Vol.67: e24240538, 2024
https://doi.org/10.1590/1678-4324-2024240538
ISSN 1678-4324 Online Edition

BABT
Brazilian Archives of Biology and Technology

Review- Human and Animal Health

Role of Artificial Intelligence and Machine Learning in Sustainable Drug Discovery

Tammanna Ravee Sahrawat[1*]
https://orcid.org/ 0000-0002-9967-9599

[1]Panjab University, Centre for Systems Biology & and Bioinformatics, UIEAST, Chandigarh, India.

Editor-in-Chief: Paulo Vitor Farago
Associate Editor: Paulo Vitor Farago

Received: 20-Jun-2024; Accepted: 26-Aug-2024

*Correspondence: tammanna@pu.ac.in; Tel.: +91-9876033577 (T.R.S.).

HIGHLIGHTS

- This review provides an overview of the theory and applications of Artificial intelligence in drug discovery.
- Sustainability in pharmaceutical development opens avenues for broader range exploration at lower costs.
- Recent applications of various active learning methods in drug discovery research.

Abstract: In recent years, big complex data sets have been generated using omics analysis, necessitating the use of various bioinformatics and computational tools for analysis and interpretation. The exploitation of artificial intelligence (AI) and big data will enable *in silico* prediction of many parameters that are presently tested in vivo. Automation and multiplexing synthetic approaches will allow a broader range of exploration of screening of thousands of compounds with better reproducibility at lower consumption and overall lower costs thereby effectively speeding up the process of drug discovery. Other advanced techniques such as Machine Learning (ML) and Deep Learning (DL) algorithms are now being used to perform virtual screening of various lead compounds that can be developed into drugs. Integration of new technology such as AI have with drug discovery process will make the design and optimization process more robust by substantially eliminating bias and human mistakes as well as the time required for testing the candidate compounds for drug development. In the years to come advancements in AI technology coupled with computational power could revolutionize the drug discovery process with a strong footling based on sustainability in pharmaceutical development. This review presents the applications of AI algorithms such as supervised learning, reinforcement learning and others can contribute at various stages of drug discovery pipelines.

Keywords: Artificial intelligence (AI); Machine learning (ML); Deep learning (DL); Drug design; Bioinformatics.

INTRODUCTION

Sustainable development is a remarkably persistent and ideal concept and its emphasis has to be paid from the cradle to the grave in all fields including drug discovery *i.e*. initial phase to the later stages of product development or clinical stages. Sustainability in drug discovery can be achieved by responsible research and innovation as one of the reasons of the decreased drug discovery productivity is the increased difficulty (and costs) in finding (and validating) novel therapeutic targets.

Artificial intelligence (AI) and big data can play a crucial role in increasing the speed and efficiency of drug development by reducing the number of test animals and the human and environmental burden [1]. High throughput process of automation and parallelization of chemical synthesis along with the use of *in silico* AI based approaches will open the possibility to explore the extremely large chemical which is not possible with manual and serial compound synthesis. Therefore systems biology and its allied fields of pharmacogenomics and metabolomics will guide drug discovery towards a more optimal balance of risks and benefits for patients [2]. AI and deep learning (DL) strategies for activity and bioavailability prediction along with the big data approaches are likely to play a more important role in the drug discovery in the years to come.

Brief history of AI

John McCarthy in 1956 first coined the term artificial intelligence to describe “The science and engineering of making intelligent machines” [3] and this description still holds good today. AI is a multidiscipline field, which involves integration of diverse discipline such as computer science, mathematics, linguistics, neuroscience, artificial psychology *etc*. A broad range of problems can be addressed using AI but they are some basic methods that play major role such as data acquisition and maintenance, knowledge representations, solution search, logic reasoning and machine learning (ML).

The first peak that AI had was in mid 1930 when the idea of Universal Turning machine was introduced that could stimulate any computer [4], followed by the second peak in the early 1980 when substantial progress was made in AI related models including feed-forward neural networks and back-propagation algorithm [5]. These tools can be used for construction of abstract models and provide a way to update the model using a given input [6, 7, 8]. AI was first used in the fields of chemistry and molecular biology for predicting secondary structure from protein sequence information in the year 1988 [9]. Thereafter, AI went on a back foot due to the emergence of faster desktop computers due to which interest in AI decreased considerably primarily due to high maintenance cost of expert systems.

A new paradigm of machine learning (ML) generated a lot of excitement among the scientific community in the 1980’s since it placed considerable emphasis on learning actionable insights from complex data instead of explicitly being programmed as required in the expert systems of AI. ML models were found to be simple as they were trained to discover various patterns in data such as identifying the molecular features present in a group of chemical structures, which in turn could be associated with the particular biological effect, thereby helping in making predictions based on such descriptions of unknown chemicals. This led to the development of quantitative structure activity relationship (QSAR) modeling which is an integral part of cheminformatics [10].

The field of drug discovery and development have greatly benefitted from the advancements in computational science. AI is widely used in both industry and academia, while ML, an essential component in AI, has been integrated into many fields, such as data generation and analytics. Mathematical and computational theory form the basis of algorithm-based techniques, such as ML and use of ML models in technologies, such as deep learning (DL) [11-13]. These computer-assisted computational techniques, first explored in the 1950s, have shown much promise in drug discovery Figure 1.

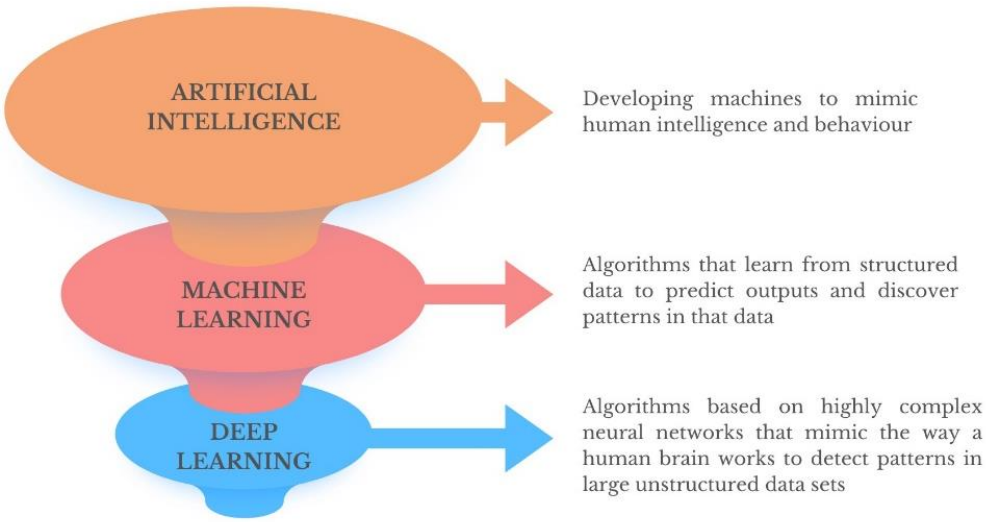

Figure 1. The Relationship between Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning (DL)

Basic principles of AI

Artificial intelligence algorithms are a broad field that encompasses a large variety of methods. To understand the applications of various AI, ML and DL methods in the process of drug development, some basic principles of AI based algorithms are discussed briefly. There are seven categories in which learning tasks and techniques employed for drug development projects are mainly classified into:

1. *Supervised learning algorithms*: Known responses called output for a given set of input data, are required, as the training process involves learning to map the output to the input date such that for output can be predicted correctly for unseen input data. Support Vector Machine (SVM) is an approach based on supervised leaning. Example: for a library of compounds, which have been classified into active or inactive, a supervised algorithm can be used to learn the relationship between molecular features present in compounds and bioactivity such that new molecules can be predicted as active or inactive compounds [14, 15].

2. *Unsupervised learning algorithms*: Only input data is available with no corresponding output variables. Therefore, the process will include extraction of structural features from the data which can then be used to group the input samples into different classes. Clustering and projection algorithms such as k-means clustering, hierarchical clustering, principal component analysis (PCA), and self-organizing map (SOMs) are some examples [16, 17].

3. *Semi-supervised learning algorithms*: Most of the drug discovery problems fall in this category due to enormous size of the chemical space (comprising $>10^{60}$ molecules) with large amount of input data is available with relatively few output data. Some common methods based on semi-supervised learning algorithms are self-training and co-training [18-20].

4. *Active learning algorithms*: It is a specialized branch of semi-supervised learning in which the algorithm can interactively query the user or some other source of information to determine the classification of input data. Active learning strategies combined with other AI algorithms have been shown to match intuitively with the properties desired in many other aspects of drug discovery [21]. Methods based on active leaning are pool- based active learning (learner chooses sample to be labeled in the pool of data that is unlabeled) and selective sampling (learner decides whether to query or discard the incoming stream of unlabeled data) [22].

5. *Reinforce learning algorithms*: It works on the principle of reward and penalties by allowing design and optimization of a system without necessarily knowing the correct or optimal approach as long as the reward function correlates with the success of the objective. Q-learning methods (such as the deep Q-network (DQN) [23] and policy gradient methods (such as trust region policy optimization (TRPO) [24] are based on reinforcement learning algorithms.

6. *Transfer learning algorithms*: These methods learn and transfer information/knowledge from source data to target data which if successful would improve the predictive performance on the target domain. The disability of transfer of knowledge from source to target depends upon the degree of relatedness reference [25].

7. *Multi-task learning algorithms*: In these methods several different but conceptually related tasks are learned in parallel making use of shared interpretation and are particularly useful for multi-target prediction [26].

Also available are various learning algorithms based on mathematical principles of machine learning algorithms such as Bayesian Algorithms, Instance-Based Methods, Decision Tree Algorithms, Ensemble Algorithms, Dimensionality Reduction, Artificial Neural Networks, (includes Feed-Forward Network, Recurrent Neural Network, Boltzmann Machine and its Variants, Autoencoder and its Variants, Convolutional Neural Networks, Attention and Attention-Based Architectures) [27].

Artificial Intelligence Application in Drug Discovery

Technologies incorporating AI are versatile tools which can be applied at various stages of drug development (Figure 2) such as identification and validation of drug targets, designing of new drugs, drug repurposing, improving R & D efficiency, aggregating and analyzing biometrical information and refining the decision making process for clinical trials reference [28-29].

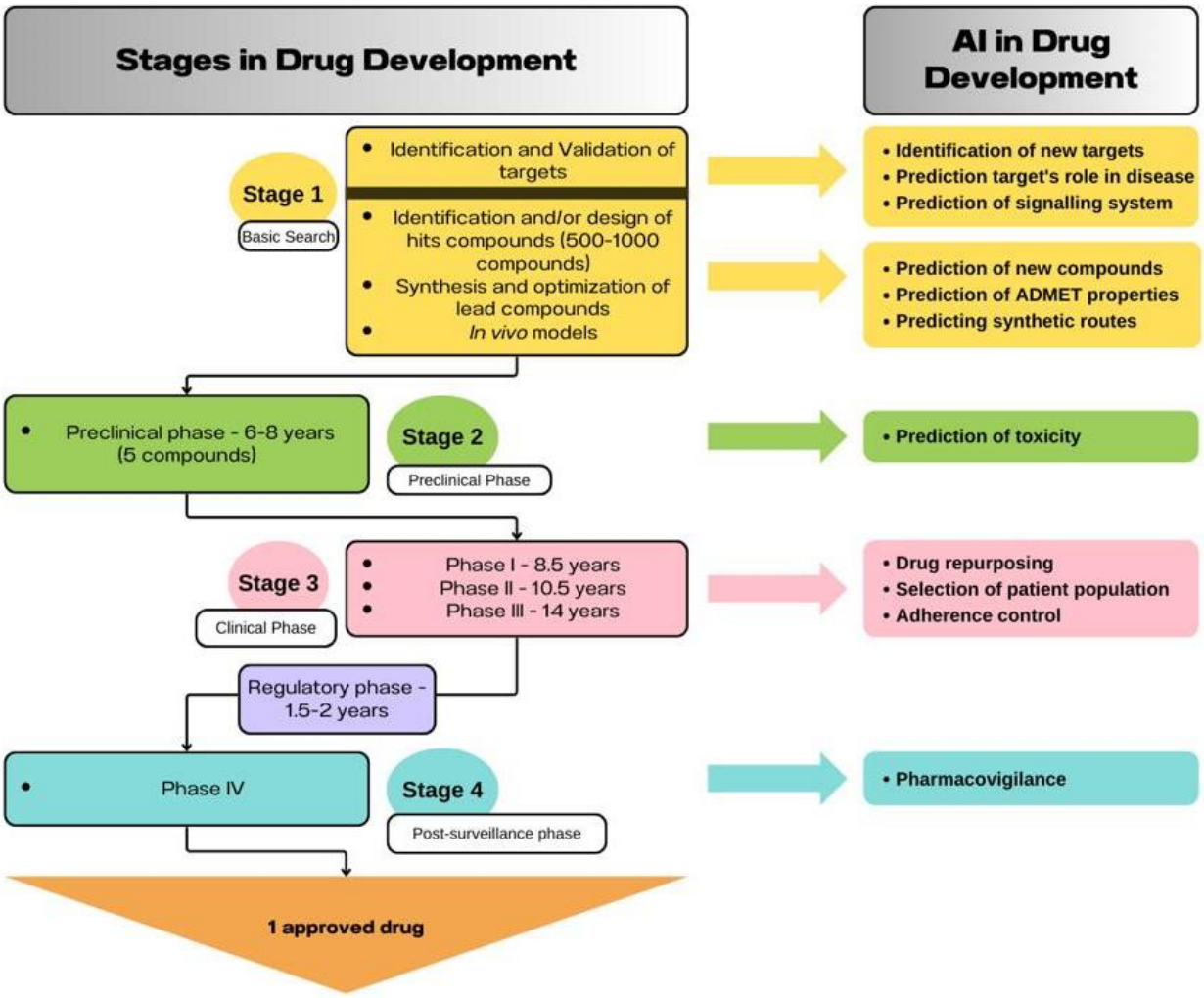

Figure 2. Role of Artificial intelligence (AI) in the various stages of drug development process

The advantages of using AI over classical drug development methods includes minimizing bias and human intervention thereby reducing the inefficiency and uncertainty in the drug development process. By using various AI methods including deep learning, success has been achieved in identifying potential drug candidates and predicting their properties accurately along with their possible toxicity risks [30]. Supplementary Table 1 enlists the various AI tools used in drug development pipeline.

Application of AI in predicting protein structure and identifying molecular targets and their pathways

Treatment of diseases using drugs involves assignment of correct target for the drug molecule. For selectively targeting of disease, it is essential to predict structure of target proteins as numerous proteins/molecular target involved in different pathways may be involved in development of disease. Structure based drug design by predicting protein structure using AI aids in drug design based on target site of the protein thereby taking into consideration the effect of drug on the target for safety considerations before synthesis. Deep Neural networks (DNN) based AI tool Alphafold has been successful in predicting 3-D protein structures [31]. Using AI it is now possible to incorporate information of genomics, biochemical attributes and target tractability during the process of pathway and target identification [32]. IBM Watson for Drug Discovery is one such AI platform that has been able to identify five new RNA binding proteins which have been linked to the pathogenesis of amyotrophic lateral sclerosis (ALS), a neurodegenerative disease [33].

Application of AI in finding hit or lead compounds

Identification of small drug like molecules is a critical step in the drug discovery process. From the large chemical space available, the initial step is to identify novel and high quality molecules using AI [34]. Identification of target specific virtual molecules and their association with their targets can also be achieved using ML techniques and predictive model software thereby optimizing the safety and efficacy attributes.

Open access databases (Figure 3) are of much use were structural data is insufficient. Algorithms that are based on network systems biology data, phenotypic data or disease can also be used. AI offers the possibility to effectively prioritize molecules based on the ease of synthesis of developed tools that are effective for optimizing the synthetic root. One of the biggest advantages offered by AI systems is reducing R & D expenditure by decreasing the number of compounds that are required to be synthesized and tested *in vivo* and *in vitro* [35].

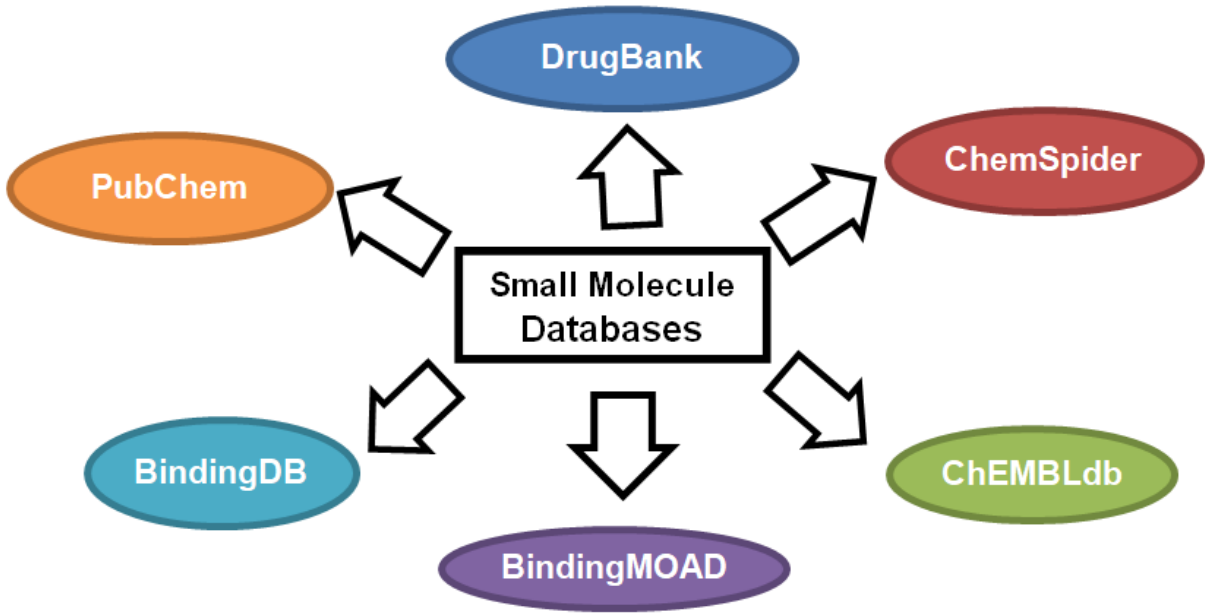

Figure 3. Open access small molecule databases used during *in silico* drug design

Application of AI in synthesis of drug-like compounds

Retro synthetic approaches are extensively used by chemists for synthesis of drug-like molecules that obey Lipinski's rule of five. It involves analysis of the target compounds recursively to sequentially convert them into smaller fragments of building blocks which can be easily prepared. This is followed by identification of the reactions that will convert these fragments into target compounds. AI would play an integral role in predicting the best reactions and one such platform is 3-NMCTS developed by combining deep neural networks with Monte Carlo research for CAOCS i.e. computer-aided organic compound synthesis, but this is not part of the Computer aided drug discovery work flow. This platform offers the advantage of being faster and better to filter out from building blocks to select only well known reactions for the synthesis of the target compounds [36].

Application of AI in predicting the mode-of-action of compounds

The ability to protect the *in vivo* safety profiles of compounds along with their on and off-target effects before their synthesis is particularly important as it would help to reduce the time cause and attribution rates of the drug development programmes. AI based programs namely DeepTox (predicts toxicity of new compounds), TargeTox (biological network target-based drug toxicity risk prediction method based on the principle that similar functional properties share similarities in biological networks) and PrOCTOR (predicts the probability of toxicity in clinical trials) [37, 38] are available and their predictive accuracy could be improved with data sets that are bigger in size and refined to have information about toxicity and therapeutic profiles. Other AI tools are SPiDER and read-across structure–activity relationships (RASAR) that predict the toxicity properties associated with molecular structures by mining a large database of chemicals [39, 40].

Application of AI in population selection for clinical trials

Specific patient population selection for a clinical trial is a crucial and its success requires development of AI approaches to identify and predict relationship between human-relevant biomarkers and in vitro phenotypes that afford a more predictable, quantifiable assessment of the uncertainty of therapeutic responses in a specific patients [41, 42].

Predictive modeling using AI can assist clinical trials by recognizing disease in patients, identifying gene targets and predicting the effect of the drug-like compounds and their corresponding on- and off-target effects. Medication adherence of Phase II trial of schizophrenia patients was tracked through AiCure, and it was reported that AiCure increased adherence 25% as compared to directly observed therapy [43].

Application of AI in predicting drug-protein interactions and drug repurposing/ polypharmacology

Accurate prediction of drug–protein interactions ensures better therapeutic efficacy [44]. Various AI methods have been useful in the prediction of drug–protein interactions which could be extended to target–disease and target–target associations to speed up the drug discovery process [45]. SVM approach was used to construct a model for predicting ligand–protein interaction based on the primary sequence of proteins and the structural features of small molecules to discover nine novel active compounds for 4 pharmacologically important targets namely GPR40, SIRT1, p38, and GSK-3β. [46]. In a computationally driven network, the ML approach is widely used, which utilizes techniques such as SVM, NN, logistic regression, and DL. PREDICT, SPACE, and other ML approaches, consider drug–drug, disease–disease similarity, the similarity between target molecules, chemical structure, and gene expression profiles for drug repurposing [47].

Polypharmacology or drug repurposing has been defined as the ability of small molecules to interact simultaneously and specifically with multiple targets [48]. Drug repurposing allows applying an already exiting drug to a new disease and is advantageous as the drug is qualified to directly go to Phase II trials without having to pass through Phase I clinical trials and toxicology testing again thereby reducing expenditure [49]. *In silico* methods through DL applications (such as deep neural networks), to predict pharmacological properties of drug and drug repurposing using transcriptomic data comprising various biological systems and conditions have been reported [50]. Generative adversarial networks (GANs) is a next-generation AI that uses DL to produce photo-realistic pictures from text descriptions to design drug molecules by imagining or creating new data modeled on real data [51]. Databases, such as ZINC, PubChem, Ligand Expo, KEGG, ChEMBL, DrugBank, STITCH, BindingDB, Supertarget, PDB, help to integrate information of diverse molecular pathways, crystal structures, binding affinities, drug targets, disease relevance, chemical properties and biological activities which can be used by AI to to design polypharmacological agents. DeepDDI has been developed using AI for understanding of drug– drug interactions, associated mechanisms and prediction of alternative drugs for intended clinical use without negative health effects [47].

CONCLUSION

This review summarizes the tools and techniques based on AI that are being used to accelerate process of drug design, which is otherwise a long and costly process based on traditional methods. In the current wave of AI has been ushered in due to advances in AI algorithms, availability of big data, along with increase in the architectural hardware. Using computational approaches based on AI, ML and DL by employing data science methods, target identification, de novo design, drug repurposing, prediction of reactivity and bio-activity can be done within reasonable time-scales thereby providing impetus to innovative drug discovery. These tools will play a significant role in sustainability of pharmaceutical practices so as to reduce the environmental footprint of pharmaceuticals.

Funding: This research received no external funding.
Acknowledgments: The author acknowledges the infrastructure and technical support of Panjab University, Chandigarh, India.
Conflicts of Interest: The author declares no conflict of interest.
Supplementary Material: Table 1 enlists the various AI tools used in drug development pipeline. Available in: https://www.documentador.pr.gov.br/documentador/pub.do?action=d&uuid=@gtf-escriba-tecpar@a5818326-2b63-4c53-bcfc-f94dbf9751ea

REFERENCES

1. Hunter J. Adopting AI is essential for a sustainable pharma industry. Drug Discov. World. 2016:69-71.
2. Klont F, Kremer D, Neto AW, Berger SP, Touw DJ, Hak E, et al. Metabolomics data complemented drug use information in epidemiological databases: pilot study of potential kidney donors. J Clin Epidemio. 2021 Jul 1;135:10-6.
3. McCarthy J, Hayes PJ. Some philosophical problems from the standpoint of artificial intelligence. In Readings in artificial intelligence. 1981 Jan 1; 431-50)
4. Gastaldi JL. Computing Cultures: Historical and Philosophical Perspectives. Minds Mach. 2024 Feb;34 (Suppl 1):1-10.
5. Gholap AD, Uddin MJ, Faiyazuddin M, Omri A, Gowri S, Khalid M. Advances in Artificial Intelligence in Drug Delivery and Development: A Comprehensive Review. Comput Biol Med. 2024 Jun 7:108702.
6. Baum EB. On the capabilities of multilayer perceptrons. J Complex.1988 Sep 1;4(3):193-215.
7. Huang SC, Huang YF. Bounds on number of hidden neurons of multilayer perceptrons in classification and recognition. In1990 IEEE International Symposium on Circuits and Systems (ISCAS) 1990 May 1 (pp. 2500-03). IEEE.
8. Rumelhart DE, Hinton GE, Williams RJ. Learning Internal Representations by Error Propagation, Parallel Distributed Processing, Explorations in the Microstructure of Cognition, ed. DE Rumelhart and J. McClelland. Vol. 1. 1986. Biometrika. 1986;71:599-607.
9. Qian N, Sejnowski TJ. Predicting the secondary structure of globular proteins using neural network models. J Mol Biol.1988 Aug 20;202(4):865-84.
10. Tyrchan C, Evertsson E. Matched molecular pair analysis in short: algorithms, applications and limitations. Comput Struct Biotechnol J. 2017 Jan 1;15:86-90.
11. Bishop CM. Model-based machine learning. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2013 Feb 13;371(1984):20120222.
12. VoPham T, Hart JE, Laden F, Chiang YY. Emerging trends in geospatial artificial intelligence (geoAI): potential applications for environmental epidemiology. Environ Health. 2018 Dec;17:1-6.
13. Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T. The rise of deep learning in drug discovery. Drug Discov Today. 2018 Jun 1;23(6):1241-50.
14. Raymond JL, Medina JF. Computational principles of supervised learning in the cerebellum. Annu Rev Neurosci. 2018 Jul 8;41:233-53.
15. Joachims T. Text categorization with support vector machines: Learning with many relevant features. InEuropean conference on machine learning 1998 Apr 21 (pp. 137-142). Berlin, Heidelberg: Springer Berlin Heidelberg.
16. Wold S, Esbensen K, Geladi P. Principal component analysis. Chemom Intell Lab Syst. 1987 Aug 1;2(1-3):37-52.
17. Kohonen T. The self-organizing map. Proceedings of the IEEE. 1990 Sep;78(9):1464-80.
18. Chapelle O, Scholkopf B, Zien A. Semi-supervised learning (Chapelle, O. et al., eds.; 2006)[book reviews]. IEEE Transactions on Neural Networks. 2009 Feb 24;20(3):542.
19. Yarowsky D. Unsupervised word sense disambiguation rivaling supervised methods. In33rd annual meeting of the association for computational linguistics 1995 Jun (pp. 189-96).
20. Blum A, Mitchell T. Combining labeled and unlabeled data with co-training. InProceedings of the eleventh annual conference on Computational learning theory 1998 Jul 24 (pp. 92-100).
21. Reker D, Schneider G. Active-learning strategies in computer-assisted drug discovery. Drug Discov Today. 2015 Apr 1;20(4):458-65.
22. Atlas L, Cohn D, Ladner R. Training connectionist networks with queries and selective sampling. Adv Neural Inf Process Syst. 1989;2.
23. Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, et al. Human-level control through deep reinforcement learning. Nature. 2015 Feb 26;518(7540):529-33.
24. Schulman J, Levine S, Abbeel P, Jordan M, Moritz P. Trust region policy optimization. InInternational conference on machine learning 2015 Jun 1 (pp. 1889-97). PMLR.
25. Pan SJ, Yang Q. A survey on transfer learning. IEEE Transactions on knowledge and data engineering. 2009 Oct 16;22(10):1345-59.
26. Dahl GE, Jaitly N, Salakhutdinov R. Multi-task neural networks for QSAR predictions. arXiv preprint arXiv:1406.1231. 2014 Jun 4.
27. Vemula D, Jayasurya P, Sushmitha V, Kumar YN, Bhandari V. CADD, AI and ML in drug discovery: A comprehensive review. Eur J Pharm Sci. 2023 Feb 1;181:106324.
28. Huang Z, Juarez JM, Li X. Data mining for biomedicine and healthcare. J Healthc Eng. 2017;2017.
29. Mamoshina P, Vieira A, Putin E, Zhavoronkov A. Applications of deep learning in biomedicine. Mol Pharm. 2016 May 2;13(5):1445-54.
30. Hughes JP, Rees S, Kalindjian SB, Philpott KL. Principles of early drug discovery. Br J Pharmacol. 2011 Mar;162(6):1239-49.
31. Hutson M. AI protein-folding algorithms solve structures faster than ever. Nature. 2019 Jul 22.
32. Wang Q, Feng Y, Huang J, Wang T, Cheng G. A novel framework for the identification of drug target proteins: Combining stacked auto-encoders with a biased support vector machine. PloS One. 2017 Apr 28;12(4):e0176486.

33. Bakkar N, Kovalik T, Lorenzini I, Spangler S, Lacoste A, Sponaugle K, et al. Artificial intelligence in neurodegenerative disease research: use of IBM Watson to identify additional RNA-binding proteins altered in amyotrophic lateral sclerosis. Acta Neuropathol.. 2018 Feb;135:227-47.
34. Reymond JL, Van Deursen R, Blum LC, Ruddigkeit L. Chemical space as a source for new drugs. Med Chem Comm. 2010;1(1):30-8.
35. Okafo G. Adapting drug discovery to artificial intelligence. Drug Target Rev. 2018;5(2):50-2.
36. Segler MH, Preuss M, Waller MP. Planning chemical syntheses with deep neural networks and symbolic AI. Nature. 2018 Mar 29;555(7698):604-10.
37. Mayr A, Klambauer G, Unterthiner T, Hochreiter S. DeepTox: toxicity prediction using deep learning. Front Environ Sci. 2016 Feb 2;3:80.
38. Gayvert KM, Madhukar NS, Elemento O. A data-driven approach to predicting successes and failures of clinical trials. Cell Chem Biol. 2016 Oct 20;23(10):1294-301.
39. Schneider P, Tanrikulu Y, Schneider G. Self-organizing maps in drug discovery: compound library design, scaffold-hopping, repurposing. Curr Med Chem. 2009 Jan 1;16(3):258-66.
40. Reker D, Rodrigues T, Schneider P, Schneider G. Identifying the macromolecular targets of de novo-designed chemical entities through self-organizing map consensus. Proceedings of the National Academy of Sciences. 2014 Mar 18;111(11):4067-72.
41. Perez-Gracia JL, Sanmamed MF, Bosch A, Patiño-Garcia A, Schalper KA, Segura V, et al.Strategies to design clinical studies to identify predictive biomarkers in cancer research. Cancer Treat Rev. 2017 Feb 1;53:79-97.
42. Harrer S, Shah P, Antony B, Hu J. Artificial intelligence for clinical trial design. Trends Pharmacol Sci. 2019 Aug 1;40(8):577-91.
43. Bain EE, Shafner L, Walling DP, Othman AA, Chuang-Stein C, Hinkle J, et al. Use of a novel artificial intelligence platform on mobile devices to assess dosing compliance in a phase 2 clinical trial in subjects with schizophrenia. JMIR mHealth and uHealth. 2017 Feb 21;5(2):e7030.
44. Wan F, Zeng J. Deep learning with feature embedding for compound-protein interaction prediction. Biorxiv. 2016 Nov 7:086033.
45. Yu H, Chen J, Xu X, Li Y, Zhao H, Fang Y, et al. A systematic prediction of multiple drug-target interactions from chemical, genomic, and pharmacological data. PloS One. 2012 May 30;7(5):e37608.
46. Wang F, Liu D, Wang H, Luo C, Zheng M, Liu H, et al. Computational screening for active compounds targeting protein sequences: methodology and experimental validation. J Chem Inf Model. 2011 Nov 28;51(11):2821-8.
47. Park K. A review of computational drug repurposing. Transl Clin Pharmacol. 2019 Jun;27(2):59.
48. Sahrawat TR, Kaur P. Polypharmacological study of Ceritinib using a structure based in silico approach. Bionatura. 2019;4(2):836-40.
49. Corsello SM, Bittker JA, Liu Z, Gould J, McCarren P, Hirschman JE, et al. The Drug Repurposing Hub: a next-generation drug library and information resource. Nat Med. 2017 Apr;23(4):405-8.
50. Aliper A, Plis S, Artemov A, Ulloa A, Mamoshina P, Zhavoronkov A. Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data. Mol Pharm. 2016 Jul 5;13(7):2524-30.
51. Khrabrov K, Zhavoronkov A, Kadurin A, Nikolenko S, Aliper A. DruGAN: An Advanced Generative Adversarial Autoencoder Model for de Novo Generation of New Molecules with Desired Molecular Properties in Silico.