalent SARS-CoV-2 main protease inhibitors. Science. 2023; 382(6671): 201.
40. Shen M-Y, et al. A comprehensive support vector machine binary hERG classification model based on extensive but biased end point hERG data sets. Chem Res Toxicol. 2011; 24(6): 934–49.
41. Hassanzadeh P. The significance of artificial intelligence in drug delivery system design. Adv Drug Delivery Rev. 2019; 151–152: 169–190. doi: 10.1016/j.addr.2019.05.001
42. Luo M. Micro-/nanorobots at work in active drug delivery. Adv Funct Mater. 2018; 28(25): 1706100.
43. Fu J, Yan H. Controlled drug release by a nanorobot. Nat Biotechnol. 2012; 30(5): 407–408. doi: 10.1038/nbt.2206
44. Calzolari D. Search algorithms as a framework for the optimization of drug combinations. PLoS Comput Biol. 2008; 4(12): e1000249. doi: 10.1371/journal.pcbi.1000249.
45. Wilson B, KM. Artificial intelligence and related technologies enabled nanomedicine for advanced cancer treatment. Nanomedicine (Lond). 2020;15(5):433–5. doi:10.2217/nnm-2019-0366. PubMed: 31997697.
46. Tsigelny IF. Artificial intelligence in drug combination therapy. Brief Bioinform. 2019; 20(4): 1434–1448. doi: 10.1093/bib/bby004
47. Insilico. (2023 Jan 10). A Phase 1 Breakthrough in AI Drug Discovery. [Online]. https://insilico.com/blog/ipf-phase1#:~:text=The%20safety%20and%20PK%20data,in%20drug%20discovery%20and%20development.
48. AlphaFold Protein Structure Database. [Online]. https://alphafold.ebi.ac.uk/
49. Benevolent AI. (2022 Apr 25). AI-Enabled Drug Discovery. [Online].

© STM Journals 2025. All Rights Reserved

https://www.benevolent.com/news-and-media/blog-and-videos/benevolentai-ai-enabled-drug-discovery/
50. Chirag Bhardwaj. (2025 Jul 17). The Potential of AI in Drug Discovery and its Impact on Healthcare. [Online]. https://appinventiv.com/blog/ai-in-drug-discovery/
51. DrugBank Team. (2024 Apr 12). Unleashing AI in Drug Discovery: Prospects and Challenges. [Online]. https://blog.drugbank.com/unleashing-ai-in-drug-discovery-prospects-and-challenges/

© STM Journals 2025. All Rights Reserved