
A Comprehensive Review of Artificial Intelligence in Drug Discovery: Current Trends and Future Perspective

Manpreet Kaur and Printy Dhadwal*

University School of Pharmaceutical Sciences, Rayat Bahra University, Mohali- 140104, India

Abstract: Artificial intelligence (AI) and machine learning (ML) have emerged as the breakthrough technologies most likely to revolutionize pharmaceutical research and development (R&D) in the last ten years. Revolutionary developments in computational technology and the concurrent removal of earlier restrictions on the gathering and processing of massive amounts of data are partly to blame for this. In the meantime, it's now too costly to introduce new medications to the markets and to patients. The pharmaceutical sectors are interested in AI/ML approaches because of their automated nature, predictive, capabilities, and the anticipated gain in efficiency that follows. Over the past 15 to 20 years, ML techniques have become more sophisticated in their applications to drug development. In the design, conduct and analysis of clinical trials, AI/ML are beginning to positively disrupt the drug development process. Because of the COVID 19 pandemic, clinical trials may use AI/ML more quickly because of the rising reliance on digital technologies. As we approach a future where AI/ML are increasingly being included into research and development, it is imperative that we move past associated jargon and commotion. This was also crucial to understand that when drawing conclusions on data, the scientific processes are still relevant, by doing this, it will become easier to distinguish reality from fiction and enable well informed choices regarding the best application of AI/ML in drug development. This study aims to demystify key concepts, present application instances, and offer a fair assessment of how AI/ML approaches should be applied in research and development.

Keywords: Machine learning, artificial intelligence, deep learning, drug repurposing, virtual screening

*Corresponding author: Printy Dhadwal

e-mail: Printy.18210@rayatbahrauniversity.edu.in

1. Introduction

Drug discovery is the sophisticated process of inventing, finding and producing new pharmaceuticals to improve human health and combat disease. This meticulous process, which is essential for inducing potent drugs into clinical practice, consists of several phases, from target identification and lead compound discovery to optimization, exacting clinical trials, and stringent preclinical testing [1]. Artificial intelligence has numerous applications including language modelling and pharmaceutical industry enhancement. AI speeds up and reduces the cost of drug discovery and development. Professionals in medication development face the challenge of developing an efficient and low risk approach for delivering therapeutic compounds to their intended target. Furthermore developing new pharmaceutical molecules is costly and time consuming. Artificial intelligence (AI) helps to reduce the cost of drug discovery and development [2]. Artificial intelligence (AI) is being used more and more, and this will probably alter the way clinical evaluation and training are conducted [3]. The success rate of lead compounds in clinical trials has been increased by the use extensive computer screening and docking [4]. AI is beginning to show itself as a trans formative force, accelerating processes, increasing accuracy, and opening doors for formerly unseen

discoveries in fields ranging from the creation of new drug targets to the complicated web of clinical trials [5] In the pharmaceutical sector, a large portion of what is referred to as artificial intelligence (AI) is more closely related to machine learning, which is defined by an algorithmic process in computers deliver better feedback, “An application of artificial intelligence(AI) that gives systems the ability to automatically learn and improve from experience without being explicitly programmed” is how machine learning is defined [6]. By integrating the developments in machine learning (ML) in a highly standardized and automated manner, artificial intelligence (AI) uses computer software applications that analyze, learn, and uncover large data related to pharmaceuticals in order to discover new drug compounds [7] Machine learning (ML) approaches are emerging as practical instruments that can support the conventional drug development process. The incorporation of algorithmic techniques into the preclinical stages of drug discovery is covered in this perspective. In particular, we highlight a variety of machine learning (ML) based initiatives in the several disease domains to speed up the finding of early hits, the clarification of mechanism of action (MOA), and the optimization of chemical properties [8]. Identification and validation of chemical compounds, target identification,

peptide synthesis , assessment of drug toxicity and physiochemical properties ,drug monitoring, drug efficacy and effectiveness and drug repositioning are all made possible by computational modelling grounded in AI and ML principles [9]. AI has broader applicability across multiple phases, its influence in the pharmaceutical industry cannot be overlooked. AI has a clear impact on pharmaceutical products at every level, from product

management to medication department. ML, deep learning, AI based quantitative structure - activity relationship (QSRL)Technologies, QSLRML, virtual screening (VS), support vector machines (SVMs), deep virtual screening, deep neural network (DNNs), recurrent neural network (RNNs), and other algorithms are among the AI technologies utilized in drug discovery for drug screening and drug design [10].

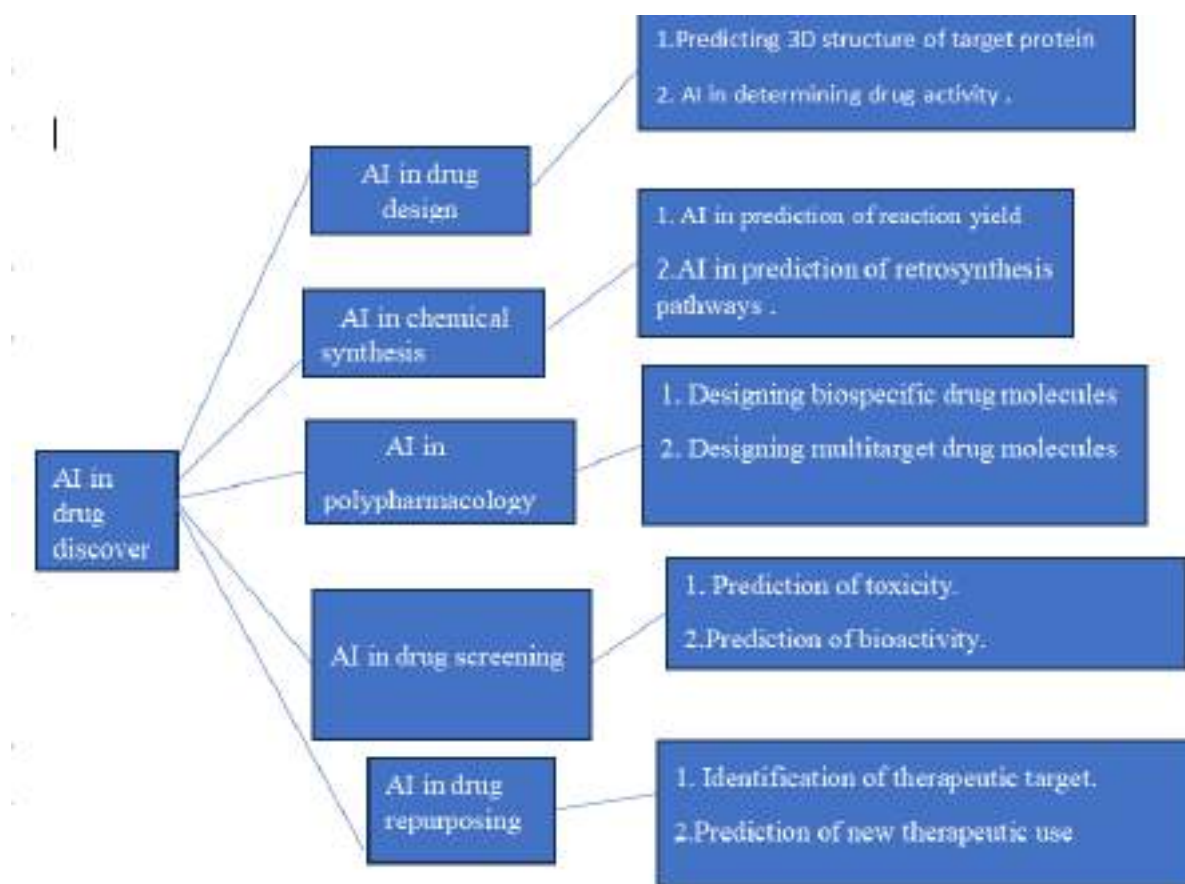


Fig. 1: Role of Artificial Intelligence in Drug Discovery [11]

2. History

The history of artificial intelligence in drug discovery is a story of steady technological

development, scientific investigation, and the changing interaction of biology, chemistry and computing. In 1950s-1960s Early AI research concentrated on fundamental ideas like

algorithms, problem solving, and fundamental machine learning method; yet, at this time, AI was not directly used in drug discovery. The majority of this time was theoretical, setting the stage for further developments in AI. In 1970s-1980s Rule based AI programs known as expert system were investigated for use in diagnosis and other medical application. Still, the process of finding new drugs was still mostly empirical and manual. The potential of AI in drug development and biological systems was acknowledged but has not yet been put into practice. [12]. In 1990s The use of machine learning techniques to forecast molecular attributes and model chemistry structure began. The biological activity of substances was predicted during this time using artificial intelligence (AI) in the form of neural network and decision trees. These techniques were first used in computational chemistry software, such as molecular docking programs, for virtual screening and molecular interaction predictions. Late 1990s: It combined high-throughput screening methods with machine learning. In order to find possible therapeutic candidates, researchers started using AI algorithms to recognize similarities in chemical databases [13]. In early 2000s Bioinformatics became a crucial field when the human genome project was finished. AI algorithms, including classification and clustering method, started to be used on biological datasets. Thanks to developments in AI, it is possible to predict

protein architectures, protein-ligand interactions, and therapeutic efficacy [14]. In 2000s AI contribution to drug target validation and identification has grown. Artificial intelligence started to study biological networks and forecast how chemicals and proteins will interact. Predicting negative drug reactions and finding biological pathways associated with particular diseases were two other benefits of AI [15]. In 2010s A subset of machine learning called deep learning started to transform the drug discovery process. The ability of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to predict chemical properties and create drugs has been demonstrated. There are now greater prospects for AI applications due to the growing availability of big datasets in chemical libraries, proteomics, and genomics [16]. In 2015-2020: companies such as Atomwise, Benevolent AI, Insilico Medicine started utilizing deep learning to forecast protein folding, find novel chemicals, and enhance the first phases of drug discovery. These businesses showed how AI may cut down on the time and expense involved in conventional drug discovery procedures [17]. In 2020s In precision medicine, AI is essential for creating medicine based on each patient's unique genetic profile. using genetic data and medical history, AI algorithms are now being utilized to forecast how patients would react to particular treatments. AI is also making clinical trial design and administration

easier, assisting in the identification of the best patient groups, lowering trial expenses and increasing success rates [18]. Particularly in reaction to the COVID -19 Pandemic, AI models were used to discover novel applications for already approved medications. AI systems showed the value of AI in speedy drug discovery by swiftly analyzing data to recommend current treatments for COVID-19 [19].

Table 1: History of artificial intelligence in drug discovery

Year	Intervention
1950	The Alan Turing study “Computing machinery and intelligence” suggested a “Universal machine with intelligent behavior [20]
1956	The Dartmouth conference brought together top scientists to debate AI research and set its aims, launching the discipline [21]
1957	Frank Rosenblatt’s Perceptron presented a single layer neural network that could learn through Perceptron learning [22]
1957	The general problem solver (GPS) Program by Newell and Simon used symbolic reasoning to solve problems [23]
1972	The invention of the PROLOG Programming language by Alain Colmerauer and Philippe Roussel brought about important breakthrough in logic programming and knowledge representation [24]
1976	The MYCIN System created by EDWARD SHORTLIFFE, which used a rule based method to show how expert systems can be used in medical diagnosis [25]
1982	The R1/XCON System by Douglas Lenat and Randal Davis was a significant expert system used in configuring computer systems, exhibiting the strength of rule -based reasoning [26]
1983	John McCarthy originated the term artificial intelligence and developed the Lisp Programming language, which became an important language for AI research [27]
1983	The CYC project, led by Douglas Lenat, aims to construct a massive knowledge base encompassing common sense thinking and understanding [28]
1988	The backpropagation algorithm, proposed by PAUL WEBBOS, allowed efficient training of multi layer neural network [29]
1997	Deep Blue Developed by IBM [30]
1998	The LeNet -5 architecture by Yann LeCun et al transformed the field of computer vision and became a basic model for image recognition Problems [31]
2009	The ImageNet project, led by Fei-Fei Li, presented a large scale dataset and benchmark for training deep convolutional neural network for image classification [32]
2011	Watson , an AI System designed by IBM [33]
2012	The ALEXNET architecture by KRIZHEVSKY , SUTSKEVER, and HINTON revolutionized Picture classification jobs and revealed the power of deep learning on GPUs [34]
2016	AlphaGo, created by DeepMind, Highlighting the advancements in machine learning and reinforcement learning [35]
2017	DeepMind’s AlphaGo Zero surpassed the performance of the original AlphaGo [36]

2018	DEVIN ET AL Created the BERT (Bidirectional Encoder Representations from Transformers) getting phenomenal results in processing in multiple natural languages that included tasks, including question answering and sentiment analysis [37]
2019	Open AI's Alpha Star upset top human players in StarCraft, Illustrating the potential of reinforcement learning in complex real time strategy games [38]
2021	AlphaFold wins the CASP Competition for guessing the 3D Protein structure [39]
2022	Organizations and projects such as the partnership on AI, OpenAI's Charter, the IEEE Global initiative on Ethics of Autonomous and Intelligent Systems have formed to address these concerns [40]
2023	ChatGPT arrives , creating great controversy [41]

3. Techniques

3.1 Supervised Learning

In the machine learning technique known as supervised learning, algorithms are trained on labelled data -that is, input data that has correspondingly accurate outputs. It is a technique to solving particular kinds of machine learning issues rather than a methodology. when the project's goal is to predict or classify using labelled data, supervised learning is essential. It can be integrated into a variety of machine learning approaches or workflows [42]. This includes instruction that model with labelled data. Unseen cases can then be classified or predicted using the learnt model. Regression and classification are two more subcategories of SL. Classification predicts a categorical label, like whether a patient is ill or healthy, using algorithms. After being a collection of labelled data, the algorithms learn to assign new data to one of the pre-established categories. For instance, a classifier may correctly divide patients into the sick and healthy classes after to

distinguish between them from the collection of medical records. Regression on the other hand, uses an algorithm to learn how to forecast the numerical value, like home prices. After learning from a dataset of labelled samples, the model produces a function that can forecast the numerical value of instances that are not visible [43].

3.2 Unsupervised Learning

Self - organizing neural networks use an unsupervised learning algorithm to find hidden patterns in unlabeled input data. Unsupervised means that the network can learn and organize information without being given an error signal to test possible solution. In supervised learning, the lack of direction for the algorithms can sometimes be good thing because it lets the algorithms look back for patterns that haven't been thought of before [44]. In general, unsupervised learning applications in DL are still uncommon nowadays. Unsupervised architectures, on the other hand, can be unexpectedly often observed in drug design,

demonstrating the method's effectiveness and the necessity of processing substantial molecular data in situations where measurable attributes are unavailable. specifically, the early neural

network protoplasts can still be identified in unsupervised learning applications for mapping molecular representations [45].

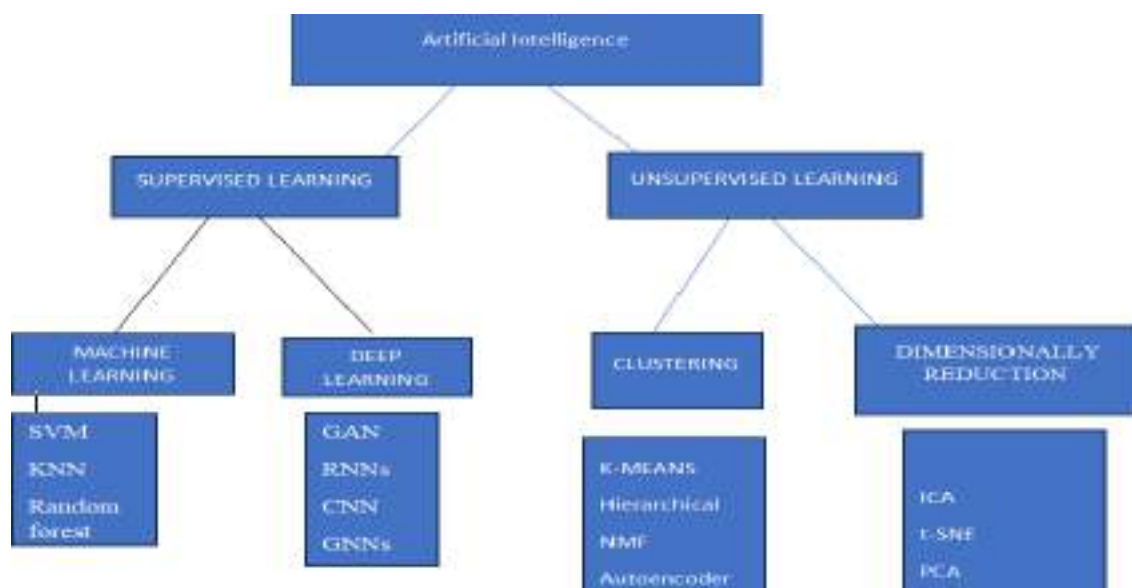


Fig 2.: Techniques of artificial intelligence in drug discovery [46]

3.2 Machine Learning

3.2.1 SVM (support vector machine)

Make predictions and models for excipient composition, processing parameters, and medication release profiles. SVM helps in predicting whether the compound is active or inactive against biological target by analyzing [47].

3.2.2 KNN (K-nearest neighbor algorithm)

Compound classification, pharmacokinetics modelling, toxicity prediction, formulation

optimization, patient satisfaction. KNN helps in identifying new drug molecules by comparing them to known drug molecules [48].

3.2.3 Random Forest

Prediction of toxicity and drug -drug interaction, drug discovery and design . Predict biological activity based on chemical structure and improve safety profile in early stage of development [49].

3.3 Deep Learning

3.3.1 GAN (Generative adversarial network)

Adverse event prediction, dosage form optimization, and generation of optimization of drug candidates [50].

RNNs (Recurrent neural network) Predicting protein / Peptide structures, Analyzing genomic data. Useful for generative models where drug like molecules is important [51].

3.3.2 GNNs (Graph neural network)

Predicting molecular properties, pharmacokinetics properties, modeling molecular relationships [52].

3.3.3 CNN (Convolutional neural network)

Prediction of toxicity and bioactivity, Identifying potential drug targets and analyzing molecular properties [53].

3.4 Clustering

3.4.1 Autoencoder

Compound screening, de novo drug design , toxicity prediction ,virtual screening .[54]

3.4.2.NMF (Non-negative Matrix Factorization)

Drug discovery and repurposing, chemical compound analysis, image analysis, pharmacokinetic modelling [55].

3.4.3 Hierarchical

Drug formulation optimization, pharmacovigilance, target identification [56].

3.4.4 K-Means

Product optimization , chemical similarity [57].

3.5 Dimensionally Reduction

3.5.1 PCA (Principal component Analysis)

Quality control analysis, facilitating formulation optimization [58].

3.5.2 T-SNE (t-distributed stochastic neighbor embedding)

Gene expression patterns, visualize molecular structures, or providing visual representation of formulation similarities [59].

3.5.3 ICA (Independent component analysis)

Used in brain imaging data and gene expression data or other types of biological data to identify underlying independent component [60].

4. Recent Advances and Techniques

4.1. Protein Structure Prediction

Techniques: Deep learning models like ALPHAFOLD.

Impact: Accurate prediction of 3D structures from amino acid sequences, accelerating drug target identification. Rapid and accurate protein structures prediction [61].

4.2 Generative Models for Molecule Design

Technique: Variational Autoencoders (VAEs), Reinforcement learning, Generative adversarial Networks (GANs) [62].

Impact: Novel drug like molecules with specific properties can design with AI Models. VAE can generate chemically diverse compound.

4.3 AI POWERED VIRTUAL SCREENING

Technique: Deep docking and machine learning algorithms

Impact: To identify potential drug candidates rapid screening of compounds is done. Predict binding affinity and helps in Structural activity relationship [63].

4.4 Natural Language Processing (NLP) In Drug Discovery

Technique: Transformers (BIOBERT, SCIBERT)

Impact: Automatic retrieval of insights and drug target interactions from biological literature [64].

4.5 Predictive Modeling for Admet Properties

Technique: Machine learning models trained on toxicological and pharmacokinetic data [65].

Impact: Helps predict absorption, distribution, metabolism, excretion and toxicity in drug design.

4.6 AI in Drug Repurposing

Technique: Graph neural network (GNN), Deep learning model [66].

Impact: Finding current medications that may be effective against novel targets like COVID-19. Helps to reduce development time and help in high success rate.

5. Challenges and Limitations

Even though drug research has benefited greatly from artificial intelligence (AI), there are still a number of issues and restrictions that need to be resolved.

5.1 Data Quality and Availability

Limited datasets: Training strong AI models is hampered by the lack of high quality , large scale datasets [68].

Data heterogeneity: It can be difficult to integrate data from many sources and formats [69].

5.2 Algorithmic Limitations

Interpretability: AI models can be hard to understand, which make it hard to comprehend how decisions are made [70].

Overfitting: AI models may perform poorly on unknown data if they overfit the training data [71].

5.3 Regulatory and Ethical Concerns

Bias and fairness: It is important to be fair and transparent because AI models might reinforce inherent biases [72].

5.4 Computational Resources

Significant computational resources are needed to train large AI models, which may be prohibitive for smaller organizations [73].

5.5 Domain Expertise

Lack of domain knowledge: AI models might not accurately represent the minute details of biological systems, requiring a high level of domain expertise to evaluate the results [74].

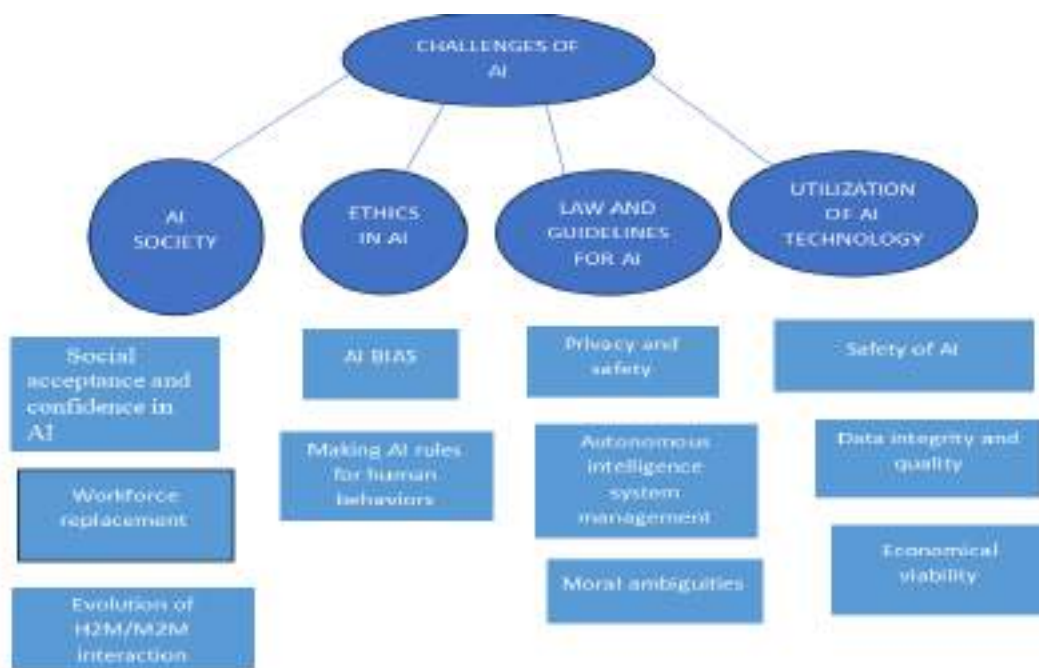


Fig. 4.: Challenges and Limitations In AI [67]

6. Conclusion

Drug discovery has been transformed by artificial intelligence (AI), which has improved clinical trial efficiency, sped up the identification of possible drug candidates and optimized the drug development process, Machine learning and deep learning are two example of AI technologies that have shown great promise in understanding complex biological data, finding new compounds, of and predicting molecular properties. Large volumes

biomedical literature can be analyzed by AI, which can help find therapeutic targets that were previously missed. To sum up, artificial intelligence (AI) greatly lowers the time and expense involved in drug discovery, improves drug design precision, and enables personalized medicine. Although there are still obstacles to overcome, like the requirement for high quality data and the intricacy of biological systems, artificial intelligence (AI) has enormous potential to improve medicine in the future by

facilitating quicker and more efficient disease treatments. AI is anticipated that combining AI with conventional drug discovery techniques may accelerate medical advancements even further, resulting in more creative treatments and better patient results. While challenges including data quality, regulatory approval, and ethical

considerations still exist, the use of AI in drug research was developing and holds promise for more effective, individualized and targeted treatment options. AI has the potential to become a more significant part of pharmaceutical research and development as the technology advances.

[Alexandre Blanco-González](#)^{1,2,3}, [Alfonso Cabezón](#)^{1,2}, [Alejandro Seco-González](#)^{1,2}, [Daniel Conde-Torres](#)^{1,2}, [Paula Antelo-Riveiro](#)^{1,2}, [Ángel Piñeiro](#)², [Rebeca Garcia-Fandino](#)¹

Reference

- [1] M. K. G. Abbas, A. Rassam, F. Karamshahi, R. Abunora & M. Abouseada, (2024), The role of AI in drug discovery, *Chembiochem*, 25(14), e202300816.
- [2] A. Blanco-Gonzalez, A. Cabezón, A. Seco-Gonzalez, A. Conde-Torres, D. Antelo-Riveiro, P. Pineiro & R. Garcia-Fandino, (2023). The role of AI in drug discovery: challenges, opportunities, and strategies. *Pharmaceuticals*, 16(6), 891.
- [3] V. Patel, & M. Shah, (2022), Artificial intelligence and machine learning in drug discovery and development. *Intelligent Medicine*, 2(3), 134-140.
- [4] A. U. Rehman, Li, M. Wu, B. Y. Ali, Rasheed, S. Shaheen, & J. Zhang, (2024). Role of artificial intelligence in revolutionizing drug discovery. *Fundamental Research*.
- [5] B. Revanth, S. S. Asrar, B. Sapkota, K. Vandana, K. S. Reddy & B. P. Kumar, (2024). The role of artificial intelligence and machine learning in drug discovery and development. *Asian J. Adv. Res.*, 7(1), 133-140.
- [6] M. J. Lamberti, M. Wilkinson, B. A. Donzanti, G. E. Wohlhieter, S. Parikh, S. R. G. Wilkins & K. Getz, (2019). A study on the application and use of artificial intelligence to support drug development. *Clinical therapeutics*, 41(8), 1414-1426.
- [7] C. Sarkar, B. Das, V. S. Rawat, J. B. Wahlang, A. Nongpiur, I. Tiewsoh & H. T. Sony, (2023). Artificial intelligence and machine learning technology driven modern drug discovery and development. *International Journal of Molecular Sciences*, 24(3), 2026.

-
- [8] D. B. Catacutan, J. Alexander, A. Arnold & J. M. Stokes, (2024). Machine learning in preclinical drug discovery. *Nature Chemical Biology*, 20(8), 960-973.
- [9] R. Gupta, D. Srivastava, M. Sahu, S. Tiwari, R. K. Ambasta & P. Kumar, (2021), Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Molecular diversity*, 25(3), 1315-1360.
- [10] S. K. Bhattamisra, P. Banerjee, P. Gupta, J. Mayuren, S. Patra & M. Candasamy, (2023). Artificial intelligence in pharmaceutical and healthcare research. *Big Data and Cognitive Computing*, 7(1), 10.
- [11] M. K. Abbas, G. Rassam, A. Karamshahi, F. Abunora, & M. Abouseada, (2024). The role of AI in drug discovery. *Chembiochem*, 25(14), e202300816.
- [12] J. McCarthy, M. L. Minsky, N. Rochester, & C. E. Shannon, (2006). A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. *AI magazine*, 27(4), 12-12.
- [13] J. A. Keith, V. Vassilev-Galindo, B. Cheng, S. Chmiela, M. Gastegger, K. R. Muller & A. Tkatchenko, (2021). Combining machine learning and computational chemistry for predictive insights into chemical systems. *Chemical reviews*, 121(16), 9816-9872.
- [14] T. Hastie, R. Tibshirani, J. H. Friedman & J. H. Friedman, (2009), *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2, pp. 1-758). New York: springer.
- [15] D. B. Searls, (2000), Using bioinformatics in gene and drug discovery. *Drug Discovery Today*, 5(4), 135-143.
- [16] S. Khaire & P. Bhaladhare, (2022), A comprehensive review on machine learning and deep learning methods in drug discovery. *Int J Recent Innov Trends Comput Commun*, 10, 1-8.
- [17] L. K. Vora, A. D. Gholap, K. Jetha, R. R. S. Thakur, H. K. Solanki, & V. P. Chavda, (2023), Artificial intelligence in pharmaceutical technology and drug delivery design. *Pharmaceutics*, 15(7), 1916.
- [18] J. Lorkowski, O. Kolaszyńska, & M. Pokorski, (2021), Artificial intelligence and precision medicine: A perspective. In *Integrative Clinical Research* (pp. 1-11). Cham: Springer International Publishing.
- [19] Y. Zhou, F. Wang, J. Tang, R. Nussinov, & F. Cheng, (2020), Artificial intelligence in COVID-19 drug repurposing, *The Lancet Digital Health*, 2(12), e667-e676.

-
- [20] A. M. Turing, (2007), Computing machinery and intelligence. In Parsing the Turing test: Philosophical and methodological issues in the quest for the thinking computer (pp. 23-65). Dordrecht: Springer Netherlands.
- [21] J. McCarthy, M. L. Minsky, N. Rochester & C. E. Shannon, (2006), A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. AI magazine, 27(4), 12-12.
- [22] F. Rosenblatt, (1958), The perceptron: a probabilistic model for information storage and organization in the brain. Psychological review, 65(6), 386.
- [23] L. Gugerty, (2006), Newell and Simon's logic theorist: Historical background and impact on cognitive modeling. In Proceedings of the human factors and ergonomics society annual meeting (Vol. 50, No. 9, pp. 880-884). Sage CA: Los Angeles, CA: SAGE Publications.
- [24] M. C. Dinu, C. Leoveanu-Condrei, M. Holzleitner, W. Zellinger & S. Hochreiter, (2024), Symbolicai: A framework for logic-based approaches combining generative models and solvers. arXiv preprint arXiv:2402.00854.
- [25] E. H. Shortliffe & B. G. Buchanan, (1975), A model of inexact reasoning in medicine. Mathematical biosciences, 23(3-4), 351-379.
- [26] R. Davis & D. B. Lenat, (1982). Knowledge-Based Systems in Artificial Intelligence: 2 Case Studies. McGraw-Hill, Inc.
- [27] J. McCarthy, P. W. Abrahams, D. J. Edwards, T. P. Hart & M. I. Levin, (1962), LISP 1.5 programmer's manual. MIT press.
- [28] D. B. Lenat, (1995), CYC: A large-scale investment in knowledge infrastructure. Communications of the ACM, 38(11), 33-38.
- [29] D. E. Rumelhart, G. E. Hinton & R. J. Williams, (1986), Learning representations by back-propagating errors. nature, 323(6088), 533-536.
- [30] M. Campbell & A. J. Hoane, Hsu Fh., (2002), Deep Blue. Artificial Intelligence, 134(1-2), 57-83.
- [31] Y. LeCun, L. Bottou, Y. Bengio & P. Haffner, (2002), Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.
- [32] J. Deng, W. Dong, R. Socher, L. J., Li & L. Fei-Fei, (2009), Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255).

-
- [33] D. Ferrucci, E. Brown, J. Chu-Carroll, J. Fan, D. Gondek, A. A. Kalyanpur & C. Welty, (2010), Building Watson: An overview of the DeepQA project. *AI magazine*, 31(3), 59-79.
- [34] A. Krizhevsky, I. Sutskever & G. E. Hinton, (2012), Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- [35] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche & D. Hassabis, (2016), Mastering the game of Go with deep neural networks and tree search. *nature*, 529(7587), 484-489.
- [36] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez & D. Hassabis, (2017), Mastering the game of go without human knowledge. *nature*, 550(7676), 354-359.
- [37] J. Devlin, M. Chang, K. W, Lee, & Toutanova, (2019), Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)* (pp. 4171-4186).
- [38] O. Vinyals, I. Babuschkin, W. M. Czarnecki, M. Mathieu, A. Dudzik, J. Chung, & D. Silver, (2019), Grandmaster level in StarCraft II using multi-agent reinforcement learning. *nature*, 575(7782), 350-354.
- [39] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger & D. Hassabis, (2021), Applying and improving AlphaFold at CASP14. *Proteins: Structure, Function, and Bioinformatics*, 89(12), 1711-1721.
- [40] S. K. Niazi, (2023), The coming of age of AI/ML in drug discovery, development, clinical testing, and manufacturing: the FDA perspectives. *Drug Design, Development and Therapy*, 2691-2725.
- [41] G. Obaido, I. D. Mienye, O. F. Egbelowo, I. D. Emmanuel, A. Ogunleye, B. Ogbuokiri & K. Aruleba, (2024), Supervised machine learning in drug discovery and development: Algorithms, applications, challenges, and prospects. *Machine Learning with Applications*, 17, 100576.
- [42] M. T. Almuqati, F. Sidi, S. N. Mohd Rum, M. Zolkepli & I. Ishak, (2024), Challenges in Supervised and Unsupervised Learning: A Comprehensive Overview. *International Journal on Advanced Science, Engineering & Information Technology*, 14(4).
- [43] R. Sathya & A. Abraham, (2013), Comparison of supervised and unsupervised learning algorithms for

- pattern classification. International Journal of Advanced Research in Artificial Intelligence, 2(2), 34-44.
- [44] J. Polanski, (2022), Unsupervised learning in drug design from self-organization to deep chemistry. International Journal of Molecular Sciences, 23(5), 2797.
- [45] L. K. Vora, A. D. Gholap, K. Jetha, R. R. S. Thakur, H. K. Solanki & V. P. Chavda, (2023), Artificial intelligence in pharmaceutical technology and drug delivery design. Pharmaceutics, 15(7), 1916.
- [46] L. K. Vora, A. D. Gholap, K. Jetha, R. R. S. Thakur, H. K. Solanki & V. P. Chavda, (2023), Artificial intelligence in pharmaceutical technology and drug delivery design. Pharmaceutics, 15(7), 1916.
- [47] K. Heikamp & J. Bajorath, (2014), Support vector machines for drug discovery. Expert opinion on drug discovery, 9(1), 93-104.
- [48] S. Dara, S. Dhamercherla, S. S. Jadav, C. M. Babu & M. J. Ahsan, (2022), Machine learning in drug discovery: a review. Artificial intelligence review, 55(3), 1947-1999.
- [49] S. Ahn, S. E. Lee & M. H. Kim, (2022), Random-forest model for drug-target interaction prediction via Kullback-Leibler divergence. Journal of Cheminformatics, 14(1), 67.
- [50] S. Tripathi, A. I. Augustin, A. Dunlop, R. Sukumaran, S. Dheer, A. Zavalny & E. Kim, (2022), Recent advances and application of generative adversarial networks in drug discovery, development, and targeting. Artificial Intelligence in the life Sciences, 2, 100045.
- [51] H. Askr, E. Elgeldawi, H. Aboul Ella, Y. A. Elshaier, M. M. Gomaa & A. E. Hassanien, (2023), Deep learning in drug discovery: an integrative review and future challenges. Artificial Intelligence Review, 56(7), 5975-6037.
- [52] Z. Zhang, L. Chen, F. Zhong, D. Wang, J. Jiang, S. Zhang & X. Li, (2022), Graph neural network approaches for drug-target interactions. Current Opinion in Structural Biology, 73, 102327.
- [53] C. Zhang, Y. Lu & T. Zang, (2022), CNN-DDI: a learning-based method for predicting drug-drug interactions using convolution neural networks. BMC bioinformatics, 23(Suppl 1), 88.
- [54] Q. Hu, M. Feng, L. Lai & J. Pei, (2018), Prediction of drug-likeness using deep autoencoder neural networks. Frontiers in genetics, 9, 585.
- [55] M. N. Wang, X. J. Xie, Z. H. You, D.W. Ding & L. Wong, (2022), A weighted non-negative matrix factorization

- p>
approach to predict potential associations between drug and disease.
- Journal of Translational Medicine*
- , 20(1), 552.
- [56] H. R. Teles, L. L. Ferreira, M. Valli, F. Coelho & A. D. Andricopulo, (2022), Hierarchical clustering and target-independent QSAR for antileishmanial oxazole and oxadiazole derivatives. *International Journal of Molecular Sciences*, 23(16), 8898.
- [57] J. Braun & D. Fayne, (2022), Mapping of Protein Binding Sites using clustering algorithms-Development of a pharmacophore based drug discovery tool. *Journal of Molecular Graphics and Modelling*, 115, 108228.
- [58] A. Giuliani, (2017), The application of principal component analysis to drug discovery and biomedical data. *Drug discovery today*, 22(7), 1069-1076.
- [59] P. Gao, J. Zhang, Y. Sun & J. Yu, (2020), Accurate predictions of aqueous solubility of drug molecules via the multilevel graph convolutional network (MGCN) and SchNet architectures. *Physical Chemistry Chemical Physics*, 22(41), 23766-23772.
- [60] S. L. Lin, (2021), The Application of machine learning ICA-VMD in an intelligent diagnosis system in a low SNR environment. *Sensors*, 21(24), 8344.
- [61] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger & D. Hassabis, (2021), Highly accurate protein structure prediction with AlphaFold. *nature*, 596(7873), 583-589.
- [62] A. Zhavoronkov, Y. A. Ivanenkov, A. Aliper, M. S. Veselov, V. A. Aladinskiy, A. V. Aladinskaya & A. Aspuru-Guzik, (2019), Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nature biotechnology*, 37(9), 1038-1040.
- [63] F. Gentile, J. C. Yaacoub, J. Gleave, M. Fernandez, A.T. Ton, F. Ban & A. Cherkasov, (2022), Artificial intelligence-enabled virtual screening of ultra-large chemical libraries with deep docking. *Nature Protocols*, 17(3), 672-697.
- [64] J. Lee, W. Yoon, S. Kim, D. Kim, S. Kim, C. H. So & J. Kang, (2020), BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, 36(4), 1234-1240.
- [65] K. Huang, T. Fu, W. Gao, Y. Zhao, Y. Roohani, J. Leskovec & M. Zitnik, (2021), Therapeutics data commons: Machine learning datasets and tasks for drug discovery and development. *arXiv preprint arXiv:2102.09548*.
- [66] S. Monteleone, T. F. Kellici, M. Southey, M. J. Bodkin & A. Heifetz, (2021),

- Fighting COVID-19 with artificial intelligence. In *Artificial Intelligence in Drug Design* (pp. 103-112). New York, NY: Springer US.
- [67] R. Dhudum, A. Ganeshpurkar & A. Pawar, (2024), Revolutionizing drug discovery: A Comprehensive review of AI applications. *Drugs and Drug Candidates*, 3(1), 148-171.
- [68] H. Chen, O. Engkvist, Y. Wang, M. Olivecrona & T. Blaschke, (2018), The rise of deep learning in drug discovery. *Drug discovery today*, 23(6), 1241-1250.
- [69] F. Cheng, F. Wang, J. Tang, Y. Zhou, Z. Fu, P. Zhang & J. Cummings, (2024), Artificial intelligence and open science in discovery of disease-modifying medicines for Alzheimer's disease. *Cell Reports Medicine*, 5(2).
- [70] R. Obeidat, I. Alsmadi, Q. B. Baker, A. Al-Njadat, S. Srinivasan, Godswill Ashong & Ifeanyi Osigwe, (2025), Researching public health datasets in the era of deep learning: a systematic literature review. *Health Informatics Journal*, 31(1), 14604582241307839.
- [71] A. Kadurin, A. Aliper, A. Kazennov, P. Mamoshina, Q. Vanhaelen, K. Khrabrov & A. Zhavoronkov, (2016), The cornucopia of meaningful leads: Applying deep adversarial autoencoders for new molecule development in oncology. *Oncotarget*, 8(7), 10883.
- [72] E. Gawehn, J. A. Hiss, & G. Schneider, (2016), Deep learning in drug discovery. *Molecular informatics*, 35(1), 3-14.
- [73] S. Zhu, M. Gilbert, I. Chetty & F. Siddiqui, (2022), The 2021 landscape of FDA-approved artificial intelligence/machine learning-enabled medical devices: an analysis of the characteristics and intended use. *International journal of medical informatics*, 165, 104828.
- [74] Z. Obermeyer, B. Powers, C. Vogeli & S. Mullainathan, (2019), Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447-453.
- [75] F. Barbault & F. Maurel, (2015), Simulation with quantum mechanics/molecular mechanics for drug discovery. *Expert Opinion on Drug Discovery*, 10(10), 104.