

Review

Revolutionizing Drug Discovery: A Comprehensive Review of AI Applications

Rushikesh Dhudum¹, Ankit Ganeshpurkar²  and Atmaram Pawar^{3,*}

¹ Department of Drug Regulatory Affairs, Poona College of Pharmacy, Bharati Vidyapeeth (Deemed to be University), Pune 411038, India

² Department of Pharmaceutical Chemistry, Poona College of Pharmacy, Bharati Vidyapeeth (Deemed to be University), Pune 411038, India

³ Department of Pharmaceutics, Poona College of Pharmacy, Bharati Vidyapeeth (Deemed to be University), Pune 411038, India

* Correspondence: atmaram.pawar@bharativedyapeeth.edu

Abstract: The drug discovery and development process is very lengthy, highly expensive, and extremely complex in nature. Considering the time and cost constraints associated with conventional drug discovery, new methods must be found to enhance the declining efficiency of traditional approaches. Artificial intelligence (AI) has emerged as a powerful tool that harnesses anthropomorphic knowledge and provides expedited solutions to complex challenges. Advancements in AI and machine learning (ML) techniques have revolutionized their applications to drug discovery and development. This review illuminates the profound influence of AI on diverse aspects of drug discovery, encompassing drug-target identification, molecular properties, compound analysis, drug development, quality assurance, and drug toxicity assessment. ML algorithms play an important role in testing systems and can predict important aspects such as the pharmacokinetics and toxicity of drug candidates. This review not only strengthens the theoretical foundation and development of this technology, but also explores the myriad challenges and promising prospects of AI in drug discovery and development. The combination of AI and drug discovery offers a promising strategy to overcome the challenges and complexities of the pharmaceutical industry.

Keywords: artificial intelligence; machine learning; deep learning; drug discovery



Citation: Dhudum, R.; Ganeshpurkar, A.; Pawar, A. Revolutionizing Drug Discovery: A Comprehensive Review of AI Applications. *Drugs Drug Candidates* **2024**, *3*, 148–171. <https://doi.org/10.3390/ddc3010009>

Academic Editor: Osvaldo Andrade Santos-Filho

Received: 24 November 2023

Revised: 18 January 2024

Accepted: 6 February 2024

Published: 13 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Drug discovery and development is a multifaceted, ravelled endeavour. This intricate process encompasses four pivotal phases, beginning with the identification and validation of targets, progressing through compound screening and lead optimization, and culminating in preclinical investigations and clinical trials. The identification of pathophysiological factors and biological targets is the first step of this process. Bioinformatics, genomics, and proteomic research are necessary to unravel cellular and genetic targets. Initially, the first molecule, i.e., hit, that exhibits activity against the specified target is identified. This can be achieved using chemical libraries or by isolating natural products from bacteria, plants, and fungi [1]. In the next step, the lead molecule that has favourable potential for the development of a drug is identified. Lead optimization is the process of further modifying the selected lead to increase its specificity and effectiveness even at lesser doses. The functional properties of newly synthesized therapeutic candidates are improved through an iterative cycle combining cellular assays and structure–activity relationships. The animal models are then used for in vivo studies, including pharmacokinetic analyses and toxicity assessments. The drug candidate is finally administered to patients in clinical trials after all preclinical testing. Clinical trials play a fundamental role in determining whether the drug candidate can deliver the intended medical benefits while ensuring the well-being of the patients involved [2]. The process is time-consuming and cumbersome.

Therefore, pharmaceutical companies seek methods to decrease expenses and speed up their initiatives. Artificial intelligence (AI) is the ability of a machine to imitate the cognitive activity of the human brain in problem-solving and learning [3]. Technology-based AI systems can replicate human intelligence using a range of cutting-edge tools and networks. AI-based technologies are increasingly being used at different stages of the drug discovery process to save time and improve profitability. These include real-time cell sorting, cell classification, calculation of compound attributes based on quantum mechanics (QM), computational organic synthesis, creation of new compounds, and a variety of other tasks [4–7]. Figure 1, illustrates the comparative analysis between conventional and AI-driven drug discovery timelines, providing a visual representation of the significant acceleration achieved by AI technologies in this critical domain.

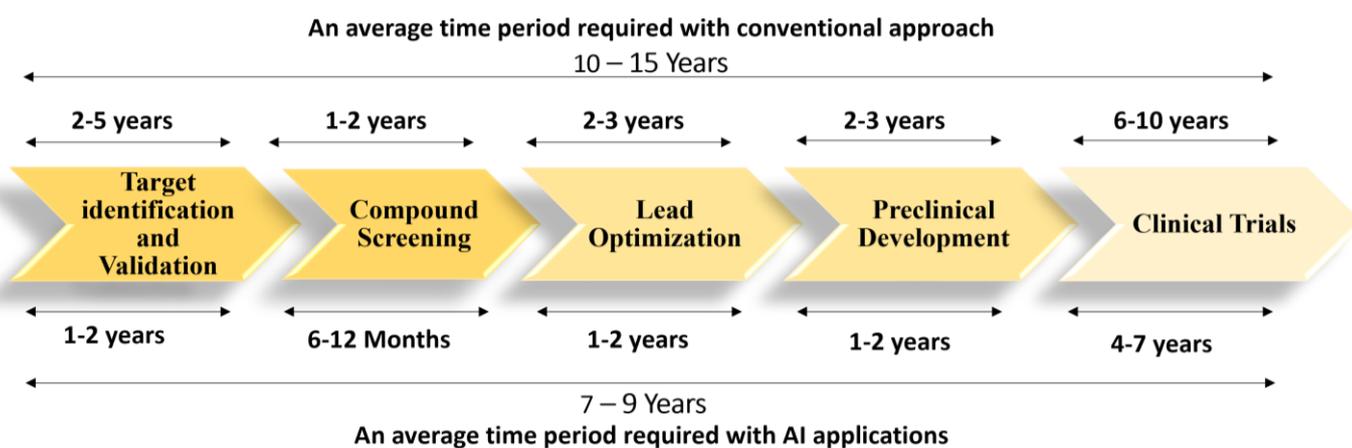


Figure 1. Impact of Artificial Intelligence on drug discovery: a comparative analysis of time required for key steps.

All phases of drug discovery and development use machine learning (ML) algorithms and software to hasten the process. McKinsey Global Institute predicts that rapid advances in AI-driven automation will fundamentally change in the way society works [8]. AI has recently gained importance in drug development as it serves to reduce the time and money spent on research, and the failure rate of clinical trials. Many AI companies are focusing on drug development due to the availability of big data for life sciences and the rapid development of ML techniques [9]. The applications of ML are now visible in novel target identification, deciphering disease-target associations, small-molecule design and optimization, uncovering disease mechanisms, understanding of disease and non-disease phenotypes, identification of unknown biomarkers for prognosis, and many more [10–14]. In recent years, the pharmaceutical industry has undergone significant data digitization, but it is difficult to acquire, verify, and use this knowledge [15]. AI can process large amounts of data through better automation and this encourages its adoption [16]. Numerous ML algorithms such as Support Vector Machine (SVM) and Random Forest (RF) have been employed in drug design to bridge the gap between random and rational drug discovery [17,18]. Hence, an increasing number of medicinal chemists have been using various AI techniques to address the fundamental problem of evaluating and predicting the biological effects of compounds [19]. The pattern recognition approach has gained special attention, and it focuses on explaining and exploring common patterns of chemical entities. It operates under the fundamental assumption that compounds sharing analogous structural formulas are likely to exhibit comparable physicochemical properties and in vitro biochemical effects [20,21].

2. The Drug Discovery Process: A Brief Overview

Various diseases or clinical conditions have no effective treatment available on the market; hence, drug development programmes still need to address these unmet clinical

needs. Early academic research often results in the identification of proteins or pathways that may have therapeutic benefit; however, the selection of a target may require further validation before drug development proceeds. Once clues are discovered, in-depth research is conducted to identify a biological or small-molecule drug candidate. It undergoes preclinical testing and clinical trials to become a commercial drug [22]. From discovery to approval, the expensive drug development process often takes more than a decade. It should be noted that the associated costs and approval times may vary [23]. Computational approaches have long played a key role in drug development and discovery. However, there are still some problems associated with traditional computing methods, including time, computing cost and reliability [24,25]. Molecular dynamics (MD) simulation is used in drug discovery to evaluate protein–ligand interactions and binding stability. One of its major disadvantages is that it can be extremely laborious and time-consuming [26]. Currently, large-scale high-throughput testing methods and randomized strategies are widely used in medicinal chemistry [27]. These methods examine many possible molecules to identify molecules with the required properties. However, these techniques can be expensive and cumbersome and may often produce inaccurate results [28]. Similarity-based approaches, e.g., ligand-based virtual screening (LBVS), are indeed characterized by significant “selectivity” in identifying various active ingredients. Its disadvantage in “real” applications is its tendency towards false-positive detections [29]. The reliance of drug discovery on trial-and-error testing results in inaccurate prediction of potential new bioactive substances. Hence, the limited traditional pharmaceutical testing techniques have proven to be effective in the past [30]. To accelerate the discovery process, industrial research and development (R&D) continually evaluates new technologies across the various research disciplines. The ever-increasing costs and delays in the entire process pave the way for investments in highly complex technologies, improved production methods and innovative research methods for tangible successes [31]. Various AI-based algorithms such as reinforcement learning, scalable, rule-based, and supervised and unsupervised techniques could help to solve these problems [32,33].

3. Basics of AI and Machine Learning (ML)

3.1. Artificial Intelligence

AI is dedicated to the computational comprehension of intelligence and the development of entities that are capable of demonstrating these actions [15]. These entities/machines can also carry out activities that need human intellect. The process involves gathering data, creating rules for applying it, coming to general or precise conclusions, and then engaging in self-correction [34]. The key benefit of an AI-based strategy is that it learns from instances and can create a model, even when we have a poor understanding of the underlying mechanism. Healthcare uses AI in the detection and treatment of several ailments. Some of the examples include IBM Watson for oncology, PathAI for pathology diagnosis, Tempus for cancer treatment, and Insilico Medicine for drug discovery [35]. The pharmaceutical community has experienced a tremendous rise in data digitalization in recent years. However, with the advent of digitalization, it has become more difficult to gather, evaluate, and apply this knowledge to complicated clinical issues [15]. This encourages the adoption of AI since it can automatically process massive volumes of data [16]. Thus, AI consists of several cutting-edge devices and networks that may imitate human intellect [36,37].

Computer modelling based on AI and ML offers excellent opportunities for target identification, peptide synthesis, toxicity and property assessment; physicochemical studies of drugs, and monitoring of drugs in terms of their efficacy and their repositioning [38]. The description of some AI applications in the pharmaceutical industry is presented in Figure 2.

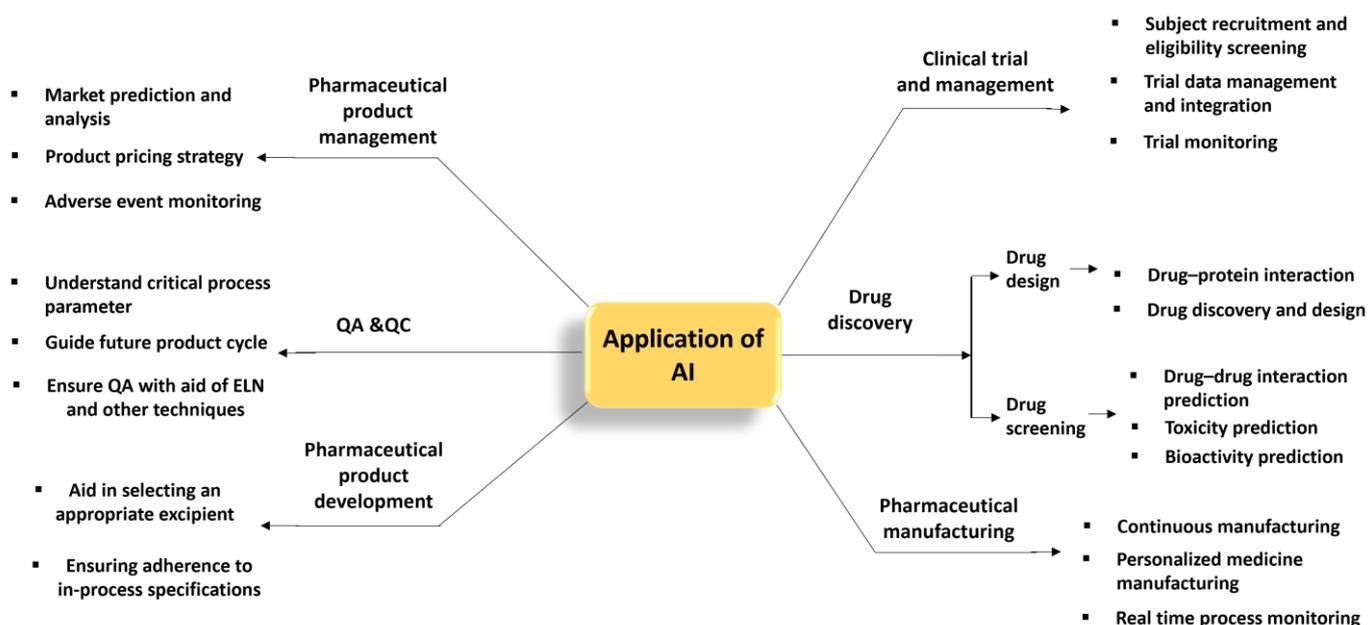


Figure 2. Application of AI in numerous subfields of pharmaceutical industries, including pharmaceutical product management and drug discovery.

3.2. Machine Learning

ML is about training computers to make sense of data and make decisions based on patterns and examples. This may be supervised learning, reinforcement learning, or unsupervised learning [39]. Any data may be employed to develop prediction models using ML. To build a model, the dataset must first be cleaned, outliers should be eliminated, and missing values must be imputed. Data selection, feature extraction, key attribute and algorithm selections, model development, and model evaluation are crucial steps [40,41].

3.3. Supervised Machine Learning

Employing labelled datasets to train algorithms for data categorization or accurate result forecasting is known as supervised ML. Once input is provided, the model uses a reinforcement learning (RL) technique to update the weights to ensure the model fits correctly. Numerous industries and organizations use supervised learning to address a wide range of practical issues. The two types of supervised learning are classification and regression [42].

3.4. Unsupervised Machine Learning

When there is an input variable but no output variable then unsupervised learning is utilized. Its primary objective is to comprehend the distribution of data to obtain a better understanding. It can be further divided into association and clustering [42].

3.5. Reinforcement Learning

RL is an intriguing subdomain of ML that has captured substantial interest within the academic and commercial realms. In contrast to supervised and unsupervised learning, RL is a method of continuous learning that preserves autonomy. RL can react swiftly to changing settings; it is employed in trading, gaming, and robotics [43]. In certain situations, supervised learning outperforms RL in classification tasks; nonetheless, RL can constantly learn with little to no human involvement, which is desirable [44,45].

3.6. Deep Learning

A subtype of ML called deep learning (DL) involves the system learning on its own from unlabelled and unstructured data. Igor Aizenberg et al. used the phrase “deep learning” to describe an artificial neural network (ANN) in the late 20th century [46]. ANNs with numerous layers of processing are used in DL [39]. DL works particularly well with difficult jobs with enormous amounts of high-dimensional data. However, it involves more computations than conventional ML and is usually called an “AI black box” as the user cannot see many of the processing levels. Hence, there are additional difficulties and dangers in understanding the model. Large Language Models (LLMs) are DL models that are trained on big datasets to produce human-like language. LLMs include chatbots based on pre-trained generative transformer (GPT) architecture such as ChatGPT and GPT-4 [47].

DL and ML algorithms are examples of AI which has emerged as a potential solution to issues and challenges in the process of drug discovery [35]. Additionally, there are several time-consuming and difficult phases involved in the discovery and development of drugs that can be handled with ease using AI [48] (Figures 3 and 4).

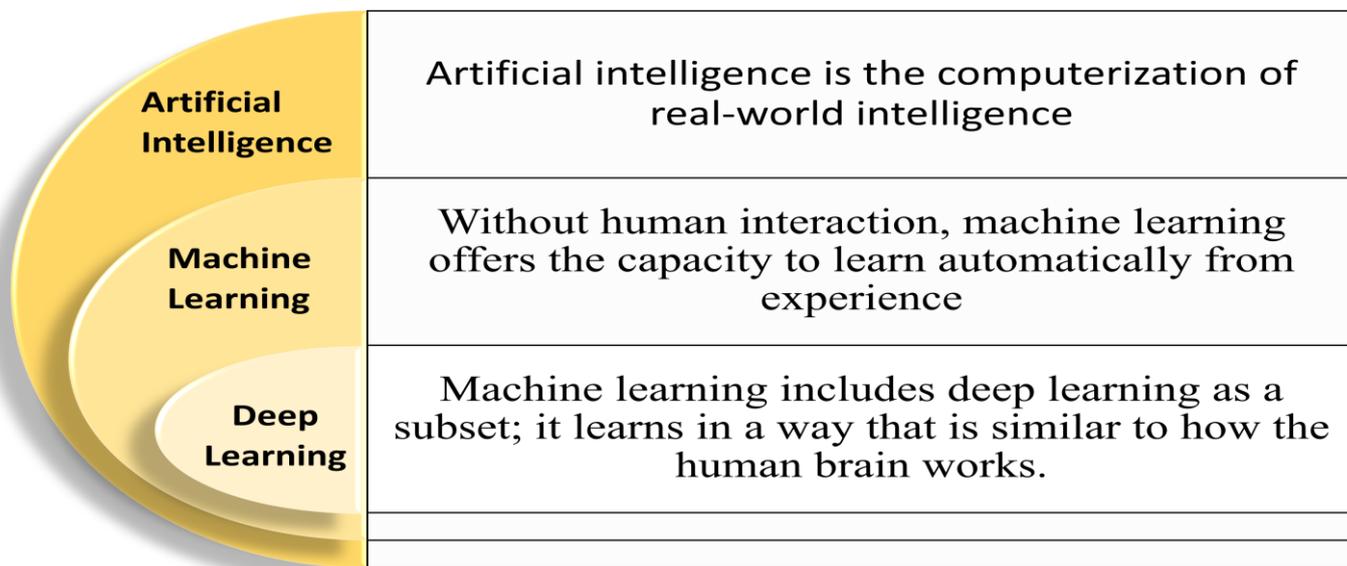


Figure 3. Machine learning and deep learning are subsets of artificial intelligence.

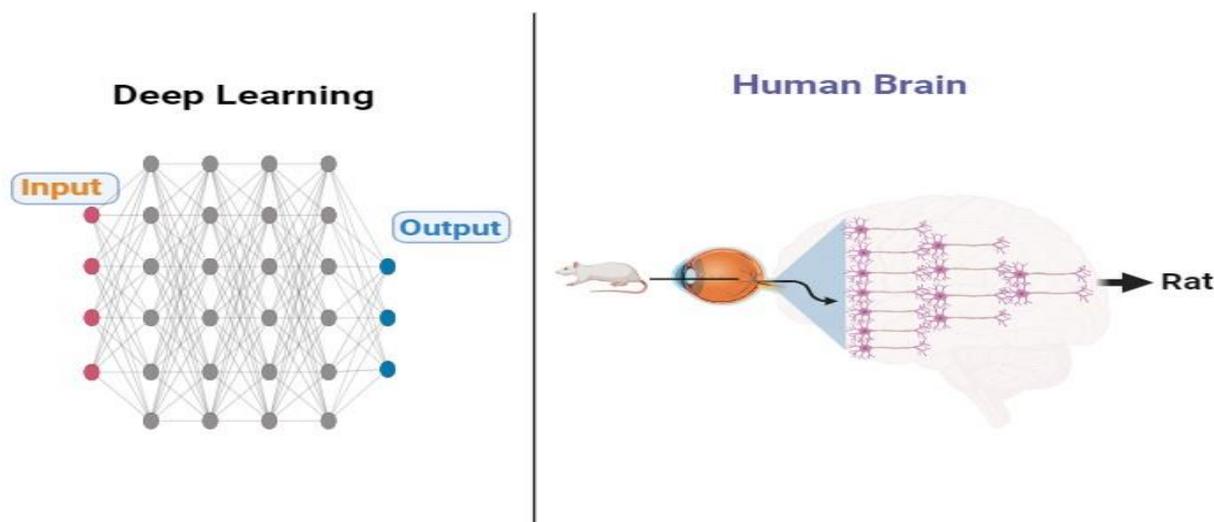


Figure 4. Comparative analysis: Deep learning mechanisms and human brain functions.

4. AI in Drug Target Identification

Drug development is a long, expensive, and risky process that requires more than 10 years and USD 2 billion to bring a new drug to market [49]. Fewer than 500 effective drug targets had been discovered by 2022, which is a negligible portion of the estimated druggable targets in humans [50–52]. Target identification is one of the most important processes to determine the biological cause of a disease and provide effective therapies [53]. It is the process of selecting appropriate biological molecules or cellular pathways that can be altered by drugs to achieve therapeutic benefits. The availability of biomedical data has increased in recent years, ranging from basic research into the causes of disease to clinical studies. However, this large amount of information poses challenges for data analysis in terms of scalability, data integration, data quality, noise, computational complexity, interpretability, and validation. AI can manage and analyse such complex networks of biological data. Recently, a promising method for target identification has been developed that combines multi-omics data with AI algorithms [54]. Puna et al. combined various bioinformatics and DL-based models trained using disease-specific text and multi-omics data to prioritize treatable genes and identify potential therapeutic targets in amyotrophic lateral sclerosis (ALS), yielding 18 potential therapeutic targets for ALS. Leveraging more than 20 AI and bioinformatics models, PandaOmics assesses targets by considering their associations with specific diseases, along with information on druggability, developmental stage, and tissue specificity. By combining omics AI scores, text-based AI scores, finance scores, and key opinion leaders (KOL) ratings, PandaOmics could forecast the target genes linked to a particular disease through the use of sophisticated DL models and AI techniques [55]. Validation of proposed AI therapeutic targets for ALS in an ALS-mimicking drosophila model revealed eight unreported targets whose elimination significantly reverses ocular neurodegeneration. Zhang et al. also developed a ML-based technique to recognize KANK1 as a novel ALS-related gene in the same therapeutic area and confirmed the neurotoxic consequences of KANK1 mutations replicated by CRISPR-Cas9 in human neurons [56]. DeepDTnet was developed by Zeng et al. to aid in the *in silico* discovery of molecular targets for already approved drugs. It is based on 15 different types of chemicals, genomic, phenotypic, and cellular networks. In a mouse model of multiple sclerosis, one of the discovered drugs targeting human ROR- γ t demonstrated specific therapeutic benefit [57].

AI has received a lot of attention, and ML-based algorithms, especially AI techniques, have produced remarkable results in the pharmaceutical industry [58]. The modern DL designs such as generative adversarial networks (GANs), recurrent neural networks (RNNs), and transfer learning approaches are attracting more attention and are used in many more healthcare applications than older ML techniques [59]. AI-powered Large Modules (LMs) help to speed up target identification and simplify searches. In addition to using a multitask learning neural network with shared character and word layers (MTM-CW), automatic biomedical named entity recognition (BioNER) is a useful method for identifying chemicals, diseases, and genes that are embedded in free-text documents [60,61]. AI-driven LMs offer potential advantages, including the capability to analyse data and assist in the identification and prioritization of targets [60].

Furthermore, Fabris et al.—using genetic or protein traits such as gene ontology terms, protein–protein interactions, and biological pathways—developed a DL approach with a new modular architecture to identify human genes associated with several age-related diseases [62] (Figure 5).

Synthetic data are data that have been intentionally created to resemble real-world patterns and characteristics. AI algorithms can be used to generate synthetic data to replicate different biological circumstances, allowing scientists to explore and test a wider range of possible outcomes [63,64]. Additionally, the predictions of AI systems may be verified using synthetic data that provide higher assurance for the target discovery process. There are two groups of AI algorithms; i.e., ML-based bioanalysis, and network-based bioanalysis algorithms. These are commonly used for predictive cancer target identification

and drug development. However, even after the use of ML algorithm, there are still three significant obstacles to target identification and drug development for cancer [65]. The difficulties are the lack of reliable data for validation, integration of disparate information, and DL models that are difficult to understand [66–68]. Although there has been significant improvement in this field, a lot more progress is required.

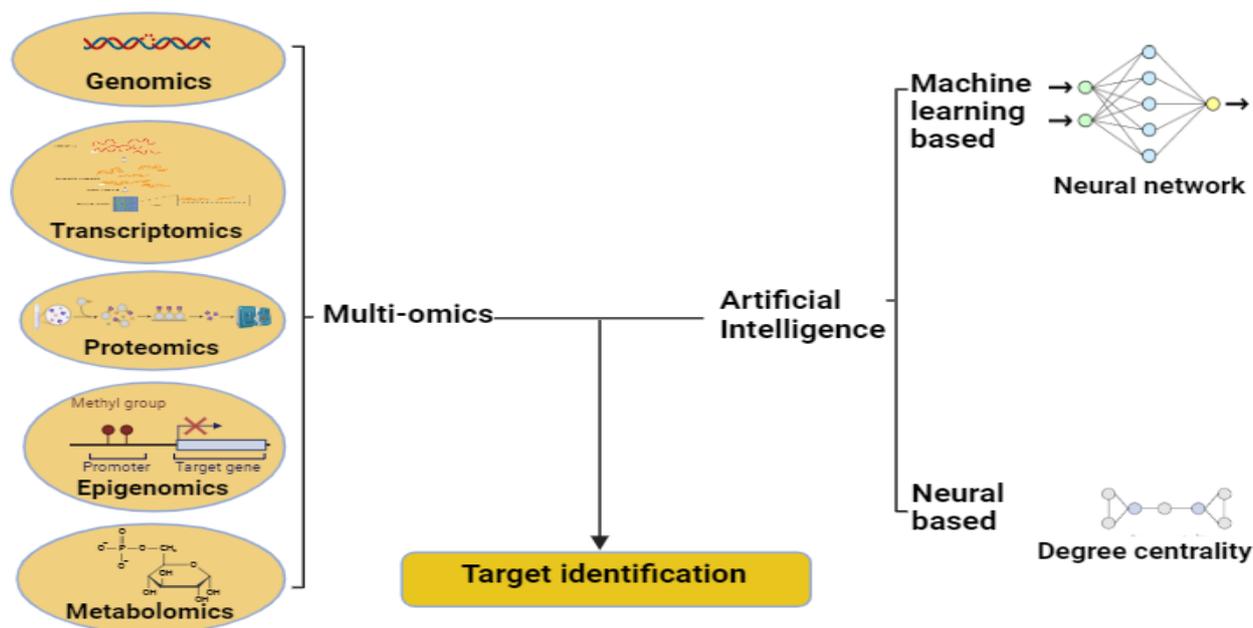


Figure 5. Artificial intelligence employs machine learning and neural network techniques to effectively identify targets. Figure created with BioRender.com.

5. AI for Insightful MD Data Analysis

Molecular dynamics (MD) are extensively employed in the field of drug discovery, playing a crucial role in understanding molecular interactions and predicting drug behaviour. Its research applications include lipid membrane permeability, protein–ligand binding, protein–protein interactions, and partition coefficients. The advantage of MD is the ability to determine thermodynamic characteristics such as the free energy of binding in atomistic detail [69]. Based on a general model of physics that governs interatomic interactions, MD predict how each atom in a protein or other molecular system will move over time [70]. The fundamental concept of MD is to determine the force that all other atoms in a biomolecular system exert on each atom by knowing the position of all. To determine the forces in an MD simulation, a model called a molecular mechanical force field is used, which is often adjusted to the results of some experimental observations and the results of quantum mechanical calculations. Additionally, these simulations can predict how biomolecules respond to changes such as mutation, phosphorylation, protonation, or the addition or removal of a ligand at the atomic level [71]. Traditional MD tools such as GROMACS 2021, Amber 20, NAMD 2.14, and GROMOS 56a7 are well established and used to study interactions between disease targets and drugs [72–75]. With increased computing power and advanced software, MD is an effective method for studying systems of biomolecules and ligands; however, it has to deal with enormous amounts of data and extremely complex simulated systems [76]. The massive amounts of data generated by MD can be processed quickly and efficiently using DL techniques. To increase the efficiency and speed of analysing MD data, DL can transform the high-dimensional structure space into a low-dimensional feature space. Surface extraction and free energy kinetics, force fields, coarse-grained molecular dynamics, thermodynamics, and other applications are just a few of the many applications of DL in molecular simulations. ML models are important tools in the field of MD. With its exceptional ability to predict chemical qualities

such as binding affinity and solubility, RF models are a useful tool for managing intricate, high-dimensional data. MD trajectories are analysed using RNNs, which capture temporal relationships and provide insights into dynamic processes. Because they are good at evaluating chemical structures, graph neural networks (GNNs) are a powerful tool for tasks such as drug-target binding and protein-protein interaction prediction. Meanwhile, the creation of new molecular structures is facilitated by GANs, which accelerates progress in materials research and drug development. These ML models including GANs, GNNs, RNNs, and RFs have transformed the field of molecular dynamics and offer invaluable resources for comprehending and investigating molecular behaviour [77]. Plante et al. used a DL technique to process large MD trajectories and classify GPCR ligands into full, partial, or reverse agonists with good accuracy. The red, green, and blue (RGB) pixels represent the X, Y, and Z coordinates of the protein atoms in the MD trajectory. The confusion matrix was used to train the neural network to predict the classification class labels of GPCR ligands. This work provided an accurate and efficient method to distinguish ligand classes and dynamic pattern differences between two GPCR targets by converting the associated MD trajectory information into image pixel representations to train a deep learning model [78]. Marchetti et al. employed the machine learning technique to classify allosteric modulators into activators or inhibitors based on conformational ensembles obtained from simulated MD trajectories of the molecular chaperone Hsp90. It shows how the machine learning technique can be expanded to explore multiple targets for allosteric diseases and can be used to develop potential allosteric drugs [79]. To develop effective cementitious binders, Lyngdoh et al. used ML with MD to study the interactions between calcium silicate gel and hydrate. Interpretability was particularly important because the best model they could obtain was an artificial neural network that yielded a probability vector. The importance of each input feature to the final prediction was determined using the SHAP approach [80].

In molecular docking, the scoring factor or scoring function (SF), is a mathematical function used to evaluate the binding affinity between a ligand and a receptor. Van der Waals forces, electrostatic interactions, hydrogen bonds, and hydrophobic interactions are only a few of the molecular interactions that are taken into account by the SF factor. Through its guidance of lead identification, virtual screening, and ligand optimization activities, it plays a critical part in the drug development process. When comparing scoring criteria between ML and traditional methodologies, ML approaches have several benefits over conventional methods. These techniques utilize sophisticated algorithms, such as neural networks or RF, to identify intricate patterns and correlations within the data. More accurate predictions are made possible by their ability to manage high-dimensional feature spaces and capture non-linear correlations. Conventional methods, on the other hand, rely on pre-established mathematical formulas that are derived from empirical or physical principles. Conventional methods may not be able to fully capture the intricacy of molecular interactions, despite their widespread usage and insightful contributions. In 2017, Ragoza et al. developed a technique using CNN scoring algorithms to automatically learn the basic binding-related properties of the PLI, given the full 3D description of the PLI as input. Their CNN scoring functions were optimized to distinguish known binders from non-ligands, and between valid and invalid binder positions [81]. A ML-based grading mechanism was created by Nguyen et al. in 2018 to choose the postures produced by GOLD, GLIDE, and Autodock Vina. Using a targeting ligand, they created a complex training dataset through WPB. They then docked the ligands again to the proteins of the selected complexes and used CNN to collect topological features, and implemented RF to understand biomolecular structure. The final predictions for this strategy were determined by combining the energy values predicted by these two machine learning algorithms [82].

By using ML to extract relevant information and understand complex interactions between molecular properties and binding affinities, ML technologies can overcome these constraints [83] (Table 1).

Table 1. Summary of machine learning methods used in scoring functions (SFs).

Sr. No.	Scoring Factor (SF)	Feature Representation	ML Model	Application/Pharmaceutical Example	Reference
1	RF-Score-VS	Terms from the three of RF-score versions	RF	VS	[84]
2	SIEVE-Score	Glide's per-residue interaction energy	RF	VS	[85]
3	Binding Affinity	Molecular description, protein–ligand interaction	RF	Drug binding prediction, VS	[72]
4	RF-Score	Pair counts for elemental atoms	RF	Scoring	[86]
5	SVM-SP	Knowledge-driven pairwise possibilities	SVM	Target-specific VS	[87]
6	Pharm-IF	Fingerprints of pharmacophore-based interaction	SVM, RF, ANN	Target-specific VS	[88]
7	PADIF-SVM	Fingerprints of interaction derived from GOLD	SVM, ANN	Docking based-target prediction	[89]
8	SVR-KB	Knowledge-based potentials for pairs	SVM	Scoring	[90]
9	K _{DEEP}	Specialized property channel	CNN	Scoring	[91]
10	Binding Activity, Affinity	Molecular descriptors	XGBoost	Drug Affinity prediction, QSAR modelling	[92]
11	RI-Score	Index of stiffness unique to an element	RF	Scoring	[93]
12	Bioactivity	Molecular fingerprint	SVM	Drug Design	[94]
13	ID-Score	50 terms linked to protein–ligand interaction	SVM	Scoring	[95]
14	Drug Classification	Chemical features, compound properties	Logistic Regression	Drug Categorization	[96]
15	EIC-Score	EIC or elemental interactive curvatures	GBDT	Scoring	[97]
16	NNScore	Information driven paired potential	ANN	VS	[98]
17	CScore	Pairs of elements are converted by two fuzzy membership functions	Modified CMAC network	Scoring	[99]
18	TNet-BP	Topological fingerprints	CNN	Scoring	[100]
19	Deep-VS	Types of amino acids, charges, distance, and atom types	CNN	VS	[101]

Abbreviations: RF, Random Forest; VS, Virtual Scoring; SVM, Support Vector Machine; CNN, Convolutional Neural Network; ANN, Artificial Neural Network; GBDT, Gradient Boosted Decision Tree.

6. Compound Screening with AI

The aim of drug discovery is to identify small molecules that can alter the function of the target protein and disease phenotypic characteristics. At the same time, there is also a need to search for small molecules with good pharmacokinetic properties and minimal toxicity. The identification of drug candidates, their validation, pharmacokinetics, and pre-clinical toxicity assessment are difficult, time-consuming, and expensive processes. It takes an average of 10 to 12 years for a drug to reach the market, and between USD 800 and USD 1.8 billion is spent on each successfully discovered drug [102,103]. The application of AI in drug screening began a revolutionary era in pharmaceutical research, significantly reducing R&D costs by 50% while increasing efficiency and accuracy [104]. AI addresses various challenges related to drug screening, including predicting physicochemical properties and assessing bioactivity and toxicity.

6.1. Primary Drug Screening: Enhancing Cell Classification and Sorting

AI can help to ease the load of repetitive and challenging tasks, making the drug development process faster. These tasks involve various activities such as sorting cells, calculating properties of small molecules, using computers to create organic compounds, designing new compounds, developing tests, and predicting the 3D shapes of target molecules [105]. Accurate identification of cell type is possible using AI-based approaches, particularly the least square vector support method (LS-SVM). With a classification accuracy of 95.34%, this method makes a significant contribution to high-throughput automated cell sorting techniques [4,105,106]. In addition, the interpretation of computerized electrocardiography (ECG), which is a fundamental step in the clinical processes of diagnosis and treatment, has shown the promise of AI [107].

Based on a CNN architecture, Nitta et al. developed the Intelligent Image-activated Cell Sorting system (iIACS), which can sort cells in real time based on cell pictures. The system combines a double membrane pump, a two-stage microfluidic chip with 3D hydraulic focusing, and a high-speed fluorescence microscope to enable automatic liquid focusing, cell sorting, and real-time detection. The intercellular contacts and intracellular localization of proteins allows sorting of blood cells and microalgae in real time. To obtain higher-quality cell images and reduce processing time, researchers modified the iIACS-based cell imaging approach with an image sensor-based optomechanical flow imaging method and improved the system hardware [4].

6.2. Secondary Drug Screening

6.2.1. Predicting Physicochemical Properties

AI is ideal for secondary drug screening, where predicting physicochemical properties is crucial for drug development. Deep neural networks (DNNs) algorithms that use chemical descriptors such as SMILES strings and potential energy measurements are used to generate viable compounds [7,108]. To predict chemical properties, these networks are divided into generation and prediction phases. Each stage is trained independently through supervised learning [7]. The study highlights the importance of considering physicochemical parameters such as melting point and logP when selecting the best drug candidates. Using data from the Environmental Protection Agency (EPA) and the Estimation Program Interface (EPI) package, Yang et al. developed a quantitative structure–property relationship (QSPR) method to determine six physicochemical properties of environmental pollutants. These include the bioconcentration factor (BCF), log P, log S, melting point (MP), boiling point (BP), and vapour pressure (log VP). The lipophilicity and solubility of various substances were predicted using neural networks based on ALGOPS software (version 2.1) and ADMET predictor. It has been shown that it is possible to predict the solubility of compounds using DL techniques such as graph-based convolutional neural networks (CVNN) and recurrent neural networks on undirected graphs [109]. Panapitiya et al. evaluated various molecular representation techniques (such as molecular descriptors, SMILES, molecular graphs, and 3D atomic coordinates) and deep learning techniques (such as fully connected neural networks, RNNs, graph neural networks, and SchNet) for solubility prediction. According to the authors' results, which were based on the same test dataset, the fully connected neural network outperformed other neural networks in predicting solubility with chemical descriptors. In addition, the researchers examined the importance of different variables in prediction and found that two-dimensional molecular descriptors made the greatest contribution [110].

6.2.2. Predicting Bioactivity: Optimizing Compound Activity

ML techniques, including matched molecular pairs (MMPs), RF, gradient boosting machines (GBMs), and DNNs, are used to predict chemical bioactivity. The performance of RF and GBMs is outperformed particularly well by the combination of MMP with DNN [101,111,112]. This method can predict multiple bioactivity characteristics including oral exposure, intrinsic clearance, ADME, and mode of action, which helps in decision-making in drug development [113–115]. Koutsoukas et al. used a multilayer perceptron model based on molecular fingerprint representation of the compounds to predict various bioactivity (pKi and pIC₅₀) against seven targets, including two GPCRs, cannabinoid receptor 1 and dopamine receptor D4. Additionally, the researchers found that MLP performs better than traditional ML techniques on large datasets. However, they pointed out that DL models require much finer tuning of hyperparameters to achieve high predictive performance [116]. Stokes et al. proposed a direct message neural network capable of predicting antimicrobial activity. They generated a feature vector for each molecule by first building a molecular graph based on its SMILES and then combining bonding information (e.g., bond type and stereochemistry) and atomic features (e.g., the number of bonds for each atom and the atom number) used. The optimized feature vector was fed into a

feedforward neural network, which acquired the antibacterial properties of the chemical by repeatedly performing the message transmission process [117].

6.2.3. Toxicity Prediction: Mitigating Risk through AI

AI-based toxicity prediction plays a key role in evaluating drug safety. The DeepTox algorithm is effective in correctly assessing the toxicity of a substance [118]. This capacity to forecast could prevent probable side effects during drug development. The application of open-source technologies such as TargeTox and PrOCTOR enhances researchers' proficiency in the identification and resolution of toxicological issues, thereby advancing the field of toxicity prediction [119,120]. However, the literature lacks essential information, even for basic molecular structures. This absence hinders the ability to conduct a fundamental environmental assessment for a therapeutic molecule's synthesis path. To address this, simple algorithms use substance interpolation based on known toxicity and similar structures. By leveraging AI expertise, researchers can efficiently discover target-specific compounds, predict physicochemical properties, assess bioactivity, and predict compound toxicity.

7. Drug Design and Optimization

The aim of drug design is to obtain small molecules that could meet various criteria, including efficacy for pharmacological purposes, an appropriate safety profile, appropriate chemical and biological properties, and sufficient innovation to secure intellectual property rights for commercial success, etc. [121]. The computational tools have been crucial to drug discovery and have completely changed the way drugs are designed. There are still several problems associated with traditional computational techniques, including input time, computational cost, and reliability [24,25]. AI could overcome all the obstacles associated with computational drug development, thereby increasing the usefulness of computational techniques in drug development.

7.1. AI in Molecular Design

The process of automatically suggesting novel chemical structures that align most effectively with a specified molecular profile is referred to as "de novo molecular design" [122]. A virtual chemical library is created for computational testing followed by synthesis, and characterization. The variational autoencoder and GANs are two significant technological advances in deep generative modelling [123,124]. "Molecular Autoencoder", studied by Gómez-Bombarelli et al., was the first to demonstrate detailed generative modelling of molecules. An interesting method is the Variational Autoencoder, which connects an encoder network with a decoder network consisting of two neural networks [33]. The encoder network transforms the chemical structures defined in SMILES into a latent space represented by a real-valued continuous vector. The decoder component can convert latent spatial vectors into chemical structures. De novo design has also been successfully applied through RNNs. They were originally developed in the field of natural language processing [125]. Sequential information is fed into the RNN. The most commonly used models for modelling and creating sequences are RNNs [126,127]. A GAN was successfully employed by Kadurin et al. to recommend potential anticancer drugs [128]. Assmann et al. discussed the practical difficulties in using de novo design to facilitate the identification of new CDK9 inhibitors. The use of the molecules suggested by the molecular generator as seeds in the similarity search of the Enamine REAL library is characterized by an improved variation of the vs. approach. Of the 69 compounds tested, seven were active against CDK9 [129]. Perron et al. have published another useful example of applying generative methods to find optimum solutions to multiparameter objectives using an RNN-based generative model [130]. Li et al. looked at the potential of RNN-based de novo design techniques to produce new molecule inhibitors in chemical space that has been thoroughly researched. Scientists announced their search for new inhibitors of CDK4 kinases and the well-studied proto-oncogene protein serine/threonine kinase 1 (PIM1). After testing

four different drugs, they identified a potent PIM1 inhibitor and two key CDK4 inhibitor compounds [131].

7.2. Predicting Pharmacokinetics and Pharmacodynamics

The key concepts of pharmacology include pharmacokinetics and pharmacodynamics. While pharmacodynamics focuses on how a drug works in the body and how it affects other systems in the body, pharmacokinetics deals with the study of drug absorption, distribution, metabolism, and elimination (ADME) [132]. The application of AI techniques in pharmacokinetics and pharmacodynamics has created new opportunities to improve drug development and personalized treatments. It can analyse complex datasets, identify trends and make predictions that could improve patient outcomes, improve drug delivery and minimize side effects [133,134]. ML and DL techniques are widely used to predict pharmacokinetic parameters. Numerous ML techniques—including Bayesian model, RF, SVM, ANN, and decision tree—have been used to predict the ADME of drugs. To predict various pharmacokinetic parameters such as drug absorption, bioavailability, clearance, volume of distribution, and half-life, DL algorithms such as Convolutional Neural Networks (CNN), Short-Term Memory (LSTM), and RNN are often used. A computational method called quantitative structure–activity relationship (QSAR) uses the chemical structure of a molecule to predict its biological activity [135–137]. With improved training data, a 47th version of admetSAR 2.0 is now available. This program also includes a module called ADMETopt, which is used to optimize lead activity based on expected ADMET attributes [138]. AI techniques facilitate the modelling of drug–receptor interactions and the prediction of drug efficacy and toxicity in the field of pharmacodynamics. The use of AI in pharmacokinetics and pharmacodynamics can significantly accelerate the drug discovery process and improve precision medicine [134,139]. Obrezanova et al. used conventional ML techniques and multitask convolutional neural networks to calculate time-dependent pharmacokinetic profiles and nine *in vivo* pharmacokinetic parameters in rats (oral and intravenous administration) based on *in vitro* measured ADME properties and molecular chemical structures of 3000 different compounds [140]. Ye et al. used transfer learning and multitasked learning to pre-train the model on over 30 million bioactivity data points. The model was then used to estimate four human pharmacokinetic parameters: oral bioavailability, plasma protein binding, V_d , and half-life, for 1104 FDA-approved small-molecule drugs. Compared to other traditional ML techniques, their DL model showed the highest performance (although not always by a significant margin) and generalization ability, achieving a mean absolute error or MAE = 0.31 for oral bioavailability and MAE = 0.17 for V_d [141]. Interestingly, Lou et al. created a model that predicts the bioavailability of mAbs administered through subcutaneous preparation in humans. A dataset of 45 clinical mAbs—with sequence and structure-based features including isoelectric point, total charge, aggregation propensity, solubility score, surface hydrophobicity spots, positive charge, and negative charge (with a threshold of 70% bioavailability)—were used to build a classification model. The study used a range of traditional Scikit-Learn ML techniques such as Adaptive Boost, Multilayer Perceptron, RF, and SVM. Among them, the tree approach showed the highest accuracy, reaching 78% [142].

Two areas that benefit greatly from the implementation of AI algorithms are drug design and optimization. De novo design, virtual screening, and structure-based drug design are just a few examples of these algorithms. The application of AI to drug development and optimization has a transformative impact on the discipline, enabling the rapid discovery of new therapeutic candidates and the more targeted and effective exploration of chemical space. Using ML, DL, and computer modelling methods, AI models can provide accurate predictions about the properties, interactions, and behaviours of potential drug candidates (Table 2).

Table 2. Some common machine learning models are used in pharmacokinetic and pharmacodynamic parameters prediction and toxicology prediction.

Sr. No.	Application	ML Model	Description	Reference
1	PK Modelling	Hidden Markov Models	Dynamic systems are represented using Sequential model.	[143]
2	PD Modelling	Logistic Regression	Utilized pharmacodynamics for binary categorization.	[144]
3	PD Modelling	Random Forest	For both regression and classification, use the ensemble approach.	[145]
4	PK and PD Modelling	Decision Tree	Nonlinear modes of decision-making for the classification of data.	[146]
5	PK and PD Modelling	Support Vector Machine (SVM)	Useful for data with non-linear connections.	[147]
6	Toxicity Prediction	Naive Bayes	A straightforward probabilistic model is frequently employed in toxicology to classify texts.	[148]
7	PK and PD Modelling	Neural Network (Deep Learning)	Sophisticated model for PK and PD forecast.	[149]
8	PK and PD Modelling	Gradient Boosting	Ensemble method that combines weak learners.	[150]
9	Toxicity Prediction	Random Forest	An ensemble approach to classify problems that can handle high-dimensional data and is resistant to over fitting.	[26]
10	Toxicity Prediction	Gradient Boosting	By merging weak learners, the ensemble learning approach improves model performance.	[151]
11	PK and PD Modelling	Bayesian Network	Graphical probabilistic models for parameter estimation	[152]

8. Assessing Drug Toxicity Using AI

The stringent safety requirements associated with drug development make it challenging to introduce new drugs to the market. Clinical trials often fail due to unexpected toxicity and post-marketing safety issues, resulting in unnecessary morbidity and mortality. Clinical trials test the safety and effectiveness of a drug before it is approved while pharmacovigilance continually verifies a drug's safety information during its usage in patients [153]. The establishment of pre-market drug safety has been shown to benefit significantly from the use of AI-based approaches, particularly in the area of toxicity assessment [154]. The vast reach of AI helps to predict the side effects, therapeutic targets, and in vivo safety of chemicals before manufacturing. Usually, after designing of the small molecule, the assays are employed to predict off-target toxicity, genotoxicity, organ toxicity, cytotoxicity, and mitochondrial toxicity [155]. The analysis of new types of data, including gene expression and cell imaging data, combined with knowledge of chemical structure, can now be used to predict the effects of in vivo toxicity [156]. Various in silico calculation methods have proven useful in calculating the toxicity of drug candidates. These methods, which include target-based predictions and QSARs, evaluate multiple pharmacological properties to predict toxicity. Various drug safety effects—such as skin/eye irritation, tissue-specific toxicity, and 50% lethal drug dose (LD₅₀) values—were modelled using QSAR techniques [157]. In particular, the QSAR model allows for examining the relationship between multiple predictors (e.g., molecular features) and responses (e.g., biological activities such as binding affinity) [158]. Early QSAR approaches assessed the chemical properties of drug candidates using multivariate linear regression [159]. Due to their excellent prediction accuracy, robustness, and readability of ensemble techniques such as RF and SVMs, they are currently the most popular options [160]. Compared to Naive Bayes, k-Nearest Neighbour (k-NN) and RF algorithms, SVM showed better performance in predicting activity values in the latest QSAR modelling of histone deacetylase (HDAC) inhibitors [161]. In addition, with the help of such QSARs, it is possible to predict activity based on objectives such as toxicity. Recently, Minerali et al. created and compared ML algorithms to predict drug-induced liver injury (DILI) using the company's Assay Central software. To do this, they used data previously collected by research teams at Pfizer and AstraZeneca, as well as data from

the FDA. The best Bayesian model based on the DILI problem category from the DILI Rank database produced results with an ROC of 81%, a sensitivity of 74%, a specificity of 76%, and an accuracy of 75% [162]. Williams et al. used ML to predict DILI with the pharmaceutical company, AstraZeneca. They were able to quantify the risk of an association being classified as low, medium, or high with an accuracy of 63%. The model provided an accuracy of 86%, a sensitivity of 87%, a specificity of 85%, a positive predictive value of 92%, and a negative predictive value of 78% for binary (yes/no) DILI prediction [163]. In addition to developing *in silico* models for EI/EC using ML techniques and molecular fingerprints, Verma et al. combined quantitative structure–toxicity relationship (QSTR) models by ANN to produce 88% sensitivity and 82% specificity for EI; and 96% sensitivity and 91% specificity for eye corrosion (EC). Manually gathering data for training from X-Mol (<http://www.x-mol.com>) and ChemIDplus yielded 95% accuracy for EI and 96% for EC [164]. Using data on the transcriptional and molecular profiles of over a thousand drugs—35% of which have known cardiotoxicities—Mamoshina et al. employed ML to predict various drug-induced cardiotoxicities. The dataset was selected from a wide range of open-source knowledge and data sources (including DrugBank), with the best predictor achieving an average of 79% for safe vs. risky drug AUC and 66% for an unknown set of drugs. AUC (80%) indicated specific cardiotoxicity for specific drug classes and AUC (76%) indicated heart failures with potential for anti-neoplastic drugs across all investigated drug categories [165]. Weibel et al. achieved greater than 70% cytotoxicity prediction accuracy using a DL strategy developed from an internal dataset of more than 34,000 compounds with less than 5% cytotoxic chemicals. When applying this technique to new compounds, care must be taken to carefully consider the scope of the model. However, one of the advantages of this method is the use of cytotoxicity maps that provide the visual meaning of the substructures of different chemicals [166]. Hunta et al. developed three ML methods based on SVM, kNNs, and neural networks to predict DDIs in non-communicable diseases (NCDs). Using data from DrugBank, they combined the functions of transport proteins and enzymes and compared the results of different methods using five-fold cross-validation. This allowed them to determine which two NN layers performed best and predict NCDs based on pharmacokinetic mechanisms with an accuracy of 83% (F-measure 85.23% and AUC 90%) [167].

9. Role of AI in Synthetic Organic Chemistry

AI-assisted synthesis planning has a long history; it began in the field of computer-aided retrosynthetic prediction in the 1970s [168]. AI approaches have transformed synthetic organic chemistry in recent years, allowing for more streamlined and efficient processes for molecule design and synthesis. AI techniques, including ML and DL, have been shown to be effective in forecasting reaction results, improving reaction conditions, and suggesting new synthetic pathways. The search and creation of novel chemical compounds with desired qualities has been greatly sped up by these developments. Natural language processing techniques such as sequence-to-sequence models and transformer models have served as inspiration for this subject. The finding that the rank distribution of fragments in biological molecules is comparable to that of English words serves as the impetus for this area of inquiry [169]. By processing products using an encoder–decoder architecture, rule-free techniques usually take into account their products in a text-based form such as SMILES. This allows for the prediction of the matching synthetic precursors at a one-step reaction distance [170]. Several teams created techniques in 2017 that leverage AI to analyse big datasets of chemical reactions and suggest novel synthetic pathways [171]. These techniques split up synthesis planning into two distinct tasks: response prediction and search. In the search job, the program looks for a sequence of chemical interactions between a target molecule and a set of easily accessible starting components that will create a retrosynthetic “path”. The network attempts to ascertain if the reactions along the path will be viable and will deliver a fair yield of the target product based on the precedent of the reaction prediction task [6]. Segler et al. employed DL and then game AI in 2017 to quickly

and effectively identify feasible pathways; the field then saw a surge in interest [172,173]. A database of known reactions was automatically transformed into reaction templates by taking into account the reaction's core and its closest neighbour atoms, in contrast to manually encoding the rules. Through training to identify which template corresponded with a particular product, an ANN immediately acquired retrosynthetic methods from the data. Additionally, the network allows for an effective tree-search for feasible routes and prioritizes them among hundreds of reaction templates. An expert panel with blinded preferences compared the routes provided by the algorithm to those found in the literature and concluded that both were equally desirable. Segler et al. reduced the barrier to entry for other research organizations by demonstrating that data-driven techniques, as opposed to manual encoding, could be used successfully [173]. More recent advances in neural networks were leveraged by Lin and colleagues in a 2020 study to enhance this outcome and produce a 63.0% accuracy for their top-ranked prediction. They tried to forecast a reaction's feasibility in the reaction prediction stage. Training a neural network to match a chemical to a previously observed (named) response is one method for solving this challenge [174].

10. Real-World Case Studies

Gupta et al. have effectively used AI to identify a new drug that can be used to cure cancer. A sizable dataset of recognized anticancer drugs with related biological activity was used to train a DL algorithm. The developed model anticipated new substances with enhanced anticancer potential for the treatment of cancer. Another example is the utilization of ML algorithms to identify new inhibitors of beta-secretase (BACE 1), an enzyme implemented the progression of Alzheimer disease (AD). The brain of AD patient contains amyloid plaques which are produced by the enzyme BACE 1. ML makes it possible to anticipate compound binding affinities and screen through huge chemical libraries; this is useful in the identification of BACE 1 inhibitors. Molecular descriptors and structural characteristics of known BACE 1 inhibitors, as well as their experimentally confirmed binding affinities, were used to train ML models. These models can then forecast the binding affinity of new chemicals towards BACE 1 by learning from these training data [175,176].

Many companies were under pressure to find the best drug in the shortest period possible due to the SARS-CoV-2 viral epidemic. To achieve their objectives, some firms have resorted to using AI in conjunction with the existing data. These are some instances of businesses that, as a result of their efforts, have been able to successfully discover workable solutions to battle the COVID-19 virus [177].

Some companies have been able to successfully discover workable solutions to battle the COVID-19 virus. Rather than using 2D or 3D molecular structures, this technique predicts the strength of the interaction between a hit and its target protein using simpler chemical sequences. An essential protein in the SARS-CoV-2 virus, the causative agent of COVID-19, has a high likelihood of binding to and obstructing the actions of atazanavir, an FDA-approved antiviral medication used as an HIV treatment. In addition to remdesivir, three more antivirals are being studied in patients. Deargen's ability to uncover antiviral drugs via DL techniques is a significant development in the field of pharmaceutical research. Another biotechnology business, Benevolent AI located in London, uses AI, ML, and medical data to expedite research on health-related topics. Thus far, six medications have been identified such as ruxolitinib, baricitinib (Olumiant), tocilizumab (Actemra), paxlovid, lagevrio and remdesivir. Ruxolitinib is allegedly undergoing clinical trials for COVID-19. The company has been using a vast medical knowledge base, together with data gathered from the scientific literature by its AI system and ML, to find potential drugs that might obstruct the process of SARS-CoV-2 viral replication [177].

11. Challenges and Limitations of AI in Drug Discovery

Notwithstanding AI's potential advantages in drug development, there are several obstacles and restrictions that need to be taken into account. One of the primary issues

is the lack of adequate data [68]. AI-based methods frequently need a lot of data for training. The consistency, quality, and availability of adequate data can frequently affect how accurate and dependable the outcomes are [33]. With each molecule having between 10 and 50 conformations, or more than 500 billion distinct entities, analysing 30 billion molecules presents substantial logistical and practical hurdles. These intricacies are outside the purview of this investigation [178]. While computers have facilitated the process of selecting and optimizing lead compounds, completely automated drug development is still not possible due to the lack of artificially intelligent scientific robots or their digital counterparts. Automated AI drug creation is still an aspirational objective [179] (Figure 6).

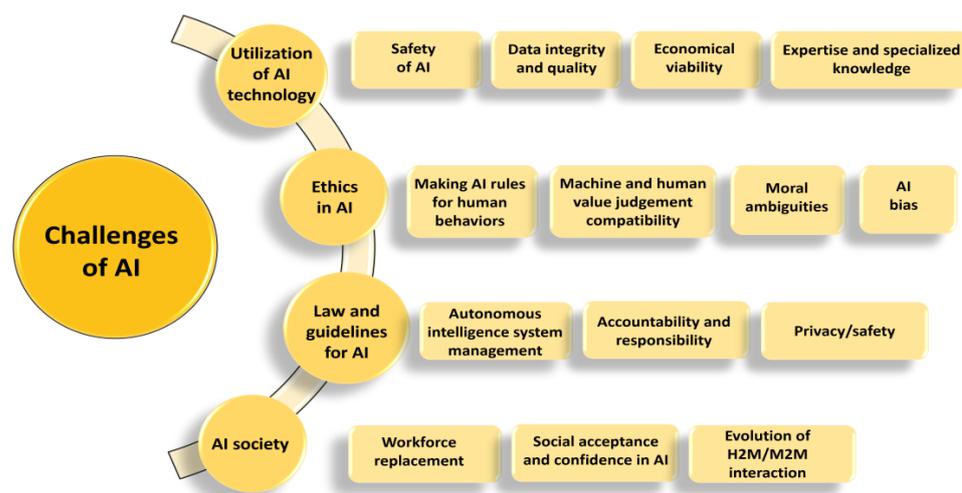


Figure 6. Challenges in the use of AI.

The similarity-based methods employed in these technologies typically do not account for heterogeneous information defined in a relationship network. Combining feature-based and similarity-based methods can help overcome this restriction. The fact that AI/ML-based models need a lot of training and that the requirements vary depending on the application is another drawback. Due to the learning qualities ingrained in the modelling data, DL models produce “black boxes” that are challenging to interpret; however, their use is growing in popularity. Nevertheless, their use is limited by the need for large amounts of high-quality data that are expensive to produce and frequently kept private. The use of inadequate quantities of high-quality data also affects the performance and dependability of DL models.

12. Future Perspective

AI has the potential to revolutionize the pharmaceutical industry by speeding up the process of finding and developing new drugs. The pharmaceutical business frequently uses both ML and DL algorithms. However, ML techniques have been the subject of many debates, especially in the fields of image analysis and omics data. The use of AI to quickly forecast novel leads and biological targets is expected to have a major impact on cutting down the length of time needed for drug discovery, as well as the likelihood of failures at subsequent stages. The identification of novel therapeutic indications for drugs might have a substantial impact on medication repurposing and reuse. In the last 10 years, DL has made considerable progress in a variety of AI research domains, such as speech and picture recognition and natural language processing. In data analysis, the application of edge detection, picture classification, and filters has been shown to yield superior outcomes. Additional investigation is necessary in the field of ML in conjunction with quantum ML to the biochemical methods utilized in the first phases of drug development. For example, quantum computers can shorten the training period of a limited Boltzmann machine. The sites of ligand binding have recently been predicted using DL methods. All the DL

methods available today, however, are built on popular convolutional/recurrent network architecture. To further enhance the precision of ligand binding studies LBS prediction, research that has effectively predicted and interpreted protein structure should serve as the foundation for next-generation DL structures. AI-enabled monitoring devices make it possible to provide remote patient care and ensure that prescriptions are followed. Continuous data collection will be facilitated by wearable devices and sensors. These data will be incorporated into AI algorithms to provide customized treatment programmes and enhance adherence. AI may also enhance clinical trial design, and patient recruitment and selection. AI algorithms may also be used to identify eligible patients, save testing expenses, and expedite therapy approval by utilizing genetic profiles, biomarkers, and electronic health information. AI has a lot of potential applications in medicine, such as multidimensional data analysis, prediction of drug bioactivity and interactions, precise decision-making, disease classification, dosage ratio optimization, acceleration of drug development, and complex problem-solving related to the creation of efficient drug delivery systems. Although there are many exciting possibilities presented by this future framework, it is critical to recognize that to fully fulfil the potential of AI in pharmaceutical product development, concerns about data quality, regulatory frameworks, and ethical standards must be resolved. Thus, with continuing research and development, as well as collaborations between businesses, academic institutions, and regulatory bodies, AI-based technologies have the potential to revolutionize the pharmaceutical sector and enhance patient outcomes in the years to come.

13. Ethical Considerations

When using these technologies in medication development, ethical considerations must be made. Since confidential medical data are frequently used to train AI models, patient privacy is a serious problem. Data security and safety are important considerations that should not be disregarded. It is important to guarantee that patients' data are gathered and handled in a manner that upholds their rights and safeguards their privacy. The possibility of bias in AI algorithms is another significant worry as it may result in unfair treatment of groups of people and uneven access to healthcare. The ideas of equality and fairness may be jeopardized. The other issue is the possibility of bias in AI algorithms, which might result in unfair treatment of groups of people and uneven access to healthcare. This might lead to a compromise in the ideas of justice and equality. Regulatory bodies have been entrusted with developing stringent standard protocols, guidelines, and assessment methods to successfully incorporate AI into drug creation. These regulations ought to address a variety of topics, including ethical dilemmas as well as concerns about patient safety and animal welfare. Every effort must be made to minimize, enhance, and replace animal models to uphold ethical principles. An important aspect of drug research is animal testing. In addressing the ethical and legal ramifications of AI in drug discovery, the FDA's release of a discussion paper titled "Use of AI and ML in Drug and Biologics Development" is a significant step. An overview of the use of AI to drug discovery, non-clinical research, and clinical trials is given in this article; practices for applying ML and AI are also provided. In general, a great deal of investigation and techniques are required to address the ethical use of AI in the pharmaceutical sector. AI may be used properly and ethically by the pharmaceutical business to address these issues.

14. Conclusions

Recent developments in ML and DL techniques have surely helped in the area of drug discovery and development. Numerous AI tools and techniques have been created to help with different stages of the drug development process, including lead optimization, clinical trials, drug repurposing, protein–ligand binding, 3D structure prediction, disease target gene prediction, drug screening, toxicity and bioactivity prediction, and lead optimization. The goal of AI development and its advanced tools is always to make life easier for pharmaceutical businesses. The use of the most recent AI-based technologies will make

automation more relevant because it will reduce time to market while also improving product quality, boosting overall production process safety, and ensuring higher resource utilization while maintaining cost-effectiveness. There are still particular challenges in putting these technologies to use in reality. Although there are not any AI-based drugs on the market now, it is anticipated that AI will eventually become an indispensable and helpful tool for the pharmaceutical sector.

Author Contributions: Conceptualization, A.G. and A.P.; data curation, R.D.; writing—original draft preparation, R.D.; writing—review and editing, A.G. and A.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work received no external funding.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

AI	Artificial intelligence
ALS	Amyotrophic lateral sclerosis
ANN	Artificial neural network
CNN	Convolutional Neural Network
DL	Deep learning
DNNs	Deep neural networks
FDA	Food and Drug Administration
GBDT	Gradient Boosted Decision Tree
GANs	Generative adversarial networks
GBMs	gradient boosting machines
GNNs	Graph Neural Networks
LMs	Large Modules
LBVS	Ligand-based virtual screening
MAE	Mean absolute error
ML	Machine learning
MMPs	Matched molecular pairs
MD	Molecular dynamics
QSAR	Quantitative structure–activity relationship
QM	Quantum mechanics
PK	Pharmacokinetic
RF	Random Forest
RNNs	Recurrent Neural Networks
RL	Reinforcement learning
SF	Scoring factor
STM	Short-Term Memory
SVM	Support Vector Machine
VS	Virtual Scoring

References

1. Zhu, T.; Cao, S.; Su, P.-C.; Patel, R.; Shah, D.; Chokshi, H.B.; Szukala, R.; Johnson, M.E.; Hevener, K.E. Hit identification and optimization in virtual screening: Practical recommendations based on a critical literature analysis: Miniperspective. *J. Med. Chem.* **2013**, *56*, 6560–6572. [[CrossRef](#)]
2. Chan, H.S.; Shan, H.; Dahoun, T.; Vogel, H.; Yuan, S. Advancing drug discovery via artificial intelligence. *Trends Pharmacol. Sci.* **2019**, *40*, 592–604. [[CrossRef](#)]
3. Goel, A.; Davies, J. Artificial Intelligence. In *Cambridge Handbook of Intelligence*; Cambridge University Press: Cambridge, UK, 2011; pp. 468–482. [[CrossRef](#)]
4. Nitta, N.; Sugimura, T.; Isozaki, A.; Mikami, H.; Hiraki, K.; Sakuma, S.; Iino, T.; Arai, F.; Endo, T.; Fujiwaki, Y. Intelligent image-activated cell sorting. *Cell* **2018**, *175*, 266–276.e213. [[CrossRef](#)] [[PubMed](#)]
5. Von Lilienfeld, O.A. Quantum machine learning in chemical compound space. *Angew. Chem. Int. Ed.* **2018**, *57*, 4164–4169. [[CrossRef](#)] [[PubMed](#)]
6. Coley, C.W.; Barzilay, R.; Jaakkola, T.S.; Green, W.H.; Jensen, K.F. Prediction of organic reaction outcomes using machine learning. *ACS Cent. Sci.* **2017**, *3*, 434–443. [[CrossRef](#)]

7. Popova, M.; Isayev, O.; Tropsha, A. Deep reinforcement learning for de novo drug design. *Sci. Adv.* **2018**, *4*, eaap7885. [[CrossRef](#)]
8. Lamberti, M.J.; Wilkinson, M.; Donzanti, B.A.; Wohlhieter, G.E.; Parikh, S.; Wilkins, R.G.; Getz, K. A study on the application and use of artificial intelligence to support drug development. *Clin. Ther.* **2019**, *41*, 1414–1426. [[CrossRef](#)]
9. Agatonovic-Kustrin, S.; Beresford, R. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *J. Pharm. Biomed. Anal.* **2000**, *22*, 717–727. [[CrossRef](#)]
10. Jeon, J.; Nim, S.; Teyra, J.; Datti, A.; Wrana, J.L.; Sidhu, S.S.; Moffat, J.; Kim, P.M. A systematic approach to identify novel cancer drug targets using machine learning, inhibitor design and high-throughput screening. *Genome Med.* **2014**, *6*, 57. [[CrossRef](#)]
11. Ferrero, E.; Dunham, I.; Sanseau, P. In silico prediction of novel therapeutic targets using gene–disease association data. *J. Transl. Med.* **2017**, *15*, 182. [[CrossRef](#)]
12. Riniker, S.; Wang, Y.; Jenkins, J.L.; Landrum, G.A. Using information from historical high-throughput screens to predict active compounds. *J. Chem. Inf. Model.* **2014**, *54*, 1880–1891. [[CrossRef](#)] [[PubMed](#)]
13. Godinez, W.J.; Hossain, I.; Lazic, S.E.; Davies, J.W.; Zhang, X. A multi-scale convolutional neural network for phenotyping high-content cellular images. *Bioinformatics* **2017**, *33*, 2010–2019. [[CrossRef](#)] [[PubMed](#)]
14. Olsen, T.G.; Jackson, B.H.; Feeser, T.A.; Kent, M.N.; Moad, J.C.; Krishnamurthy, S.; Lunsford, D.D.; Soans, R.E. Diagnostic performance of deep learning algorithms applied to three common diagnoses in dermatopathology. *J. Pathol. Inform.* **2018**, *9*, 32. [[CrossRef](#)] [[PubMed](#)]
15. Ramesh, A.; Kambhampati, C.; Monson, J.R.; Drew, P. Artificial intelligence in medicine. *Ann. R. Coll. Surg. Engl.* **2004**, *86*, 334. [[CrossRef](#)] [[PubMed](#)]
16. Miles, J.; Walker, A.J. The potential application of artificial intelligence in transport. In Proceedings of the IEE Proceedings-Intelligent Transport Systems, Toronto, ON, Canada, 17–20 September 2006; pp. 183–198.
17. Hearst, M.A.; Dumais, S.T.; Osuna, E.; Platt, J.; Scholkopf, B. Support vector machines. *IEEE Intell. Syst. Their Appl.* **1998**, *13*, 18–28. [[CrossRef](#)]
18. Ho, T.K. Random decision forests. In Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, Canada, 14–16 August 1995; pp. 278–282.
19. Wessel, M.D.; Jurs, P.C.; Tolan, J.W.; Muskal, S.M. Prediction of human intestinal absorption of drug compounds from molecular structure. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 726–735. [[CrossRef](#)] [[PubMed](#)]
20. Abe, H.; Kumazawa, S.; Taji, T.; Sasaki, S.I. Applications of computerized pattern recognition: A survey of correlations between pharmacological activities and mass spectra. *Biomed. Mass Spectrom.* **1976**, *3*, 151–154. [[CrossRef](#)] [[PubMed](#)]
21. Stuper, A.J.; Jurs, P.C. Classification of psychotropic drugs as sedatives or tranquilizers using pattern recognition techniques. *J. Am. Chem. Soc.* **1975**, *97*, 182–187. [[CrossRef](#)]
22. Gulati, K.; Aw, M.S.; Losic, D. Nanoengineered drug-releasing Ti wires as an alternative for local delivery of chemotherapeutics in the brain. *Int. J. Nanomed.* **2012**, *7*, 2069–2076.
23. Rautio, J.; Kumpulainen, H.; Heimbach, T.; Oliyai, R.; Oh, D.; Järvinen, T.; Savolainen, J. Prodrugs: Design and clinical applications. *Nat. Rev. Drug Discov.* **2008**, *7*, 255–270. [[CrossRef](#)]
24. Baldi, A. Computational approaches for drug design and discovery: An overview. *Syst. Rev. Pharm.* **2010**, *1*, 99. [[CrossRef](#)]
25. Lavecchia, A.; Cerchia, C. In silico methods to address polypharmacology: Current status, applications and future perspectives. *Drug Discov. Today* **2016**, *21*, 288–298. [[CrossRef](#)]
26. De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. Role of molecular dynamics and related methods in drug discovery. *J. Med. Chem.* **2016**, *59*, 4035–4061. [[CrossRef](#)]
27. Wess, G.; Urmann, M.; Sickenberger, B. Medicinal chemistry: Challenges and opportunities. *Angew. Chem. Int. Ed.* **2001**, *40*, 3341–3350. [[CrossRef](#)]
28. Pu, L.; Naderi, M.; Liu, T.; Wu, H.-C.; Mukhopadhyay, S.; Brylinski, M. eToxPred: A machine learning-based approach to estimate the toxicity of drug candidates. *BMC Pharmacol. Toxicol.* **2019**, *20*, 2. [[CrossRef](#)] [[PubMed](#)]
29. Lang, P.T.; Kuntz, I.D.; Maggiora, G.M.; Bajorath, J. Evaluating the high-throughput screening computations. *J. Biomol. Screen.* **2005**, *10*, 649–652. [[CrossRef](#)] [[PubMed](#)]
30. Bannigan, P.; Aldeghi, M.; Bao, Z.; Häse, F.; Aspuru-Guzik, A.; Allen, C. Machine learning directed drug formulation development. *Adv. Drug Deliv. Rev.* **2021**, *175*, 113806. [[CrossRef](#)] [[PubMed](#)]
31. Al Qaraghuli, M.M.; Alzahrani, A.R.; Niwasabutra, K.; Obeid, M.A.; Ferro, V.A. Where traditional drug discovery meets modern technology in the quest for new drugs. *Ann. Pharmacol. Pharm.* **2017**, *2*, 1061.
32. Chen, R.; Liu, X.; Jin, S.; Lin, J.; Liu, J. Machine learning for drug-target interaction prediction. *Molecules* **2018**, *23*, 2208. [[CrossRef](#)]
33. Gómez-Bombarelli, R.; Wei, J.N.; Duvenaud, D.; Hernández-Lobato, J.M.; Sánchez-Lengeling, B.; Sheberla, D.; Aguilera-Iparraguirre, J.; Hirzel, T.D.; Adams, R.P.; Aspuru-Guzik, A. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Cent. Sci.* **2018**, *4*, 268–276. [[CrossRef](#)]
34. Mak, K.-K.; Pichika, M.R. Artificial intelligence in drug development: Present status and future prospects. *Drug Discov. Today* **2019**, *24*, 773–780. [[CrossRef](#)] [[PubMed](#)]
35. Duch, W.; Swaminathan, K.; Meller, J. Artificial intelligence approaches for rational drug design and discovery. *Curr. Pharm. Des.* **2007**, *13*, 1497–1508. [[CrossRef](#)] [[PubMed](#)]
36. Yang, Y.; Siau, K.L. A qualitative research on marketing and sales in the artificial intelligence age. In Proceedings of the Thirteenth Annual Midwest Association for Information Systems Conference (MWAIS 2018), St. Louis, MO, USA, 17–18 May 2018.

37. Wirtz, B.W.; Weyerer, J.C.; Geyer, C. Artificial intelligence and the public sector—Applications and challenges. *Int. J. Public Adm.* **2019**, *42*, 596–615. [[CrossRef](#)]
38. Zhong, F.; Xing, J.; Li, X.; Liu, X.; Fu, Z.; Xiong, Z.; Lu, D.; Wu, X.; Zhao, J.; Tan, X. Artificial intelligence in drug design. *Sci. China Life Sci.* **2018**, *61*, 1191–1204. [[CrossRef](#)] [[PubMed](#)]
39. Sidey-Gibbons, J.A.; Sidey-Gibbons, C.J. Machine learning in medicine: A practical introduction. *BMC Med. Res. Methodol.* **2019**, *19*, 64. [[CrossRef](#)] [[PubMed](#)]
40. Carvalho, T.F.M.; Silva, J.C.F.; Calil, I.P.; Fontes, E.P.B.; Cerqueira, F.R. Rama: A machine learning approach for ribosomal protein prediction in plants. *Sci. Rep.* **2017**, *7*, 16273. [[CrossRef](#)] [[PubMed](#)]
41. Silva, J.C.F.; Teixeira, R.M.; Silva, F.F.; Brommonschenkel, S.H.; Fontes, E.P. Machine learning approaches and their current application in plant molecular biology: A systematic review. *Plant Sci.* **2019**, *284*, 37–47. [[CrossRef](#)] [[PubMed](#)]
42. Manne, R. Machine learning techniques in drug discovery and development. *Int. J. Appl. Res.* **2021**, *7*, 21–28. [[CrossRef](#)]
43. Polydoros, A.S.; Nalpantidis, L. Survey of model-based reinforcement learning: Applications on robotics. *J. Intell. Robot. Syst.* **2017**, *86*, 153–173. [[CrossRef](#)]
44. Quah, K.H.; Quek, C.; Leedham, G. Reinforcement learning combined with a fuzzy adaptive learning control network (FALCON-R) for pattern classification. *Pattern Recognit.* **2005**, *38*, 513–526. [[CrossRef](#)]
45. Abdulhai, B.; Pringle, R.; Karakoulas, G.J. Reinforcement learning for true adaptive traffic signal control. *J. Transp. Eng.* **2003**, *129*, 278–285. [[CrossRef](#)]
46. Kubat, M.; Kubat, J. *An Introduction to Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2017; Volume 2.
47. Birhane, A.; Kasirzadeh, A.; Leslie, D.; Wachter, S. Science in the age of large language models. *Nat. Rev. Phys.* **2023**, *5*, 277–280. [[CrossRef](#)]
48. Zhang, L.; Tan, J.; Han, D.; Zhu, H. From machine learning to deep learning: Progress in machine intelligence for rational drug discovery. *Drug Discov. Today* **2017**, *22*, 1680–1685. [[CrossRef](#)] [[PubMed](#)]
49. Hinkson, I.V.; Madej, B.; Stahlberg, E.A. Accelerating therapeutics for opportunities in medicine: A paradigm shift in drug discovery. *Front. Pharmacol.* **2020**, *11*, 770. [[CrossRef](#)]
50. Zhou, Y.; Zhang, Y.; Lian, X.; Li, F.; Wang, C.; Zhu, F.; Qiu, Y.; Chen, Y. Therapeutic target database update 2022: Facilitating drug discovery with enriched comparative data of targeted agents. *Nucleic Acids Res.* **2022**, *50*, D1398–D1407. [[CrossRef](#)]
51. Kana, O.; Brylinski, M. Elucidating the druggability of the human proteome with e findsite. *J. Comput.-Aided Mol. Des.* **2019**, *33*, 509–519. [[CrossRef](#)] [[PubMed](#)]
52. Finan, C.; Gaulton, A.; Kruger, F.A.; Lumbers, R.T.; Shah, T.; Engmann, J.; Galver, L.; Kelley, R.; Karlsson, A.; Santos, R. The druggable genome and support for target identification and validation in drug development. *Sci. Transl. Med.* **2017**, *9*, eaag1166. [[CrossRef](#)] [[PubMed](#)]
53. Schenone, M.; Dančik, V.; Wagner, B.K.; Clemons, P.A. Target identification and mechanism of action in chemical biology and drug discovery. *Nat. Chem. Biol.* **2013**, *9*, 232–240. [[CrossRef](#)]
54. Pun, F.W.; Ozerov, I.V.; Zhavoronkov, A. AI-powered therapeutic target discovery. *Trends Pharmacol. Sci.* **2023**, *44*, 561–572. [[CrossRef](#)]
55. Pun, F.W.; Liu, B.H.M.; Long, X.; Leung, H.W.; Leung, G.H.D.; Mewborne, Q.T.; Gao, J.; Shneyderman, A.; Ozerov, I.V.; Wang, J. Identification of therapeutic targets for amyotrophic lateral sclerosis using PandaOmics—An AI-Enabled Biological Target Discovery Platform. *Front. Aging Neurosci.* **2022**, *14*, 638. [[CrossRef](#)]
56. Zhang, S.; Cooper-Knock, J.; Weimer, A.K.; Shi, M.; Moll, T.; Marshall, J.N.; Harvey, C.; Nezhad, H.G.; Franklin, J.; dos Santos Souza, C. Genome-wide identification of the genetic basis of amyotrophic lateral sclerosis. *Neuron* **2022**, *110*, 992–1008.e1011. [[CrossRef](#)]
57. Zeng, X.; Zhu, S.; Lu, W.; Liu, Z.; Huang, J.; Zhou, Y.; Fang, J.; Huang, Y.; Guo, H.; Li, L. Target identification among known drugs by deep learning from heterogeneous networks. *Chem. Sci.* **2020**, *11*, 1775–1797. [[CrossRef](#)] [[PubMed](#)]
58. McCulloch, W.S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **1943**, *5*, 115–133. [[CrossRef](#)]
59. Zhavoronkov, A.; Mamoshina, P.; Vanhaelen, Q.; Scheibye-Knudsen, M.; Moskalev, A.; Aliper, A. Artificial intelligence for aging and longevity research: Recent advances and perspectives. *Ageing Res. Rev.* **2019**, *49*, 49–66. [[CrossRef](#)] [[PubMed](#)]
60. Giorgi, J.M.; Bader, G.D. Towards reliable named entity recognition in the biomedical domain. *Bioinformatics* **2020**, *36*, 280–286. [[CrossRef](#)]
61. Wang, X.; Zhang, Y.; Ren, X.; Zhang, Y.; Zitnik, M.; Shang, J.; Langlotz, C.; Han, J. Cross-type biomedical named entity recognition with deep multi-task learning. *Bioinformatics* **2019**, *35*, 1745–1752. [[CrossRef](#)]
62. Fabris, F.; Palmer, D.; Salama, K.M.; de Magalhaes, J.P.; Freitas, A.A. Using deep learning to associate human genes with age-related diseases. *Bioinformatics* **2020**, *36*, 2202–2208. [[CrossRef](#)]
63. Shayakhmetov, R.; Kuznetsov, M.; Zhebrak, A.; Kadurin, A.; Nikolenko, S.; Aliper, A.; Polykovskiy, D. Molecular generation for desired transcriptome changes with adversarial autoencoders. *Front. Pharmacol.* **2020**, *11*, 269. [[CrossRef](#)] [[PubMed](#)]
64. Viñas, R.; Andrés-Terré, H.; Liò, P.; Bryson, K. Adversarial generation of gene expression data. *Bioinformatics* **2022**, *38*, 730–737. [[CrossRef](#)]
65. Li, G.; Luo, J.; Xiao, Q.; Liang, C.; Ding, P.; Cao, B. Predicting microrna-disease associations using network topological similarity based on deepwalk. *IEEE Access* **2017**, *5*, 24032–24039. [[CrossRef](#)]
66. Camacho, D.M.; Collins, K.M.; Powers, R.K.; Costello, J.C.; Collins, J.J. Next-generation machine learning for biological networks. *Cell* **2018**, *173*, 1581–1592. [[CrossRef](#)]

67. Li, Z.; An, L.; Li, H.; Wang, S.; Zhou, Y.; Yuan, F.; Li, L. Identifying novel genes and chemicals related to nasopharyngeal cancer in a heterogeneous network. *Sci. Rep.* **2016**, *6*, 25515. [[CrossRef](#)]
68. Vamathevan, J.; Clark, D.; Czodrowski, P.; Dunham, I.; Ferran, E.; Lee, G.; Li, B.; Madabhushi, A.; Shah, P.; Spitzer, M. Applications of machine learning in drug discovery and development. *Nat. Rev. Drug Discov.* **2019**, *18*, 463–477. [[CrossRef](#)] [[PubMed](#)]
69. Bennett, W.D.; He, S.; Bilodeau, C.L.; Jones, D.; Sun, D.; Kim, H.; Allen, J.E.; Lightstone, F.C.; Ingólfsson, H.I. Predicting small molecule transfer free energies by combining molecular dynamics simulations and deep learning. *J. Chem. Inf. Model.* **2020**, *60*, 5375–5381. [[CrossRef](#)] [[PubMed](#)]
70. Karplus, M.; McCammon, J.A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **2002**, *9*, 646–652. [[CrossRef](#)] [[PubMed](#)]
71. Hollingsworth, S.A.; Dror, R.O. Molecular dynamics simulation for all. *Neuron* **2018**, *99*, 1129–1143. [[CrossRef](#)] [[PubMed](#)]
72. Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*, 19–25. [[CrossRef](#)]
73. Case, D.A.; Cheatham III, T.E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K.M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R.J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688. [[CrossRef](#)] [[PubMed](#)]
74. Phillips, J.C.; Hardy, D.J.; Maia, J.D.; Stone, J.E.; Ribeiro, J.V.; Bernardi, R.C.; Buch, R.; Fiorin, G.; Hémin, J.; Jiang, W. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **2020**, *153*, 044130. [[CrossRef](#)] [[PubMed](#)]
75. Scott, W.R.; Hünenberger, P.H.; Tironi, I.G.; Mark, A.E.; Billeter, S.R.; Fennen, J.; Torda, A.E.; Huber, T.; Krüger, P.; Van Gunsteren, W.F. The GROMOS biomolecular simulation program package. *J. Phys. Chem. A* **1999**, *103*, 3596–3607. [[CrossRef](#)]
76. Wang, Y.; Ribeiro, J.M.L.; Tiwary, P. Machine learning approaches for analyzing and enhancing molecular dynamics simulations. *Curr. Opin. Struct. Biol.* **2020**, *61*, 139–145. [[CrossRef](#)]
77. Noé, F.; Tkatchenko, A.; Müller, K.-R.; Clementi, C. Machine learning for molecular simulation. *Annu. Rev. Phys. Chem.* **2020**, *71*, 361–390. [[CrossRef](#)]
78. Plante, A.; Shore, D.M.; Morra, G.; Khelashvili, G.; Weinstein, H. A machine learning approach for the discovery of ligand-specific functional mechanisms of GPCRs. *Molecules* **2019**, *24*, 2097. [[CrossRef](#)]
79. Marchetti, F.; Moroni, E.; Pandini, A.; Colombo, G. Machine learning prediction of allosteric drug activity from molecular dynamics. *J. Phys. Chem. Lett.* **2021**, *12*, 3724–3732. [[CrossRef](#)]
80. Lyngdoh, G.A.; Li, H.; Zaki, M.; Krishnan, N.A.; Das, S. Elucidating the constitutive relationship of calcium–silicate–hydrate gel using high throughput reactive molecular simulations and machine learning. *Sci. Rep.* **2020**, *10*, 21336. [[CrossRef](#)]
81. Ragoza, M.; Hochuli, J.; Idrobo, E.; Sunseri, J.; Koes, D.R. Protein–ligand scoring with convolutional neural networks. *J. Chem. Inf. Model.* **2017**, *57*, 942–957. [[CrossRef](#)]
82. Nguyen, D.D.; Cang, Z.; Wu, K.; Wang, M.; Cao, Y.; Wei, G.-W. Mathematical deep learning for pose and binding affinity prediction and ranking in D3R Grand Challenges. *J. Comput.-Aided Mol. Des.* **2019**, *33*, 71–82. [[CrossRef](#)] [[PubMed](#)]
83. Ashtawy, H.M.; Mahapatra, N.R. Machine-learning scoring functions for identifying native poses of ligands docked to known and novel proteins. *BMC Bioinform.* **2015**, *16*, S3. [[CrossRef](#)] [[PubMed](#)]
84. Wójcikowski, M.; Ballester, P.J.; Siedlecki, P. Performance of machine-learning scoring functions in structure-based virtual screening. *Sci. Rep.* **2017**, *7*, 46710. [[CrossRef](#)] [[PubMed](#)]
85. Yasuo, N.; Sekijima, M. Improved method of structure-based virtual screening via interaction-energy-based learning. *J. Chem. Inf. Model.* **2019**, *59*, 1050–1061. [[CrossRef](#)] [[PubMed](#)]
86. Ballester, P.J.; Mitchell, J.B. A machine learning approach to predicting protein–ligand binding affinity with applications to molecular docking. *Bioinformatics* **2010**, *26*, 1169–1175. [[CrossRef](#)]
87. Li, L.; Khanna, M.; Jo, I.; Wang, F.; Ashpole, N.M.; Hudmon, A.; Meroueh, S.O. Target-specific support vector machine scoring in structure-based virtual screening: Computational validation, in vitro testing in kinases, and effects on lung cancer cell proliferation. *J. Chem. Inf. Model.* **2011**, *51*, 755–759. [[CrossRef](#)] [[PubMed](#)]
88. Sato, T.; Honma, T.; Yokoyama, S. Combining machine learning and pharmacophore-based interaction fingerprint for in silico screening. *J. Chem. Inf. Model.* **2010**, *50*, 170–185. [[CrossRef](#)] [[PubMed](#)]
89. Nogueira, M.S.; Koch, O. The development of target-specific machine learning models as scoring functions for docking-based target prediction. *J. Chem. Inf. Model.* **2019**, *59*, 1238–1252. [[CrossRef](#)] [[PubMed](#)]
90. Li, L.; Wang, B.; Meroueh, S.O. Support vector regression scoring of receptor–ligand complexes for rank-ordering and virtual screening of chemical libraries. *J. Chem. Inf. Model.* **2011**, *51*, 2132–2138. [[CrossRef](#)] [[PubMed](#)]
91. Jiménez, J.; Škalič, M.; Martínez-Rosell, G.; Fabritiis, G.D. KDEEP: Protein-Ligand Absolute Binding Affinity Prediction via 3D-Convolutional Neural Networks. *J. Chem. Inf. Model.* **2018**, *58*, 287–296. [[CrossRef](#)] [[PubMed](#)]
92. Thafar, M.A.; Alshahrani, M.; Albaradei, S.; Gojobori, T.; Essack, M.; Gao, X. Affinity2Vec: Drug-target binding affinity prediction through representation learning, graph mining, and machine learning. *Sci. Rep.* **2022**, *12*, 4751. [[CrossRef](#)] [[PubMed](#)]
93. Duy Nguyen, D.; Xiao, T.; Wang, M.; Wei, G.-W. Rigidity strengthening is a vital mechanism for protein-ligand binding. *arXiv* **2017**, arXiv:1704.05883.
94. Weston, J.; Perez-Cruz, F.; Bousquet, O.; Chapelle, O.; Elisseeff, A.; Schölkopf, B. Feature selection and transduction for prediction of molecular bioactivity for drug design. *Bioinformatics* **2003**, *19*, 764–771. [[CrossRef](#)]
95. Li, G.-B.; Yang, L.-L.; Wang, W.-J.; Li, L.-L.; Yang, S.-Y. ID-Score: A new empirical scoring function based on a comprehensive set of descriptors related to protein–ligand interactions. *J. Chem. Inf. Model.* **2013**, *53*, 592–600. [[CrossRef](#)]

96. Porath, A.J.; Beirness, D.J. Predicting categories of drugs used by suspected drug-impaired drivers using the Drug Evaluation and Classification Program tests. *Traffic Inj. Prev.* **2019**, *20*, 255–263. [[CrossRef](#)]
97. Nguyen, D.D.; Wei, G.W. DG-GL: Differential geometry-based geometric learning of molecular datasets. *Int. J. Numer. Methods Biomed. Eng.* **2019**, *35*, e3179. [[CrossRef](#)]
98. Durrant, J.D.; McCammon, J.A. NNScore: A neural-network-based scoring function for the characterization of protein–ligand complexes. *J. Chem. Inf. Model.* **2010**, *50*, 1865–1871. [[CrossRef](#)]
99. Ouyang, X.; Handoko, S.D.; Kwok, C.K. Cscore: A simple yet effective scoring function for protein–ligand binding affinity prediction using modified cmac learning architecture. *J. Bioinform. Comput. Biol.* **2011**, *9*, 1–14. [[CrossRef](#)]
100. Cang, Z.; Wei, G.-W. TopologyNet: Topology based deep convolutional and multi-task neural networks for biomolecular property predictions. *PLoS Comput. Biol.* **2017**, *13*, e1005690. [[CrossRef](#)]
101. Pereira, J.C.; Caffarena, E.R.; Dos Santos, C.N. Boosting docking-based virtual screening with deep learning. *J. Chem. Inf. Model.* **2016**, *56*, 2495–2506. [[CrossRef](#)]
102. Paul, S.M.; Mytelka, D.S.; Dunwiddie, C.T.; Persinger, C.C.; Munos, B.H.; Lindborg, S.R.; Schacht, A.L. How to improve R&D productivity: The pharmaceutical industry’s grand challenge. *Nat. Rev. Drug Discov.* **2010**, *9*, 203–214. [[PubMed](#)]
103. Myers, S.; Baker, A. Drug discovery—An operating model for a new era. *Nat. Biotechnol.* **2001**, *19*, 727–730. [[CrossRef](#)]
104. Wang, L.; Ding, J.; Pan, L.; Cao, D.; Jiang, H.; Ding, X. Artificial intelligence facilitates drug design in the big data era. *Chemom. Intell. Lab. Syst.* **2019**, *194*, 103850. [[CrossRef](#)]
105. Tripathy, R.K.; Mahanta, S.; Paul, S. Artificial intelligence-based classification of breast cancer using cellular images. *Rsc Adv.* **2014**, *4*, 9349–9355. [[CrossRef](#)]
106. Samui, P.; Kothari, D. Utilization of a least square support vector machine (LSSVM) for slope stability analysis. *Sci. Iran.* **2011**, *18*, 53–58. [[CrossRef](#)]
107. Fernández-Ruiz, I. Artificial intelligence to improve the diagnosis of cardiovascular diseases. *Nat. Rev. Cardiol.* **2019**, *16*, 133. [[CrossRef](#)] [[PubMed](#)]
108. Hessler, G.; Baringhaus, K.-H. Artificial intelligence in drug design. *Molecules* **2018**, *23*, 2520. [[CrossRef](#)] [[PubMed](#)]
109. Yang, X.; Wang, Y.; Byrne, R.; Schneider, G.; Yang, S. Concepts of artificial intelligence for computer-assisted drug discovery. *Chem. Rev.* **2019**, *119*, 10520–10594. [[CrossRef](#)]
110. Panapitiya, G.; Girard, M.; Hollas, A.; Sepulveda, J.; Murugesan, V.; Wang, W.; Saldanha, E. Evaluation of deep learning architectures for aqueous solubility prediction. *ACS Omega* **2022**, *7*, 15695–15710. [[CrossRef](#)]
111. Sheridan, R.P.; Wang, W.M.; Liaw, A.; Ma, J.; Gifford, E.M. Extreme gradient boosting as a method for quantitative structure–activity relationships. *J. Chem. Inf. Model.* **2016**, *56*, 2353–2360. [[CrossRef](#)] [[PubMed](#)]
112. Wallach, I.; Dzamba, M.; Heifets, A. AtomNet: A deep convolutional neural network for bioactivity prediction in structure-based drug discovery. *arXiv* **2015**, arXiv:1510.02855.
113. Leach, A.G.; Jones, H.D.; Cosgrove, D.A.; Kenny, P.W.; Ruston, L.; MacFaul, P.; Wood, J.M.; Colclough, N.; Law, B. Matched molecular pairs as a guide in the optimization of pharmaceutical properties; a study of aqueous solubility, plasma protein binding and oral exposure. *J. Med. Chem.* **2006**, *49*, 6672–6682. [[CrossRef](#)]
114. Paixão, P.; Gouveia, L.F.; Morais, J.A. Prediction of the in vitro intrinsic clearance determined in suspensions of human hepatocytes by using artificial neural networks. *Eur. J. Pharm. Sci.* **2010**, *39*, 310–321. [[CrossRef](#)]
115. Keefer, C.E.; Chang, G.; Kauffman, G.W. Extraction of tacit knowledge from large ADME data sets via pairwise analysis. *Bioorganic Med. Chem.* **2011**, *19*, 3739–3749. [[CrossRef](#)]
116. Koutsoukas, A.; Monaghan, K.J.; Li, X.; Huan, J. Deep-learning: Investigating deep neural networks hyper-parameters and comparison of performance to shallow methods for modeling bioactivity data. *J. Cheminform.* **2017**, *9*, 42. [[CrossRef](#)]
117. Stokes, J.M.; Yang, K.; Swanson, K.; Jin, W.; Cubillos-Ruiz, A.; Donghia, N.M.; MacNair, C.R.; French, S.; Carfrae, L.A.; Bloom-Ackermann, Z. A deep learning approach to antibiotic discovery. *Cell* **2020**, *180*, 688–702.e613. [[CrossRef](#)]
118. Jain, S.; Siramshetty, V.B.; Alves, V.M.; Muratov, E.N.; Kleinstreuer, N.; Tropsha, A.; Nicklaus, M.C.; Simeonov, A.; Zakharov, A.V. Large-scale modeling of multispecies acute toxicity end points using consensus of multitask deep learning methods. *J. Chem. Inf. Model.* **2021**, *61*, 653–663. [[CrossRef](#)]
119. Lysenko, A.; Sharma, A.; Boroevich, K.A.; Tsunoda, T. An integrative machine learning approach for prediction of toxicity-related drug safety. *Life Sci. Alliance* **2018**, *1*. [[CrossRef](#)] [[PubMed](#)]
120. Basile, A.O.; Yahi, A.; Tatonetti, N.P. Artificial intelligence for drug toxicity and safety. *Trends Pharmacol. Sci.* **2019**, *40*, 624–635. [[CrossRef](#)] [[PubMed](#)]
121. Segall, M. Advances in multiparameter optimization methods for de novo drug design. *Expert Opin. Drug Discov.* **2014**, *9*, 803–817. [[CrossRef](#)] [[PubMed](#)]
122. Meyers, J.; Fabian, B.; Brown, N. De novo molecular design and generative models. *Drug Discov. Today* **2021**, *26*, 2707–2715. [[CrossRef](#)] [[PubMed](#)]
123. Kingma, D.P.; Welling, M. Auto-encoding variational bayes. *arXiv* **2013**, arXiv:1312.6114.
124. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **2014**, *27*.
125. Bengio, Y. Learning deep architectures for AI. *Found. Trends®Mach. Learn.* **2009**, *2*, 1–127. [[CrossRef](#)]
126. Géron, A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*; O’Reilly Media, Inc.: Sebastopol, CA, USA, 2022.

127. Heaton, J.; Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; The MIT Press: Cambridge, MA, USA, 2016; 800p; ISBN 0262035618. [[CrossRef](#)]
128. Kadurin, A.; Aliper, A.; Kazennov, A.; Mamoshina, P.; Vanhaelen, Q.; Khrabrov, K.; Zhavoronkov, A. The cornucopia of meaningful leads: Applying deep adversarial autoencoders for new molecule development in oncology. *Oncotarget* **2017**, *8*, 10883. [[CrossRef](#)]
129. Assmann, M.; Bal, M.; Craig, M.; D'Oyley, J.; Phillips, L.; Triendl, H.; Bates, P.A.; Bashir, U.; Ruprah, P.; Shaker, N. A novel machine learning approach uncovers new and distinctive inhibitors for cyclin-dependent kinase 9. *BioRxiv* 2020. [[CrossRef](#)]
130. Perron, Q.; Mirguet, O.; Tajmouati, H.; Skiredj, A.; Rojas, A.; Gohier, A.; Ducrot, P.; Bourguignon, M.P.; Sansilvestri-Morel, P.; Do Huu, N. Deep generative models for ligand-based de novo design applied to multi-parametric optimization. *J. Comput. Chem.* **2022**, *43*, 692–703. [[CrossRef](#)] [[PubMed](#)]
131. Li, X.; Xu, Y.; Yao, H.; Lin, K. Chemical space exploration based on recurrent neural networks: Applications in discovering kinase inhibitors. *J. Cheminform.* **2020**, *12*, 42. [[CrossRef](#)] [[PubMed](#)]
132. Zhavoronkov, A.; Vanhaelen, Q.; Oprea, T.I. Will artificial intelligence for drug discovery impact clinical pharmacology? *Clin. Pharmacol. Ther.* **2020**, *107*, 780–785. [[CrossRef](#)] [[PubMed](#)]
133. van Laar, S.A.; Gombert-Handoko, K.B.; Guchelaar, H.J.; Zwaveling, J. An electronic health record text mining tool to collect real-world drug treatment outcomes: A validation study in patients with metastatic renal cell carcinoma. *Clin. Pharmacol. Ther.* **2020**, *108*, 644–652. [[CrossRef](#)]
134. Noorbakhsh-Sabet, N.; Zand, R.; Zhang, Y.; Abedi, V. Artificial intelligence transforms the future of health care. *Am. J. Med.* **2019**, *132*, 795–801. [[CrossRef](#)]
135. Bhattamisra, S.K.; Banerjee, P.; Gupta, P.; Mayuren, J.; Patra, S.; Candasamy, M. Artificial Intelligence in Pharmaceutical and Healthcare Research. *Big Data Cogn. Comput.* **2023**, *7*, 10. [[CrossRef](#)]
136. Westreich, D.; Lessler, J.; Funk, M.J. Propensity score estimation: Neural networks, support vector machines, decision trees (CART), and meta-classifiers as alternatives to logistic regression. *J. Clin. Epidemiol.* **2010**, *63*, 826–833. [[CrossRef](#)]
137. Daoui, O.; Elkhatabi, S.; Chtita, S.; Elkhlabi, R.; Zgou, H.; Benjelloun, A.T. QSAR, molecular docking and ADMET properties in silico studies of novel 4, 5, 6, 7-tetrahydrobenzo [D]-thiazol-2-Yl derivatives derived from dimedone as potent anti-tumor agents through inhibition of C-Met receptor tyrosine kinase. *Heliyon* **2021**, *7*, e07463. [[CrossRef](#)] [[PubMed](#)]
138. Yang, H.; Lou, C.; Sun, L.; Li, J.; Cai, Y.; Wang, Z.; Li, W.; Liu, G.; Tang, Y. admetSAR 2.0: Web-service for prediction and optimization of chemical ADMET properties. *Bioinformatics* **2019**, *35*, 1067–1069. [[CrossRef](#)]
139. Kim, H.; Kim, E.; Lee, I.; Bae, B.; Park, M.; Nam, H. Artificial intelligence in drug discovery: A comprehensive review of data-driven and machine learning approaches. *Biotechnol. Bioprocess Eng.* **2020**, *25*, 895–930. [[CrossRef](#)] [[PubMed](#)]
140. Obrezanova, O. Artificial intelligence for compound pharmacokinetics prediction. *Curr. Opin. Struct. Biol.* **2023**, *79*, 102546. [[CrossRef](#)]
141. Ye, Z.; Yang, Y.; Li, X.; Cao, D.; Ouyang, D. An integrated transfer learning and multitask learning approach for pharmacokinetic parameter prediction. *Mol. Pharm.* **2018**, *16*, 533–541. [[CrossRef](#)] [[PubMed](#)]
142. Lou, H.; Hageman, M.J. Machine learning attempts for predicting human subcutaneous bioavailability of monoclonal antibodies. *Pharm. Res.* **2021**, *38*, 451–460. [[CrossRef](#)]
143. Lavielle, M. Pharmacometrics models with hidden Markovian dynamics. *J. Pharmacokinet. Pharmacodyn.* **2018**, *45*, 91–105. [[CrossRef](#)] [[PubMed](#)]
144. Lee, H.; Kimko, H.C.; Rogge, M.; Wang, D.; Nestorov, I.; Peck, C.C. Population pharmacokinetic and pharmacodynamic modeling of etanercept using logistic regression analysis. *Clin. Pharmacol. Ther.* **2003**, *73*, 348–365. [[CrossRef](#)]
145. Farhana, N.; Afendi, F.; Fitrianto, A.; Wijaya, S. Classification modeling of support vector machine (SVM) and random forest in predicting pharmacodynamics interactions. *J. Phys. Conf. Ser.* **2021**, *1863*, 012067. [[CrossRef](#)]
146. Cheng, F.; Zhao, Z. Machine learning-based prediction of drug–drug interactions by integrating drug phenotypic, therapeutic, chemical, and genomic properties. *J. Am. Med. Inform. Assoc.* **2014**, *21*, e278–e286. [[CrossRef](#)]
147. Seok, K.H.; Shim, J.; Cho, D.; Noh, G.-J.; Hwang, C. Semiparametric mixed-effect least squares support vector machine for analyzing pharmacokinetic and pharmacodynamic data. *Neurocomputing* **2011**, *74*, 3412–3419. [[CrossRef](#)]
148. Zhang, H.; Ma, J.-X.; Liu, C.-T.; Ren, J.-X.; Ding, L. Development and evaluation of in silico prediction model for drug-induced respiratory toxicity by using naïve Bayes classifier method. *Food Chem. Toxicol.* **2018**, *121*, 593–603. [[CrossRef](#)]
149. Chou, W.-C.; Lin, Z. Machine learning and artificial intelligence in physiologically based pharmacokinetic modeling. *Toxicol. Sci.* **2023**, *191*, 1–14. [[CrossRef](#)] [[PubMed](#)]
150. Keutzer, L.; You, H.; Farnoud, A.; Nyberg, J.; Wicha, S.G.; Maher-Edwards, G.; Vlasakakis, G.; Moghaddam, G.K.; Svensson, E.M.; Menden, M.P. Machine learning and pharmacometrics for prediction of pharmacokinetic data: Differences, similarities and challenges illustrated with rifampicin. *Pharmaceutics* **2022**, *14*, 1530. [[CrossRef](#)] [[PubMed](#)]
151. Grenet, I.; Yin, Y.; Comet, J.-P.; Gelenbe, E. *Machine Learning to Predict Toxicity of Compounds*; Cham: Rhodes, Greece, 2018; pp. 335–345.
152. Mendonça, V. Bayesian Pharmacokinetics: Pharmacodynamics Modeling & Simulation. *ProQuest* 2019.
153. Tannenbaum, C.; Day, D. Age and sex in drug development and testing for adults. *Pharmacol. Res.* **2017**, *121*, 83–93. [[CrossRef](#)]
154. Raies, A.B.; Bajic, V.B. In silico toxicology: Computational methods for the prediction of chemical toxicity. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2016**, *6*, 147–172. [[CrossRef](#)]
155. Hasselgren, C.; Myatt, G.J. Computational Toxicology and Drug Discovery. *Methods Mol. Biol.* **2018**, *1800*, 233–244. [[CrossRef](#)]

156. Vo, A.H.; Van Vleet, T.R.; Gupta, R.R.; Liguori, M.J.; Rao, M.S. An Overview of Machine Learning and Big Data for Drug Toxicity Evaluation. *Chem. Res. Toxicol.* **2020**, *33*, 20–37. [[CrossRef](#)] [[PubMed](#)]
157. Patlewicz, G.; Fitzpatrick, J.M. Current and Future Perspectives on the Development, Evaluation, and Application of in Silico Approaches for Predicting Toxicity. *Chem. Res. Toxicol.* **2016**, *29*, 438–451. [[CrossRef](#)] [[PubMed](#)]
158. Zakariya Yahya Algamal, M.H.L. High-dimensional QSAR prediction of anticancer potency of imidazo[4,5-b]pyridine derivatives using adjusted adaptive LASSO. *J. Chemom.* **2015**, *29*, 547–556. [[CrossRef](#)]
159. Luco, J.M.; Ferretti, F.H. QSAR Based on Multiple Linear Regression and PLS Methods for the Anti-HIV Activity of a Large Group of HEPT Derivatives. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 392–401. [[CrossRef](#)] [[PubMed](#)]
160. Ma, J.; Sheridan, R.P.; Liaw, A.; Dahl, G.E.; Svetnik, V. Deep Neural Nets as a Method for Quantitative Structure–Activity Relationships. *J. Chem. Inf. Model.* **2015**, *55*, 263–274. [[CrossRef](#)] [[PubMed](#)]
161. Shi, J.; Zhao, G.; Wei, Y. Computational QSAR model combined molecular descriptors and fingerprints to predict HDAC1 inhibitors. *Med. Sci.* **2018**, *34*, 52–58. [[CrossRef](#)] [[PubMed](#)]
162. Minerali, E.; Foil, D.H.; Zorn, K.M.; Lane, T.R.; Ekins, S. Comparing machine learning algorithms for predicting drug-induced liver injury (DILI). *Mol. Pharm.* **2020**, *17*, 2628–2637. [[CrossRef](#)] [[PubMed](#)]
163. Williams, D.P.; Lazic, S.E.; Foster, A.J.; Semenova, E.; Morgan, P. Predicting drug-induced liver injury with Bayesian machine learning. *Chem. Res. Toxicol.* **2019**, *33*, 239–248. [[CrossRef](#)] [[PubMed](#)]
164. Verma, R.P.; Matthews, E.J. Estimation of the chemical-induced eye injury using a weight-of-evidence (WoE) battery of 21 artificial neural network (ANN) c-QSAR models (QSAR-21): Part I: Irritation potential. *Regul. Toxicol. Pharmacol.* **2015**, *71*, 318–330. [[CrossRef](#)]
165. Mamoshina, P.; Bueno-Orovio, A.; Rodriguez, B. Dual transcriptomic and molecular machine learning predicts all major clinical forms of drug cardiotoxicity. *Front. Pharmacol.* **2020**, *11*, 639. [[CrossRef](#)]
166. Webel, H.E.; Kimber, T.B.; Radetzki, S.; Neuenschwander, M.; Nazaré, M.; Volkamer, A. Revealing cytotoxic substructures in molecules using deep learning. *J. Comput.-Aided Mol. Des.* **2020**, *34*, 731–746. [[CrossRef](#)]
167. Hunta, S.; Yooyativong, T.; Aunsri, N. A novel integrated action crossing method for drug-drug interaction prediction in non-communicable diseases. *Comput. Methods Programs Biomed.* **2018**, *163*, 183–193. [[CrossRef](#)]
168. Wipke, W.T.; Ouchi, G.I.; Krishnan, S. Simulation and evaluation of chemical synthesis—SECS: An application of artificial intelligence techniques. *Artif. Intell.* **1978**, *11*, 173–193. [[CrossRef](#)]
169. Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
170. Liu, B.; Ramsundar, B.; Kawthekar, P.; Shi, J.; Gomes, J.; Luu Nguyen, Q.; Ho, S.; Sloane, J.; Wender, P.; Pande, V. Retrosynthetic reaction prediction using neural sequence-to-sequence models. *ACS Cent. Sci.* **2017**, *3*, 1103–1113. [[CrossRef](#)] [[PubMed](#)]
171. Coley, C.W.; Green, W.H.; Jensen, K.F. Machine learning in computer-aided synthesis planning. *Acc. Chem. Res.* **2018**, *51*, 1281–1289. [[CrossRef](#)] [[PubMed](#)]
172. Segler, M.H.; Waller, M.P. Neural-symbolic machine learning for retrosynthesis and reaction prediction. *Chem.–A Eur. J.* **2017**, *23*, 5966–5971. [[CrossRef](#)] [[PubMed](#)]
173. Segler, M.H.; Preuss, M.; Waller, M.P. Planning chemical syntheses with deep neural networks and symbolic AI. *Nature* **2018**, *555*, 604–610. [[CrossRef](#)] [[PubMed](#)]
174. Lin, K.; Xu, Y.; Pei, J.; Lai, L. Automatic retrosynthetic route planning using template-free models. *Chem. Sci.* **2020**, *11*, 3355–3364. [[CrossRef](#)] [[PubMed](#)]
175. Gupta, R.; Srivastava, D.; Sahu, M.; Tiwari, S.; Ambasta, R.K.; Kumar, P. Artificial intelligence to deep learning: Machine intelligence approach for drug discovery. *Mol Divers* **2021**, *25*, 1315–1360. [[CrossRef](#)] [[PubMed](#)]
176. Dhamodharan, G.; Mohan, C.G. Machine learning models for predicting the activity of AChE and BACE1 dual inhibitors for the treatment of Alzheimer’s disease. *Mol. Divers.* **2022**, *26*, 1501–1517. [[CrossRef](#)] [[PubMed](#)]
177. Askr, H.; Elgeldawi, E.; Aboul Ella, H.; Elshaier, Y.A.M.M.; Gomaa, M.M.; Hassanien, A.E. Deep learning in drug discovery: An integrative review and future challenges. *Artif. Intell. Rev.* **2023**, *56*, 5975–6037. [[CrossRef](#)]
178. Tropsha, A. Best Practices for QSAR Model Development, Validation, and Exploitation. *Mol Inf.* **2010**, *29*, 476–488. [[CrossRef](#)] [[PubMed](#)]
179. Schneider, G. Automating drug discovery. *Nat. Rev. Drug Discov.* **2018**, *17*, 97–113. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.