IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, December 2024

AI-Driven Insights: Paving the Path to Next-Generation Therapeutics

Prof. Dhiraj S. Kalyankar[1], Mr. Chaitanya G. Bhagat[2], Mr. Ayush D. Kadu[3], Ms. Anjali P. Tambade[4], Mr. Kaustubh S. Dhoran[5]
Department of Computer Science and Engineering[1-5]
Dr. Rajendra Gode Institute of Technology & Research, Amravati, Maharashtra, India
dhiraj.kalyankar50@gmail.com, chaitanyabhagat143@gmail.com, ayushkadu993@gmail.com, anjalitambade51@gmail.com, kaustubhdhoran24@gmail.com

Abstract: *This chapter thoroughly examines the critical role of artificial intelligence (AI) in drug discovery and development, covering its potential, methodologies, real-world applications, and the challenges it presents. It begins with a comprehensive introduction to AI and its subfields, including machine learning (ML), deep learning (DL), and natural language processing (NLP). The chapter then outlines various AI algorithms such as regression, support vector machines, and neural networks. It also explains approaches for optimizing and validating AI models, with a focus on metrics used for their quantitative assessment.*

Next, the chapter highlights the impact of AI across different stages of the drug discovery and development process, showcasing examples of its use in AI-driven drug discovery companies and their innovative platforms. Challenges such as limited data availability, ethical concerns, and integrating AI with traditional methods are discussed, along with potential solutions like data augmentation and explainable AI (XAI). It also explores regulatory perspectives, particularly from the United States Food and Drug Administration (FDA), illustrating the growing relationship between AI and regulatory science.

The chapter concludes with a forward-looking view on AI's future in drug discovery. AI is revolutionizing the field by automating tasks such as image analysis in pathology and radiology, improving diagnostic accuracy, and reducing human error. In clinical trials, AI is used to optimize trial design, select appropriate patient groups, and monitor real-time data, leading to faster decision-making. AI also plays a key role in analyzing scientific literature, helping researchers stay current with new advancements.

Keywords: natural language processing

I. INTRODUCTION

Approximately 6-7% of the global GDP (equating to around 8.5 to 9 trillion USD) is spent annually on healthcare, with the development of a new drug costing more than $1 billion and potentially taking up to 14 years. The overall success rate in drug development, from phase I clinical trials to drug approval, is quite low across all therapeutic fields, with about 97% of cancer drugs failing in clinical trials. This leads to high risks for investors and increases the prices of successful drugs to account for the costs of failed attempts. With the digitization of medical records, there is significant potential for advancements in clinical trials, precision medicine, drug discovery, and health policy through data-driven methods. In recent years, drug discovery has undergone a significant transformation due to advancements in the analytical techniques and computational tools.

As a result, there is increasing interest in using artificial intelligence (AI) to enhance various stages of the drug discovery pipeline, including molecular design and optimization, structure-based drug design, and both preclinical and clinical development. Biomedical data, such as genomic profiles, imaging data, and chemical and drug databases, can be combined with deep learning models to streamline the discovery of new drugs and their clinical applications.

Numerous reviews have examined the role of AI in drug discovery. For example, GPU computing and deep learning models have been explored for drug discovery, as well as the impact of AI in precision medicine and material property predictions. Other reviews have highlighted the progress made since the completion of the Human Genome Project and discussed how machine learning helps understand biological interactions. AI methods applied to 3D structure-based

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, December 2024

drug discovery, molecular dynamics simulations, and protein-ligand docking have also been explored, as have AI's role in graph-based formulations of therapeutic problem.

Despite these advancements, there are significant challenges in data representation and prediction—key issues in drug design where AI can excel. One difficulty is that many drug discovery tasks do not easily fit into machine learning frameworks due to the lack of standardized datasets and knowledge representations. For instance, a drug can be represented in various formats, such as SMILES strings, ECFP, or graphs, while proteins can be depicted as 1D amino acid sequences, protein sequence representations, or 3D structures. There are also issues related to the scarcity of labeled data, inconsistencies in labels, and small sample sizes.

In that review, we discuss the potential of machine learning libraries and molecular representations, the role of graph neural networks in drug discovery, and how to address issues like data collection, labeling inconsistencies, and small sample sizes. Significant progress has been made in recent years with deep learning applications in drug discovery, including the development of open-source tools, AI-ready benchmark datasets, and deep learning platforms tailored to drug design.

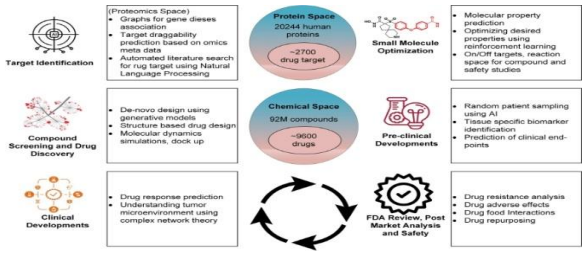

Fig. 1: Applications of AI-based at different stages of a drug discovery

AI-based methods can be applied at various stages of the drug discovery pipeline. There are around 2,700 known drug target proteins in the human body and approximately 9,600 FDA-approved small-molecule drugs. Machine learning models can identify targeted proteins, graph neural networks (GNNs) can predict drug-target interactions and binding affinities, and reinforcement learning can optimize molecular properties. Computer vision techniques can assess the spatial environment of tumors, generative models can design novel molecules, and simulations can predict the stability and dynamics of protein-drug complexes. natural language processing (NLP) can be used to mine scientific literature for drug repurposing, FDA reviews, and post-market analyses.

II. LITERATURE REVIEW

AI techniques have enabled researchers to efficiently identify and validate biological targets, such as proteins and genes, that play a crucial role in disease mechanisms. Traditionally, target identification was an experimental and time-consuming process. AI can now analyze vast datasets, including genomic, transcriptomic, and proteomic data, to predict targets with higher accuracy.

Mamoshina et al. (2016) Demonstrated how machine learning can be applied to analyze omics data, identifying disease-related biomarkers that serve as drug targets. The use of AI can rapidly screen potential targets and assess their drug ability, thereby streamlining early drug discovery.

Lavecchia (2015) Showcased how AI-based computational methods for target prediction, such as support vector machines (SVM) and random forests, significantly increase the efficiency of finding new targets. The integration of artificial intelligence (AI) in drug development is transforming the pharmaceutical industry by accelerating various stages of research, including drug discovery, preclinical testing, and clinical trials. The literature highlights the remarkable impact AI has made, especially in streamlining processes that have traditionally been time-consuming and expensive.

AI-powered algorithms for molecular modeling and de novo drug design have drastically improved the process of identifying promising lead compounds. Techniques such as generative adversarial networks (GANs) and reinforcement learning allow AI systems to generate novel molecules with desired properties.

Zhavoronkov et al. (2019) Reported the successful application of generative AI to propose novel molecular structures, optimizing key features like bioavailability and binding affinity. This has enabled the discovery of drug candidates that

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, December 2024

would have otherwise required years of experimentation. Popova et al. (2018) Described the use of variational autoencoders (VAEs) to design drug-like molecules, improving the speed and accuracy of lead generation by creating a continuous representation of molecular structures.

In Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) properties are critical in determining whether a drug will be effective and safe in humans. AI models, particularly neural networks, have demonstrated the ability to predict these properties with greater accuracy than traditional empirical methods. Cheng et al. (2013) Showed how machine learning models can be applied to large ADMET datasets to predict drug metabolism and toxicity, providing early insights into potential drug failures.

Maltarollo et al. (2017) Demonstrated the use of deep learning for ADMET property prediction, outperforming older statistical models by using large chemical databases to train predictive systems.Virtual screening, powered by AI, involves computational techniques to screen large libraries of chemical compounds to identify those most likely to bind to a specific drug target. Deep learning models and AI algorithms have significantly reduced the number of compounds that need to be tested experimentally.

Ragoza et al. (2017) Developed a deep learning-based docking model that outperformed traditional molecular docking techniques in virtual screening, providing faster and more accurate predictions of drug-target interactions. Jorgensen et al. (2016) Highlighted the importance of AI-based molecular docking methods in virtual screening, discussing their role in reducing the computational cost and time required for screening large compound libraries.

AI models are being used to streamline patient recruitment in clinical trials by analyze electronic health records (EHRs) and genetic information. This helps identify patients who meet the trial criteria more efficiently and reduces the time it takes to enroll participants.

Weng et al. (2016) Discussed how machine learning algorithms can enhance patient recruitment by predicting eligibility based on health records, significantly reducing the time required for patient matching.

Topol (2019) Explored AI's potential to personalize medicine, emphasizing its use in patient stratification during clinical trials to ensure more precise and effective treatments. AI-driven models are also being used to predict clinical trial outcomes by analyzing patient data and real-time results. This allows for adaptive trial designs, where trials can be modified based on interim results to improve success rates.

Fleming et al. (2020) Described how AI and real-world evidence are being integrated into clinical trial designs to create more adaptive and efficient trials, increasing the likelihood of a trial success.

Once a drug is approved, AI can help monitor post-market safety by detecting adverse drug reactions (ADRs) through natural language processing (NLP) techniques that analyze social media, medical records, and patient reports. This is crucial for identifying potential safety issues that may not have surfaced during clinical trials. Liu et al. (2019) Developed an AI system using NLP to detect ADRs from unstructured medical records and patient forums. The model provided faster ADR detection compared to conventional pharmacovigilance methods.

III. ARCHITECTURE AND WORKING

3.1 AI in drug discovery

AI has revolutionized drug discovery so that it is capable of overcoming all the challenges related to the enormous chemical space embracing more than 10^{60} molecules. The conventional drug development process is slow and long-winding, whereas the AI accelerates the identification of potential candidates as drugs, validates drug targets, and optimizes the design of drugs in general. In this way, AI models, specifically ML and DL models, are a good efficient source in the identification of promising hit and lead compounds to dramatically reduce time and cost in drug discovery.

AI tools have proved to be beneficial across most of the drug development process. It has been utilized in predicting the structure of drug molecules and assessing their efficacy and safety. Thus, in VS, AI performs the task of sifting through very large chemical libraries to determine bioactive compounds.

The open-access chemical spaces involved are PubChem, DrugBank, and ChemDB, among others. Those AI algorithms for drug design help in rapid selection of lead compounds based upon their physical, chemical, and toxicological properties through molecular fingerprint recognition and QSAR models.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, December 2024

3.2 AI in Drug Screening

The deployment of AI tools improves the drug screening efficiency by predicting physicochemical properties, bioactivity, and toxicity of compounds. For instance, in VS, DNNs and SVMs algorithms are employed to predict synthesis feasibility as well as in vivo activity and toxicity. The major pharmaceutical companies have already exploited AI to advance their drug discovery platforms in immuno-oncology and cardiovascular disease.

Prediction of Physicochemical Property Based on AI Model: These AI models predict major critical properties, such as solubility, partition coefficient, and permeability, which in turn decide a drug's pharmacokinetic parameters. Machine learning models can generate feasible molecules through large sets of compound optimization data and predict their properties.

Several methods have been developed that fall into the category of ANN-based and graph-based models. Acid dissociation constants and lipophilicity have been predicted using these approaches. For example, some DL methods, such as undirected graph recursive neural networks, use solubility predictions. Others use AI predictors fed by cellular data coming from in vitro assays to model the permeability of molecules.

Bioactivity Prediction: Efficacy in drug performance is also related to drug molecules interacting with target proteins. For example, DeepDTA and PADME are AI-based models that assess the binding affinity of drugs for proteins based on features like the chemical structure of the drug and the sequence of the protein. These models, sometimes more efficient than the traditional methods, can predict drug-target interactions without requiring a complete 3D protein structure. Some of the popular AI tools used include XenoSite and SMARTCyp to determine metabolism sites, as well as methods such as MANTRA that classify drugs according to gene expression profiles, thus enabling the prediction of therapeutic effects.

Toxicity prediction: The potential toxicity of a compound is one entity that needs to be predicted in order to avoid adverse effects. AI tools such as DeepTox rely on enormous datasets of chemical descriptors to predict toxicity better by identifying the static and dynamic features that are present in molecular structures. Such online tools, including LimTox and Toxtree, assist further in predicting toxicity, thereby minimizing expensive in vitro and animal testing.

3.3 AI in Drug Design

Target Protein Structure Prediction: AI helped predict the 3D protein structure, which is an important step in the structure-based drug discovery process. In this regard, AI models, like AlphaFold, study the chemical environment around a protein to determine its precise structure. DL models, such as RGN, and many other techniques have been used for effectively predicting accurate 3D protein structures to help researchers target the specific proteins caused by the disease.

Drug-Protein Interaction Prediction: In drug-protein interaction prediction, correct prediction of a drug's interactions is important for the efficacy of drugs and avoidance of off-target effects. AI-based techniques such as SVM and RF models scan thousands of protein-ligand interactions and explore new drugs along with their corresponding interactions with several receptors. They also predict polypharmacology when several drugs interact with a given set of receptors responsible for side effects.

3.4 De Novo Drug Design

AI has revolutionized the process of de novo design for drug molecules, which can synthesize new compounds. DL has substituted the age-old methods of traditional design, such as predicting the bioactivity of new molecules. Synthia programmes encode chemical synthesis rules to generate possible synthesis routes for drug compounds, while RANC is a platform that uses reinforcement learning for small organic molecules design. These approaches have also been developed quicker and cheaper than other older one falciparum.

IV. APPLICATION

Nanorobots for AI-Assisted Drug Delivery

Nanorobots are microscopic machines composed of integrated circuits, sensors, and power sources, requiring advanced computational capabilities, such as AI, to perform their functions. These robots can be programmed to avoid collisions, identify target locations, attach to cells, and exit the body after completing their task. Recent developments have

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, December 2024

enabled nanorobots to navigate more precisely, using factors like pH levels to reach targeted areas, reducing systemic side effects and enhancing treatment success. A more advanced application involves implantable nanorobots that deliver drugs or genes through controlled mechanisms, such as adjustable dosing or sustained release. This is made possible through AI tools, including neural networks and fuzzy logic, which can fine-tune the drug release process.

AI in Predicting Synergism and Antagonism in Drug Combinations

Many complex diseases, like cancer and tuberculosis, are treated using drug combinations that produce synergistic effects, leading to faster recovery. AI technologies, such as artificial neural networks (ANN), logistic regression, and network-based models, can streamline the process of screening and selecting optimal drug combinations. For example, Rashid et al. developed a quadratic phenotype optimization platform that identified the most effective drug combinations for treating bortezomib-resistant multiple myeloma. Using 114 FDA-approved drugs, the platform recommended decitabine and mitomycin C as the most effective combination.

AI-Powered Data Mining Bots for Research and Analytics

Data mining bots harness AI to analyze large datasets from research studies, clinical trials, and scientific literature. These bots can identify hidden patterns and trends within unstructured data, offering researchers valuable insights that might otherwise go unnoticed. By automating data collection and analysis, researchers can focus more on interpreting results rather than sifting through data. For example, data mining bots can identify potential drug interactions or repurpose existing drugs based on previous studies. They can also track emerging public health trends that may influence future drug development. By providing comprehensive data analysis, these bots support researchers and pharmaceutical companies in making informed decisions, ultimately leading to innovations that benefit both patients and the pharmaceutical industry.

Telemedicine Bots Virtual Health Consultations

Telemedicine bots facilitate virtual consultations between patients and healthcare providers, improving access to care and overall efficiency. These bots collect patient symptoms by asking relevant questions and provide initial diagnoses and treatment suggestions. They can also prioritize urgent cases and recommend appropriate care levels. By integrating with electronic health records (EHRs), telemedicine bots ensure that healthcare providers have access to a patient's medical history during consultations, thereby enhancing care quality. These bots can also schedule follow-up visits and referrals, improving healthcare resource allocation, especially during times of high patient load or pandemics. As telemedicine becomes more widespread, such bots will play a crucial role in patient engagement and access to healthcare.

V. CONCLUTION

AI-based methods are transforming healthcare by offering low-cost, intelligent, and adaptable solutions in areas such as drug design, clinical decision-making, diagnosis, prevention, and medical recommendations. Earlier, AI applications were seen as less effective compared to methods like high-throughput screening and combinatorial chemistry. Designing new chemical compounds from scratch using software relied largely on the hope that desirable properties would emerge early in the process. However, this traditionally lengthy and costly approach has evolved into a more efficient process, breaking down drug design into distinct stages: target identification, de novo molecular design, drug repositioning, retrosynthesis, and predictions of reactivity and bioactivity, followed by FDA approval and post-market analysis. The use of AI is growing in the pharmaceutical industry, with revenue from AI-based solutions projected to reach $2.199 billion by 2022. Deep Neural Networks (DNNs) enhance AI's predictive capabilities when analyzing small molecules, while one-shot learning is helpful when experimental data is limited. It's important to recognize the impact of technical and human errors, labeling limitations, and biological variability in the data. While converting experimental data into numerical or computer-assisted formats is challenging, AI is helping create trial representations that enable data classification and the development of predictive model.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, December 2024

REFERENCES

[1]. Organization W.H. World Health Organization; 2017. New perspectives on global health spending for universal health coverage. Tech. Rep. [Google Scholar]

[2]. Wouters O.J., McKee M., Luyten J. Estimated research and development investment needed to bring a new medicine to market, 2009-2018. JAMA. 2020;323(9):844–853. doi: 10.1001/jama.2020.1166. [DOI] [PMC free article] [PubMed] [Google Scholar]

[3]. Khanna I. Drug discovery in pharmaceutical industry: productivity challenges and trends. Drug Discov. Today. 2012;17(19–20):1088–1102. doi: 10.1016/j.drudis.2012.05.007. [DOI] [PubMed] [Google Scholar]

[4]. Lin A., Giuliano C.J., Palladino A., John K.M., Abramowicz C., Yuan M.L., Sausville E.L., Lukow D.A., Liu L., Chait A.R., et al. Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. Sci. Transl. Med. 2019;11(509) doi: 10.1126/scitranslmed.aaw8412. [DOI] [PMC free article] [PubMed] [Google Scholar]

[5]. Freedman D.H., et al. Hunting for new drugs with ai. Nature. 2019;576(7787):S49–S53. doi: 10.1038/d41586-019-03846-0. [DOI] [PubMed] [Google Scholar]

[6]. Davenport T., Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc. J. 2019;6(2):94. doi: 10.7861/futurehosp.6-2-94. [DOI] [PMC free article] [PubMed] [Google Scholar]

[7]. Yu K.-H., Beam A.L., Kohane I.S. Artificial intelligence in healthcare. Nat. Biomed. Eng. 2018;2(10):719–731. doi: 10.1038/s41551-018-0305-z. [DOI] [PubMed] [Google Scholar]

[8]. Pandey M., Fernandez M., Gentile F., Isayev O., Tropsha A., Stern A.C., Cherkasov A. The transformational role of gpu computing and deep learning in drug discovery. Nat. Mach. Intell. 2022;4(3):211–221. [Google Scholar]

[9]. Sierra-Sosa D., Garcia-Zapirain B., Castillo C., Oleagordia I., Nuño-Solinis R., Urtaran-Laresgoiti M., Elmaghraby A. Scalable healthcare assessment for diabetic patients using deep learning on multiple gpus. IEEE Trans. Ind. Inform. 2019;15(10):5682–5689. [Google Scholar]

[10]. Powles J., Hodson H. Google deepmind and healthcare in an age of algorithms. Health Technol. 2017;7(4):351–367. doi: 10.1007/s12553-017-0179-1. [DOI] [PMC free article] [PubMed] [Google Scholar]

[11]. Chen H., Engkvist O., Wang Y., Olivecrona M., Blaschke T. The rise of deep learning in drug discovery. Drug Discov. Today. 2018;23(6):12411250.doi:10.1016/j.drudis.2018.01.039. [DOI] [PubMed] [Google Scholar]

[12]. Vamathevan J., Clark D., Czodrowski P., Dunham I., Ferran E., Lee G., Li B., Madabhushi A., Shah P., Spitzer M., et al. Applications of machine learning in drug discovery and development. Nat. Rev. Drug Discov. 2019;18(6):463–477. doi: 10.1038/s41573-019-0024-5. [DOI] [PMC free article] [PubMed] [Google Scholar]

[13]. Zhang S., Bamakan S.M.H., Qu Q., Li S. Learning for personalized medicine: a comprehensive review from a deep learning perspective. IEEE Rev. Biomed. Eng. 2018;12:194–208. doi: 10.1109/RBME.2018.2864254. [DOI] [PubMed] [Google Scholar]

[14]. Gibbs R.A. The human genome project changed everything. Nat. Rev. Genet. 2020;21(10):575–576. doi: 10.1038/s41576-020-0275-3. [DOI] [PMC free article] [PubMed] [Google Scholar]

[15]. Sanchez-Lengeling B., Aspuru-Guzik A. Inverse molecular design using machine learning: generative models for matter engineering. Science. 2018;361(6400):360–365. doi: 10.1126/science.aat2663. [DOI] [PubMed] [Google Scholar]

[16]. Singh M., Sharma D., Garg M., Kumar A., Baliyan A., Rani R., Kumar V. Current understanding of biological interactions and processing of dna origami nanostructures: role of machine learning and implications in drug delivery. Biotechnol. Adv. 2022 doi: 10.1016/j.biotechadv.2022.108052. [DOI] [PubMed] [Google Scholar]

[17]. Ganesan A., Coote M.L., Barakat K. Molecular dynamics-driven drug discovery: leaping forward with confidence. Drug Discov. Today. 2017;22(2):249–269. doi: 10.1016/j.drudis.2016.11.001. [DOI][PubMed] [Google Scholar]

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, December 2024

[18]. Gaudelet T., Day B., Jamasb A.R., Soman J., Regep C., Liu G., Hayter J.B., Vickers R., Roberts C., Tang J., et al. Utilizing graph machine learning within drug discovery and development. Brief. Bioinform. 2021;22(6) doi: 10.1093/bib/bbab159. [DOI] [PMC free article] [PubMed] [Google Scholar]

[19]. Waring J., Lindvall C., Umeton R. Automated machine learning: review of the state-of-the-art and opportunities for healthcare. Artif. Intell. Med. 2020;104 doi: 10.1016/j.artmed.2020.101822. [DOI] [PubMed] [Google Scholar]

[20]. Deng J., Yang Z., Ojima I., Samaras D., Wang F. Artificial intelligence in drug discovery: applications and techniques. Brief. Bioinform. 2022;23(1) doi: 10.1093/bib/bbab430. [DOI] [PubMed] [Google Scholar]

[21]. Dara S., Dhamercherla S., Jadav S.S., Babu C., Ahsan M.J. Machine learning in drug discovery: a review. Artif. Intell. Rev. 2021:1–53. doi: 10.1007/s10462-021-10058-4. [DOI] [PMC free article] [PubMed] [Google Scholar]

[22]. Shen C., Ding J., Wang Z., Cao D., Ding X., Hou T. From machine learning to deep learning: advances in scoring functions for protein–ligand docking. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2020;10(1) [Google Scholar]

[23]. Li H., Sze K.-H., Lu G., Ballester P.J. Machine-learning scoring functions for structure-based virtual screening. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2021;11(1) doi: 10.1002/wcms.1225. [DOI] [PMC free article] [PubMed] [Google Scholar]

[24]. Huang K., Fu T., Glass L.M., Zitnik M., Xiao C., Sun J. Deeppurpose: a deep learning library for drug–target interaction prediction. Bioinformatics. 2020;36(22–23):5545–5547. doi: 10.1093/bioinformatics/btaa1005. [DOI] [PMC free article] [PubMed] [Google Scholar]

[25]. K. Huang, T. Fu, W. Gao, Y. Zhao, Y. Roohani, J. Leskovec, C.W. Coley, C. Xiao, J. Sun, M. Zitnik, Therapeutics data commons: machine learning datasets and tasks for therapeutics, arXiv e-prints, 2021.

[26]. Zhu Z., Shi C., Zhang Z., Liu S., Xu M., Yuan X., Zhang Y., Chen J., Cai H., Lu J., et al. Torchdrug: a powerful and flexible machine learning platform for drug discovery. 2022. arXiv:2202.08320 preprint.

[27]. Kim S., Thiessen P.A., Bolton E.E., Chen J., Fu G., Gindulyte A., Han L., He J., He S., Shoemaker B.A., et al. Pubchem substance and compound databases. Nucleic Acids Res. 2016;44(D1):D1202–D1213. doi: 10.1093/nar/gkv951. [DOI] [PMC free article] [PubMed] [Google Scholar]

[28]. Consortium U. Uniprot: a hub for protein information. Nucleic Acids Res. 2015;43(D1):D204–D212. doi: 10.1093/nar/gku989. [DOI] [PMC free article] [PubMed] [Google Scholar]