HİTİT JOURNAL OF SCIENCE

e-ISSN: -
Volume: 1 • Number: 1
July 2024

Artificial Intelligence-assisted Drug Development

Irmak Sevval TOPCU[1,*] | Orcun AVSAR[1]

[1]Hitit University, Faculty of Art and Sciences, Department of Molecular Biology and Genetics, 19040, Corum, Türkiye.

Corresponding Author

Irmak Sevval TOPCU
E-mail: irmaktopcu00@gmail.com Phone: +90 553 546 3671
RORID: ror.org/01x8m3269

Article Information

Article Type: Review
Doi: -
Received: 15.06.2024
Accepted: 12.07.2024
Published: 31.07.2024

Cite As

Topçu I.S, Avşar O. Artificial Intelligence-assisted Drug Development. Hitit journal of Science. 2024;1(1):8-21

Peer Review: Evaluated by independent reviewers working in at least two different institutions appointed by the field editor.
Ethical Statement: Not available.
Plagiarism Checks: Yes - iThenticate
Conflict of Interest: Authors approve that to the best of their knowledge, there is not any conflict of interest or common interest with an institution/organization or a person that may affect the review process of the paper.

CRedit Author Statement
Irmak Şevval Topcu: Design of the study, writing, reviewing, editing
Orcun Avşar: Design of the study, supervisio

Artificial Intelligence-assisted Drug Development

Irmak Sevval TOPCU[1,*] | Orcun AVSAR[1]
[1]Hitit University, Faculty of Art and Sciences, Department of Molecular Biology and Genetics, 19040, Corum, Türkiye.

Abstract
Deep learning and machine learning algorithms, two types of artificial intelligence, have come to light as potential solutions to issues and roadblocks in the drug design and discovery process. Both in vitro and in silico techniques have the potential to significantly lower drug development costs when compared to conventional animal models. Early on in the drug research and development process, drug candidates with relevant therapeutic activities can be identified, unsuitable compounds with unwanted side effects can be excluded, and in vitro and in silico techniques can be used to limit the number of drug poisonings. Drug discovery procedures, illness modeling, target identification, artificial intelligence, drug screening, and molecular design can all be completed far more quickly and affordably than with conventional techniques.

Keywords: Artificial Intelligence, In silico approaches, Target identification, Target verification, Target interactions.

OVERVIEW
The multi-step process of drug research and development includes clinical testing, manufacturing approval, and drug discovery. It is costly, time-consuming, complex, and has a high attrition rate (Waring M.J. et al. 2015). Clinical trial drug attrition results in a significant loss of resources (Fleming N. 2018). Chemical and biological scientists have faced a significant problem over the past 20 years: creating effective and sophisticated systems for the targeted administration of therapeutic substances with maximum efficiency and minimal danger (Lipinski CF., 2019). Another obstacle in the process of designing and developing new drugs is the expense and duration involved in creating novel therapeutic agents (Hamet P., Tremblay J., 2017). Researchers all over the world have resorted to computational techniques like virtual screening (VS) and molecular docking, also referred to as traditional approaches, to try and minimize these difficulties and barriers; however, these methods have also introduced problems like inaccuracy and inefficiency (Hassanzadeh P. et al., 2019). Long and intricate processes like target identification and validation, therapeutic screening, lead compound optimization, preclinical and clinical trials, and manufacturing applications are all part of the drug research and design process. The process of identifying the medication that works best to treat a condition is further complicated by all these procedures. Consequently, controlling process speed and cost is the largest issue facing pharmaceutical businesses (Zhang L. et al., 2017). By providing straightforward, scientific solutions to all of these issues, artificial intelligence shortens the process's time and expense (Jordan A.M., 2018).

Machine intelligence, another name for artificial intelligence, is the capacity of computer systems to learn from inputs or past data. When a machine simulates cognitive behavior linked to learning and problem-solving in the human brain, it is said to be artificial intelligence (Goel AK, Davies J (2019) Artificial intelligence. In: Cambridge Handbook of Intelligence. Cambridge). The fields of logic, statistics, cognitive psychology, decision theory, cybernetics, computer engineering, neuroscience, and linguistics are the foundations of artificial intelligence (AI). A better understanding of AI will help to mitigate its negative effects on worker safety, health, and welfare as well as its opportunities and challenges for the future of work (Russell S.J.; Norvig, P., 2016).

1.THE EMERGENCE OF ARTIFICIAL INTELLIGENCE
Robotics is widely acknowledged as the source of artificial intelligence. The Czech word "robota," which means "robot," was originally used in the science fiction drama "Rossum's Universal Robots" by author Karel Capek in 1921. The term "robot" was made famous by Isaac Asimov in the middle of the 20th century while compiling a collection of contemporary science fiction short stories. But the earliest record of a humanoid automaton dates back to the third century in China, when Yan Shi, a mechanical expert, gave the Zhou Emperor Mu a handcrafted, humanoid mechanical figure composed of wood, leather, and synthetic materials (Hamet P, Tremblay J., 2017). Al-Jazari, a Muslim philosopher from the Golden Age, invented a humanoid robot that could strike cymbals in the 12th century. Leonardo da Vinci studied human anatomy in great detail throughout the Renaissance in order to construct his humanoid robot. Only in the 1950s were his 1495 sketches unearthed. Driven by pulleys and wires, Leonardo's robot was a knight-like device that could sit, stand, swing its arms, and move its head and jaw.

From Charles Babbage, who created the first mechanical computer in 1850, to the question "can machines think?" in 1950, computer scientists and science fiction authors were captivated by the notion of machine intelligence comparable to human intelligence. Alan Turing proposed a machine intelligence test in 1950. This test, often known as the Turing test, assesses a machine's capacity for intelligent behavior that is on par with or identical to that of a human. If "a human interrogator, after some written questions have been posted, cannot tell whether the written answers come from a human or a machine," then the computer passes the test. In order to pass the Turing test, a computer needs to be able to recognize speech, store information from what it hears or knows (knowledge representation), utilize that information to answer questions and make inferences (automatic reasoning), and recognize new patterns in order to adapt to changing conditions (ML). The computer will meet the requirements of the so-called Total Turing test if it is equipped with two more capabilities: computer vision and physical interaction. The primary focuses of AI research and development at the moment are represented by these six capabilities (Howard J., 2019).

When Arthur L. Samuel created an IBM checkers software in 1952, he popularized the phrase "machine learning." "The science and engineering of making intelligent machines" is how John McCarthy defined artificial intelligence (AI) when he first used the word in 1955. He had a significant impact on

AI's early development. He and his colleagues created the field of artificial intelligence during a 1956 conference held at Dartmouth College. This event gave rise to the term that became a new field of study spanning multiple disciplines and served as the conceptual foundation for all ensuing computer-related research and development projects (Hamet P., Tremblay J., 2017). Frank Rosenblatt created the perceptron in 1957 with the purpose of recognizing images (Rosenblatt F., 1957). The continuous back-propagation model was created by Henry J. Kelley in 1960, and Stuart Dreyfus created a more straightforward version in 1962 based solely on the chain rule (Kelley H.J., 2012; Dreyfus S., 1962). The first functional deep learning networks were created in 1965 by Ivakhnenko and Lapa (Gupta R., et al., 2021). Around 1980, Kunihiko Fukushima created the first convolutional neural network (CNN), which was modeled after the structure of an animal's visual cortex (Fukushima K., 1988).

1.1. SYSTEMATIC LEARNING

Machine learning (ML) is a subfield of artificial intelligence that allows computers to learn from data. It has become popular for using computers to make predictions, acquire cognitive insights, and assist in decision-making (Jordan M.I., Mitchell T.M., 2015). ML is a break from earlier artificial intelligence techniques, which worked by hand-coding a full set of logic rules into software in an effort to foresee every scenario that could arise. With machine learning (ML), computers can use cutting-edge software techniques to extract their own rules (Haugeland J. Artificial Intelligence: The Very Idea. Cambridge).

1.1.2. GUIDED EXPERIENCE

Using a training dataset that has been precisely labeled by a human expert, supervised learning looks for patterns and makes predictions (Maini V. et al., 2019). A radiographic data image classification algorithm can learn the correct relationship between an input image (X-ray, for example) and an output label (lung cancer) using a supervised learning training dataset. It can then use this relationship to classify unlabeled images that the computer has never seen before (Choy G. et al., 2018).

1.1.3.UNSUPERVISED LEARNING

There is no usage of a preset training dataset. The learning algorithm receives unlabeled data; without human assistance, it then finds the data's hidden structure and groups the data into clusters (Hinton G., Sejnowski T.J., 1999; James G. et al., 2017).

1.1.4.SEMI-SUPERVISED LEARNING

It's a machine learning technique for better comprehending a dataset's structure. Currently, a variety of industry sectors are producing large volumes of unlabeled data from text, audio, and images (Chapelle O. et al., 2006).

1.1.5. LEARNING REINFORCEMENT

Reinforcement learning is a type of computer experimentation that is derived from basic learning theory in psychology. It is a training approach that is based on rewarding good behaviors and penalizing undesired ones (Thorndike E., 1932, Varian H., 2019). With reinforcement learning, a machine may be taught the right answers by applying incentives and penalties in the same way that humans learn by making mistakes (Sutton RS, Barto AG., 2018). Reinforcement learning can be utilized in conjunction with neural networks to train a robot to grasp objects it has never seen before or to operate autonomous vehicles (Knight W., 2017).

1.1.6. DEEP NEURAL NETWORKS

Neural networks that are fully connected and have several hidden layers. There are several nonlinear processing units in each hidden layer. DNNs use several neurons in numerous layers to automatically extract features at hierarchical levels (D'Souza S., et al., 2020).

1.1.7.DEEP LEARNING

According to Goodfellow, Bengio, and Courville (2016), deep learning is a subtype of neural networks that recognizes patterns using many processing layers of coupled neurons between input and output layers. In the areas of speech recognition, image identification, and natural language understanding, deep learning algorithms have made great progress (Krizhevsky A., et al, 2012; Hinton G., et al, 2012; Hirschberg J., et al, 2015).

2.GENERAL USAGE AREAS OF ARTIFICIAL INTELLIGENCE

2.1. ELECTRONIC DEVICES

Functional sensors are not as valuable as advanced or smart sensors. To monitor various parameters, these smart sensors can be surgically implanted in the body, fastened to safety gear, or fastened to any item (Nag A., et al., 2017; Howard J., 2019). The Internet of Things (IoT) can be created by connecting any product or device with integrated sensors to the internet and other similar devices (Chui M., et al., 2010). Applications of artificial intelligence are being brought into a wide range of industries, including banking, insurance, criminal justice, healthcare, and national security (West D.M., Allen J.R., 2018). Cutting-edge sensor systems can "sense" their surroundings using deep learning models, much like how humans perceive sound and vision (Howard J., 2019).

2.2.ROBOTIC DEVICES

"Cloud robotics" allows one AI-enabled robotic device to upload its learning experience to all other robots that are linked. (B. Kehoe and others, 2015). All cloud-connected robotic devices can be updated to use the new procedure when a robot's output reveals a safer way to complete a task at work. Robotics can learn more effectively through universal robotic upgradability in a cloud-connected network than through human learning, which is individually dependant (Pratt G.A., 2015).

2.3.DECISION SUPPORT SYSTEMS

A multipurpose informative tool with AI support can be used to extract information from data for applications involving decisionmaking (Howard J., 2019). Utilizing data already recorded in management information systems, technologies are being used to support business decisions as a result of the notion that computers should support decision makers (Bonini C.P., 1963; Pervan G., Willcocks L., 2005). Many industry

sectors, particularly the medical field, use ML-enabled DSSs for decision-making (Kononenko I., 2001; Topol E., 2019). Clinical DSSs are marketed as having the ability to increase diagnostic accuracy and assist physicians in understanding the intricate relationships between clinical variable scores. The healthcare industry generates large amounts of data, which make them ideal learning inputs for ML-enabled DSSs (Ehteshami Bejnordi B., et al., 2017; Obermeyer Z, Emanuel E.J., 2016; Phillips-Wren G., 2012). Several studies that have used ML-enabled DSSs to date include:
Lung cancer screening (Ardila D., et al., 2019),
Detection of pulmonary tuberculosis (Lakhani P., Sundaram B., 2017),
Determination of diabetic retinopathy (Gulsen V., et al., 2016; Kanagasingam Y., et al., 2018),
Skin cancer diagnosis (Esteva A., et al., 2017),
Anticancer medication response prediction in precision oncology therapy (Azuaje F., 2019; Tan M., 2016),
Progress has been made in areas such using retinal pictures to predict cardiovascular risk factors (Poplin R., et al., 2018).

Transforming research findings into clinical advancements is still a difficult undertaking, despite the early research accomplishments using machine learning to huge medical datasets holding significant potential in enhancing the quality of healthcare (Deo R.C., 2015). For instance, an AI-enabled ML image classifier for melanoma skin cancer that is trained solely on fair skin types will reinforce current health disparities rather than serve as a means of eradicating them (Adamson A.S., Smith A., 2018).

3.DRUG DISCOVERY PROCESS AND ARTIFICIAL INTELLIGENCE

3.1. PROCESS OF DRUGS DISCOVERY

3.1.1.DISEASE MODELLING AND TARGET IDENTIFICATION

The success rate of drug development is greatly impacted by disease modeling and target identification, which is a crucial initial phase in the drug discovery process (Pun F.W., et al., 2023). Furthermore, target identification helps researchers understand the mode of action of unknown medications, which makes it a critical step in the discovery and development of new drugs (Schenone M., et al., 2013). Researchers can more effectively tailor a medication for a specific ailment or disease by identifying the molecular target of that medication (McFedries A., et al., 2013; Hughes J.P., et al., 2010). To maximize medication selectivity and minimize possible adverse effects, target identification is also crucial (Schenone M., et al., 2013; Hughes J.P., et al., 2010).

A molecule must be “druggable” in order to have even the remotest chance of being a target for medication. The field of drug development is shifting toward the application of novel design principles to molecules, connecting them to difficult biological targets for novel medications of the future or novel approaches to dosage modification. The conventional pharmaceutical industry concentrates on creating tiny compounds that are orally bioavailable and have specific objectives (Sarkar C., et al., 2023).

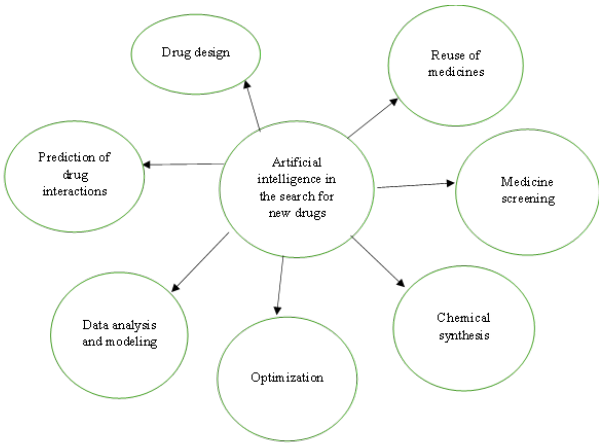

Figure 1. Artificial Intelligence’s role in medication discovery. Various steps of drug discovery, including as drug design, chemical synthesis, drug reuse, drug screening, drug interaction prediction, optimization, data analysis, and modeling, can benefit from the application of artificial intelligence.

Millions of molecules may be present in datasets used by pharmaceutical companies for medication research, but conventional machine learning techniques may not be able to handle this volume of data. Though computational models based on the quantitative structure-activity relationship (QSAR) can rapidly predict a large number of compounds or basic physicochemical parameters like logP (partition coefficient), they are not very good at predicting complex biological properties. Additionally, QSAR-based models have issues with experimental data error and insufficient experimental validation on training sets. In order to address these issues, new AI techniques like deep learning (DL) and associated modeling investigations can be used for large data modeling and analysis-based safety and efficacy evaluations of pharmaceutical compounds (Paul D., et al., 2021).

3.1.2. DRUG SCREENING WITH ARTIFICIAL INTELLIGENCE

3.1.2.1. PHYSICAL AND CHEMICAL PROPERTIES PREDICTION

When developing a new drug, physicochemical characteristics like solubility, intrinsic permeability, degree of ionization, and partition coefficient (logP) should be taken into account as they have an indirect impact on the pharmacokinetic characteristics of the drug and its target receptor family (Zang Q., et al., 2017). A variety of AI-based instruments are available for physicochemical property prediction. For instance, ML trains the program using massive data sets produced during prior chemical optimization (Yang X., et al., 2019). Drug design algorithms use chemical descriptors, such as coordinates of 3D atoms, electron density surrounding the molecule, and SMILES sequences, to produce appropriate molecules via DNN and subsequently predict their attributes (Baringhaus K.H., Hessler G., 2018).

3.1.2.2. THE BIOACTIVITY PREDICTION

Drug molecules’ ability to generate a therapeutic response is contingent upon their affinity for the target protein or receptor; those that do not exhibit any interaction with the

targeted protein will not be effective. Toxic interactions between produced medication molecules and undesirable proteins or receptors can also occur in certain situations. As a result, drug-target interaction prediction greatly depends on drug target binding affinity (DTBA). AI-based techniques can calculate a drug's binding affinity by considering the characteristics or similarities between the drug and its target. While similarity-based interactions consider the similarity between the drug and the target and presume that similar drugs will interact with the same targets, feature-based interactions identify the chemical moieties of the drug and the target to determine feature vectors (Öztürk H., et al., 2018).

To predict drug-target interactions, a variety of techniques, such as machine learning and deep learning, have been employed. To determine DTBA, machine learning (ML) techniques like Kronecker regularized least squares (KronRLS) assess how similar medicines and protein molecules are. Similarly, SimBoost took into account both feature-based and similarity-based interactions while predicting DTBA using regression trees (Öztürk H., et al., 2018).

3.1.2.3. TOXICITY PREDICTION

Any pharmacological molecule's predicted toxicity can be utilized as a guide to prevent harmful consequences, and cell-based in vitro experiments are frequently employed as pilot research. The expense of drug discovery rises when research on animals are carried out to ascertain a compound's toxicity right after. Cutting-edge AI-based methods either predict a compound's toxicity based on input features or search for commonalities between compounds. By identifying static and dynamic properties like molecular weight and Van der Waals volume within the chemical descriptors of molecules, an ML algorithm named DeepTox outperformed all other methods. It was also able to predict a molecule's toxicity with high efficiency using 2500 predefined toxicophore properties (Mayr A., et al., 2016).

3.2. DESIGN OF DRUG MOLECULES WITH ARTIFICIAL INTELLIGENCE

The necessity of developing novel medications is underscored by the advent of pandemics and epidemics, as well as the growth of grave illnesses like cancer and heart disease. Target selection, validation, high-throughput screening, animal studies, safety and efficacy protocols, clinical trials, and regulatory approval are all necessary steps in the often multi-step process of drug discovery (Vamathevan J., et al., 2019). Certain phases of this process, like finding new targets, assessing drug-target interactions, researching disease mechanisms, and enhancing small-molecule drug design and optimization, can benefit from the application of artificial intelligence-based techniques (Jeon J., et al., 2014; Katsila T., et al., 2016; Lee L., et al., 2019; Nicolaou C.A., Brown N., 2013; Vamathevan J., et al., 2019). These techniques can also be applied to the investigation of pharmacological efficacy, response, and resistance as well as the identification and development of prognostic biomarkers (Qureshi R., et al., 2022).

3.2.1. IDENTIFICATION OF THE TARGET IN DRUG DISCOVERY

Target identification is the process of finding molecules (typically proteins) that have the ability to change a disease state. Numerous data sources, such as gene expression profiles, protein-protein interaction networks, genomic, and proteomic data, can be evaluated using machine learning (ML) methods to find possible targets that may be involved in disease pathways (Sliwoski G., et al., 2014).

Determining the cause of the illness and the target is the first stage in defining a target (Lv B. M., et al., 2014). Tree-based approaches, GNNs, and graphs can be used to determine the causal links between genes and diseases. It was also suggested to identify genes linked to druggable morbidity using a decision tree-based meta-classifier that was trained on a network topology that included protein-protein, metabolic and transcription relationships, tissue expression of proteins, and subcellular localization (Qureshi R., et al., 2023). Key characteristics from the decision tree included regulation by several transcription factors, centrality in metabolic pathways, and extracellular placement. Based on characteristics including protein-protein interaction, gene expression, DNA copy number, and the presence of mutations, ML-based techniques categorized proteins as therapeutic targets or non-targets for particular diseases like lung, pancreatic, and ovarian cancer (Jeon J., et al., 2014).

3.2.2. PREDICTION OF TARGET PROTEİN STRUCTURE

The development of the disease involves many proteins, some of which are overexpressed. Predicting the target protein structure while creating a therapeutic molecule is therefore crucial for the selective targeting of disease. By anticipating the 3D protein structure as the design aligns with the target protein region's chemical environment, artificial intelligence can support structurebased drug discovery by assisting in the prediction of a compound's effect on the target as well as safety concerns prior to synthesis (Wann F., Zeng J.M., 2016). In order to predict the 3D target protein structure, AlphaFold, an artificial intelligence tool based on DNNs, was used to analyze the angles of peptide bonds and the distances between adjacent amino acids. It demonstrated excellent results, correctly predicting 25 out of 43 structures (Paul D., et al., 2021).

3.2.3. DRUG-PROTEIN INTERACTION PREDICTION

The effectiveness of therapy is greatly dependent on drug-protein interactions. Understanding drug efficacy, permitting the bait and switch of medications, and avoiding polypharmacology all depend on the ability to predict how a drug will interact with a receptor or protein. Different AI techniques have improved therapeutic efficacy by accurately predicting ligand-protein interactions (Wann F., Zeng J.M., 2016).

Because AI can anticipate drug-target interactions, it can also be used in Phase II clinical trials to assist minimize polypharmacology and reuse current medications (Mak K.K., Pichika M.R., 2019). This also lowers costs because it is more expensive to relaunch a current drug than to introduce a brand-new medicinal entity (Paul D., et al., 2021). The potential

for polypharmacology—a drug molecule's propensity to bind with many receptors and cause off-target adverse effects—can also be predicted by drug-protein interactions (Li X., et al., 2017). By using polypharmacology logic to design novel molecules, artificial intelligence can contribute to the production of safer pharmaceutical molecules (Reddy A.S., Zhang S., 2014).

4. ARTIFICIAL INTELLIGENCE ALGORITHMS USED IN DRUG DISCOVERY PROCESS

4.1. MACHINE LEARNING (ML) ALGORITHIMS

Supervised and unsupervised learning are the two primary categories of machine learning algorithms. Unsupervised learning detects patterns in a set of instances, frequently without labels for the instances, and the data is frequently transformed to a lower dimension to recognize patterns in high dimensional data using unsupervised learning algorithms before recognizing patterns. Supervised learning learns by training instances with known labels to determine the labels for new instances. Not only is unsupervised learning more effective in a low-dimensional space, but dimensionality reduction also makes the recognized model easier to understand. Semi-experienced and reinforcement learning can be created by combining supervised and unsupervised learning; both functions can be used to different types of data (Rifaioglu A.S., et al., 2019).

ML algorithms are utilized in the drug development process to create a variety of models that forecast the chemical, physical, and biological characteristics of substances (Patel L., et al., 2020). All phases of the drug discovery process, including identifying novel drug uses, forecasting drug-protein interactions, determining drug efficacy, supplying safety biomarkers, and maximizing molecular bioactivity, can benefit from the application of machine learning (ML) algorithms (Patel L. et al., 2020).

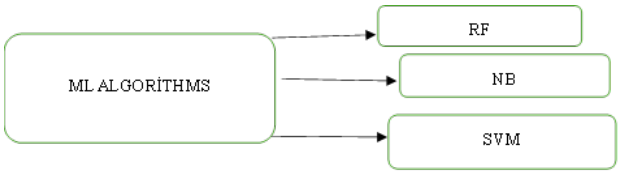

Figure 2. Commonly used ML algorithms.

RANDOM FOREST (RF)

RF is a popular method that is specifically made for big datasets with plenty of characteristics. It makes things easier by eliminating outliers (Outliers are values that deviate significantly from the general trend in the data. They need to be taken into account because they can mislead the ML model, affect its accuracy and cause poor performance. Random forest performs better when predicting variables like the Human Development Index (HDI) when techniques like winsorizing and random oversampling are used to handle outliers and imbalanced data (Notodiputro K.A., Sartono B., Zubedi F., 2022)) and categorizes and identifies datasets according to the relative features that are classified for a certain algorithm. In addition to being trained for accessibility using a variety of huge inputs, variables, and data gathering from numerous databases, it is helpful in a variety of contexts, including referring to missing data, working with outliers (For instance, the random forest method can be requested to choose the most valuable characteristic out of x attributes. If desired, this information can then be utilized in another desired model), and predicting features for classification (Breiman L., 2001). Many independent decision trees make up the mathematical process that underpins RF as a whole; each tree determines a forecast, and the tree with the greatest number of votes is deemed optimal (Sarica A., et al., 2017). By combining numerous predictions instead of concentrating on just one, multiple decision trees reduce individual errors (Patel L. et al., 2020). Regression, classifiers, and feature selection are the three main uses of RFs in drug discovery. Accelerating the training process, employing fewer parameters, loading missing data, and merging nonparametric data can be added to the list of primary factors that go along with RF in drug development (Cano G., et al., 2017). Multivariate RFs are experts in reducing error by calculating different error estimation methods inside the system. By feeding in data with combinations of genetic and epigenetic characteristics, the computational framework enables the framework to predict the mean and confidence interval of medication reactions. This is a crucial characteristic needed to analyze any medication that will be used in clinical trials (Rahman R., et al., 2017).

NAİVE BAYESİAN (NB)

A subset of supervised learning techniques known as NB algorithms is now a vital tool for classification in predictive modeling. Depending on the input features, factor correlation, and dimensionality of the data, standard NB algorithms can be one of the most effective methods for classifying dataset features. These methods increase the accuracy of retrieved datasets, which frequently come from large, mixed sources (Bielza C., Larrañaga P., 2014; Gilboa E., et al., 2013; Kim S.B., et al., 2006; Ratanamahatana C., Gunopulos D., 2010; Sun H., 2005).

SUPPORT VECTOR MACHINE (SVM)

SVMs are supervised learning algorithms that are used in the drug discovery process to derive a hyperplane and divide classes of compounds according to a feature selector. It creates an endless number of hyperplanes by taking use of commonalities across classes. It trains on linear data by projecting classes of chemicals into chemical feature space, based on features that are chosen. A hyperplane used to categorize data points by establishing decision boundaries is an ideal hyperplane that is obtained by eliminating the largest margin between classes in N (number of features) dimensional space (Heikamp K., Bajorath J., 2013). SVM's capacity to differentiate between active and inactive compounds and rank compounds in each database makes it a crucial tool for drug discovery. Regression models are essential for figuring out how a medicine and ligand interact since they make predictions by running a query against databases. Multiple situations can be associated with SVM when multiple active compounds of interest are screened against a single protein. SVM classification's primary focus is on binary class prediction, which includes a subset that can differentiate between active and inactive chemicals and substances (Patel L. et al., 2020).

SVM is especially made to incorporate ligands and target proteins as an essential part of modeling drug-target interaction (Heikamp and Bajorath, 2013). It can rate compounds from various databases according to their likelihood of being active for any computational screening in the drug discovery process. By training the algorithm with different descriptors for feature selectors, such as target protein and 2D fingerprints, the procedure can be altered. Depending on which way a chemical is positioned relative to the hyperplane, a negative or positive class label is created, resulting in a ranking of compounds from most selective to least selective (Wassermann A.M., et al., 2010; Hinselmann G., et al., 2011). For non-linear data, kernel functions are employed to optimize outcomes. According to Patel L. et al. (2020), kernel functions plot data in a higher dimensional space that allows for class classification.

4.2. DL ALGORITHMS AND ARTIFICIAL NEURAL NETWORKS

The goal of artificial neural networks (ANNs) is to simulate how neurons behave in the natural world. Several artificial neurons arranged in ordered layers make up a common artificial neural network architecture (Yang X., et al., 2019). Its most basic configuration comprises of three layers of neurons that communicate with one another, just like the human brain does. Data input occurs on the first layer, information processing occurs on the hidden layer, and output is the last layer. When every node in one layer of a feed-forward network is connected to every other layer, the neurons in an artificial neural network (ANN) are said to be dense or fully connected. Only these types of networks are referred to as multilayer perceptrons (MLPs; multiple hidden layers), dense neural networks, or complete neural networks. Stated otherwise, a network is deeper the more hidden levels it contains. The depth of the model is determined by the length of the chain connecting the many functions that make up these networks. This idea gives rise to the term “deep learning,” which describes learning systems with several information processing layers that may simulate high-level abstractions in data (Lavecchia A., 2019).

In practically every scientific and technological discipline, deep learning algorithms are acknowledged as one of the most advanced areas of development and research. DL algorithms have made it possible for computer models to learn how to represent multidimensional data through abstraction and have helped ML algorithms overcome a number of obstacles (Patel L. et al., 2020).

DL algorithms are now the standard approach for lead molecule, target, and drug activity prediction in the drug discovery process. Neural network systems, which are used to construct systems capable of complicated data recognition, interpretation, and production, are frequently the foundation of deep learning. Deep neural networks (DNNs), recurrent neural networks (RNNs), and convolutional neural networks (CNNs) are the primary subsets of neural networks that are being utilized in drug discovery (Dana D. et al., 2018; Korotcov A., et al., 2017; Ekins S., 2016).

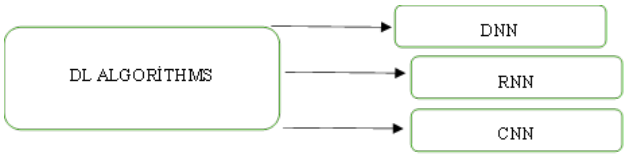

Figure 3. Commonly used DL algorithms.

DEEP NEURAL NETWORKS (DNN)

From the input layer to the hidden layer and finally to the output layer, DNNs operate on a single path data stream. Typically, supervised learning algorithms that have been trained are used to identify the outputs. A DNN may be trained to accomplish complicated tasks using guided reinforcement learning and supervised learning techniques. While a predictive DNN can forecast the chemical characteristics of novel compounds, a generative DNN may create new chemical compounds from preexisting libraries and training sets (D’Souza S., et al., 2020; Baskin I.I., et al., 2016). The correlation between these substances’ chemical structure and activity is ascertained by the utilization of QSAR models. One of the most sophisticated applications of deep learning (DL)-based artificial intelligence (AI) in drug discovery and development today is QSAR analysis, which gives scientists access to two-dimensional chemical structures and physicochemical characteristics that are associated with a molecule’s activity. Additional research into the geometric structure influencing ligand-target interactions has been made possible by 3D-QSAR (Chen R., et al., 2018; Ghasemi F., et al., 2018).

RECIPIENT NEURAL NETWORKS (RNN)

Sequence prediction was the original purpose of RNN creation. These networks only accept an input stream with varying lengths (Askr R., et al., 2023). Self-iterative or feedback connections between neurons in various levels are what distinguish them. such loops in a network, they feature feedback components to reuse internal information and function especially well with sequential data, such text, phrases, and protein sequence data. To get around the challenges of storing long-term data, they are additionally outfitted with an internal memory.

The chemical synthesis and characterisation phase becomes significant after the initial work on target discovery has been finished and a more effective technique for target-molecule interaction has been created. The majority of algorithms for new drug design and discovery use the descriptive simplified molecular input line input system (SMILES) nomenclature, which is a crucial aspect at this time. The lengthy short-term memory subset of the RNN type has evolved into a dependable, standardized technique for constructing novel chemical structures. When it comes to utilizing neurons connected to the same hidden layer to create an inputoutput processing loop, RNNs are far more beneficial algorithms than DNNs and feed-forward neural networks (Patel L. et al., 2020).

CONVOLUTIONARY NEURAL NETWORKS (CNN)

Developed to handle growing levels of complexity as well as

data preparation and aggregation, CNNs are a high-potential type of artificial neural network (ANN) that receives inputs, weights some of the inputs, and then enhances the ability to distinguish data (Yamashita R. et al., 2018). A convolutional layer with parameters made up of a collection of filters, or kernels, is what distinguishes convolutional neural networks (CNNs) from other types of neural networks. CNNs are designed to resemble the receptive field of the human visual cortex, where neurons react to stimuli. Local filters are what these cells do throughout the input space.

CNNs may process data in four steps and are among the most versatile algorithms for handling both image and non-image data (Askr H., et al., 2023):

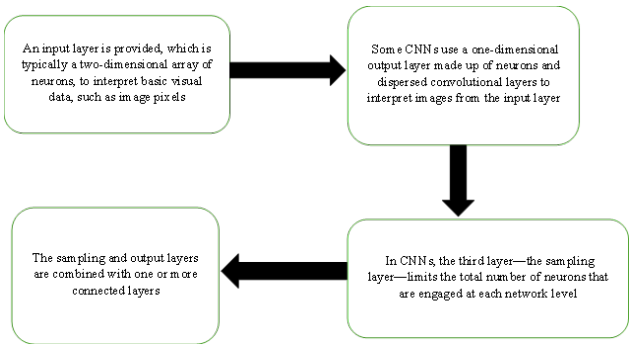

Figure 4. Stages of processing CNNs.

This idea of a network may make it easier to retrieve pertinent visual data in smaller, more manageable pieces. Neurons in a CNN are in charge of the preceding layer's group of neurons (Askr H., et al., 2023).

Four steps are involved in building the CNN when the input data is integrated into the convolutional model (Askr H., et al., 2023):

Convolution: Using the given data, a feature map is created and then put through an objective.

Maximum Pooling: Based on the supplied modifications, this aids CNN in identifying an image.

Flattening: At this point, the data is standardized for CNN's analysis.

Complete Linking: The process of generating a model's loss function is frequently referred to as the "hidden layer".
Image recognition, image analysis, video analysis, picture segmentation (splitting an image into regions with distinct features), and natural language processing (NLP) are among the tasks performed by CNNs (Chauhan et al. 2018; Tajbakhsh et al. 2016; Mohamed et al. 2020).

CNNs are among the most useful tools in the drug development process for target and lead identification and characterisation, protein-ligand scoring, and in silico target-lead interaction screening. Furthermore, CNNs have been utilized in the development of motility models that depict how cancer cells respond to various forms of therapy (Dana D., et al., 2018; Mencattini A., et al. 2020; Ragoza M., et al. 2017; Rathi P.C., et al. 2020; Reher R., et al., 2020).

5. *In Silico* APPROACH

These days, with the aid of modern computers and information technology, the procedures involved in medication development, optimization, and discovery have changed due to the rapid evolution of technology. In the biomedical field, the optimization process from hit detection and hit to routing has been facilitated and accelerated by the use of computer-aided or in silico design utilizing computational tools (Ekins S., et al., 2009).

To find hit and lead compounds, the drug discovery industry often employs one of two models: the phenotype- or target-based method. These vary in ways that help identify therapeutic targets and choose or optimize small molecules (Dodd F.S., 2005; Swinney D.C. and Anthony J., 2011). The phrase "therapeutic target" refers to the location of the substance's binding that will facilitate the substance's biological activity (Andrade E.L. et al., 2016).

The phenotype screening strategy, also known as advanced or classical pharmacology, uses better disease-relevant tests (such as isolated tissue or animal models of the disease, cell-based phenotypic analysis) to identify drugs based on their physiological effects. Through the interaction of several targets (receptors, transcription factors, enzymes, etc.) of a previously undisclosed target, this strategy may lead to the identification of a molecule that modifies the illness phenotype (Dodd F.S., 2005; Swinney D.C., 2012).

The two broad categories of approaches utilized in computer-aided drug design (CADD) are ligand-based and structure-based. Structure-based CADD is recommended when the target protein's structure is known, particularly for soluble proteins that crystallize readily. In the event that target structure information is lacking, ligand-based CADD is utilized by building predictive, quantitative structure-activity relationship (QSAR) models and using the knowledge of known active and inactive compounds through chemical similarity searches (Kapetanovic I.M., 2008; Katsila T., et al., 2016). Drug productivity, speed, and costeffectiveness can all be rationalized and enhanced by using ligand- and structure-based steps in the discovery process, such as compound generation by virtual screening, predicting the binding free energy between a ligand and a receptor, and optimizing high affinities (Sliwoski G., et al., 2014).

5.1. TARGET IDENTIFICATION AND VALIDATION FOR THERAPEUTICS

In target-based drug development, targets are found through a range of molecular techniques and instruments, such as the evaluation of the genome and proteins (proteomics) linked to a disease in humans. Targets related to human pathology can be found by utilizing a variety of molecular techniques, including RNA interference, zinc finger proteins, antisense oligonucleotides, tissue and cell microarrays, nucleic acid microarrays, and protein microarrays (Terstappen G.C., et al., 2007; Wang S., et al., 2004). The target is identified first in the

phenotypic-based approach, which observes the substance's activity beforehand.

Reverse convolution is another term for the target identification procedure in this method. Chemical proteomics-based methods (affinity chromatography, activity-based protein profiling, label-free techniques), expression cloning methods, in silico methods, and others can be used to identify targets (Terstappen G.C., et al., 2007; Lee J., Bogyo M., 2013). Validation of the treatment target is necessary after identification. Here, the objective is to determine whether altering the therapeutic target will result in a believable biological response. To this end, validation methods include altering the target in disease-affected humans as well as using whole animal models and in vitro tools (Hughes J.P., et al., 2010; Terstappen G.C., Reggiani A., 2001).

Three categories comprise the most commonly recognized standards for target validation in drug discovery (Andrade E.L., 2016):
1-Expression of the target protein or mRNA in appropriate cell types, animal models, or patient target tissues
2-Target modulation produces the intended functional effect in cell systems.
3-Prove that the target is responsible for the disease phenotype in patients or animal models.

Typically, in vivo or in vitro experiments are used to get the first steps of therapeutic target validation. These are then followed by the use of immunohistochemistry or in situ hybridization techniques to express messenger RNA or proteins in human samples, respectively. Though the first method that springs to mind is protein characterization, this approach may be hampered by the absence of particular antibodies directed against a particular target; additionally, target validation is rarely, if ever, thought to be achieved solely by the target protein's association with diseased or target tissue (Lindsay M.A., 2003). It's also necessary for the target to have functional significance to disease modification. Using small molecule inhibitors, antisense oligonucleotides, and siRNA, target validation can also be studied in transgenic and gene knockout animals; however, it should be noted that animal models frequently do not exhibit the exact disease phenotype or share the same pathophysiology as observed in patients. Targets frequently result in differing tissue expression and distribution in animal models than in human models. Moreover, pathogenic pathways in humans can have a distinct mechanism of action and differ evolutionarily from those in animal models. It is best to confirm a target using at least two distinct methods before moving on to the rigorous clinical stage of drug development in order to prevent all of these issues (Andrade E.L., 2016).

Like the more widely used biological phrases *in vivo* and *in vitro*, the term "*in silico*" refers to investigations carried out by computers. It explains how data is utilized to build computational models or simulations that can be used to forecast outcomes, put forth theories, and eventually result in new medical discoveries or advancements in therapy. The benefits of in silico investigations are their low cost, quick implementation, and capacity to minimize animal exploitation. This technique has been employed as a means of expediting the identification of promising novel therapeutics. Toxicology and pharmacokinetic research, as well as the investigation of structure-activity connections, are all included in the construction of *in silico* drug prototypes (Ekins S., et al., 2009). To effectively direct the development of new drugs through the execution of *in vitro* and *in vivo* research, *in silico* studies are crucial.

Homology modeling in the context of in silico pharmacodynamics is predicated on amino acid sequence homology, which offers details on structural and functional similarities. Therapeutic target structures are mapped using this technique, which also covers the three-dimensional structure of the targets (Ekins S., et al., 2009).

Molecular docking, which predicts the bioactive conformation of a small molecule at the binding site of a macromolecule, is another technique frequently employed for pharmacodynamic evaluation. This approach determines the relevant binding affinity after providing a good approximation of the predicted shape and fit of the ligand in the protein cavity (Lengauer T., Rarey M., 1996).

Via the use of three-dimensional macromolecular data on the topological arrangement of biological information as a prerequisite for detailed information, ligand-based virtual screening is based on virtual screening. Target-based virtual screening, which is based on receptor structure, selects compounds for biochemical or biological testing by analyzing vast compound databases using molecular docking techniques to establish an ideal chemical and biological space (Andrade E.L., 2016).

5.3. COMPUTER AIDED DRUG DESIGN (CADD)

Using a variety of computer tools, CADD integrates computational chemistry, molecular modeling, molecular design, and rational drug design to find and create a therapeutic development lead (Muegge et al., 2017). CADD employs two distinct methodologies, namely structure-based drug design (SBDD) and ligand-based drug design (LBDD), contingent upon the accessibility of three-dimensional protein or ligand structures (Vemula D., et al., 2023).

Structure-Based Drug Design (SBDD): Characterizing the binding site cavity and having access to the therapeutic target protein's three-dimensional structure are the two primary components of structure-based drug design (Kawato et al., 2015). SBDD has surfaced as a potential method in the pharmaceutical sector for ligand generation and optimization (Gurung et al., 2021; Jorgensen W.L., 2004; Park H., et al., 2012).

Ligand-Based Drug Design (LBDD): This approach is employed in situations where three-dimensional receptor data is unavailable. Understanding the chemicals that attach to the desired biological target is the foundation of the technique. By using a known ligand as a target, LBDD techniques establish a structure-activity relationship (SAR) between the ligand's activities and physicochemical characteristics. This information can be used to guide the creation of novel

medications with increased activity or to improve currently available ones (Yu and MacKerell, 2017).

6. ARTIFICIAL INTELLIGENCE IN DRUG DOSAGE FORM DESIGN

For biological compartments in the human body system to comprehend the impact of drug delivery, physicochemical barriers are essential. Depending on the route of administration, one of the most crucial parameters for keeping an eye on a successful drug delivery system is the penetration rate. For instance, after entering the stomach, the medication taken orally needs to pass through the intestinal or gastric epithelium. This step is crucial for the drug's continued bloodstream dissemination. The process of delivery involves moving the medication through the bloodstream to a specified tissue or site (Bhhatarai B., et al., 2019; Chavda V.P., 2019; Siepmann J., Siepmann F., 2012; Das P.J., et al., 2016; Colombo S., 2020). The way a medicine interacts with biological components greatly affects how the drug behaves in the body at the end. The drug's molecular characteristics control the process up to the final state. Drugs can either actively or passively aid in their penetration. Drug distribution is predicted via computational analysis using in silico models, which are based on the molecular characteristics of the drug. Passive permeation is ineffective for small, physiologically active compounds and necessitates a specific delivery method. Membrane transport drives the process of active permeation, which is dependent on intricate biological interactions. The pharmacokinetic properties of the drug delivery system can be studied with the aid of numerous specific parameters employed in this intricate process. Research units can be better understood and multi-layered data can be thoroughly analyzed thanks to artificial intelligence. In order to discover the best outcomes with parameter evaluation, the model to be applied methodically is based on a number of criteria, including simulation, scoring, and refinement at each stage of the inquiry. Moreover, AI is used to investigate how a drug delivery method affects the drug's pharmacokinetics in order to improve data prediction for continuous improvement, precise comprehension of the medication's interaction with biology, and efficient comprehension of toxicity and distribution. AI gathers data from many sources and creates indicators according to the chosen drug delivery system's performance. The efficacy of treatment is contingent upon the precision with which AI selects drug delivery devices. The goal of artificial intelligence is to apply current treatments to newly discovered diseases. It is helpful in the drug discovery process in addition to the drug reuse approach. Formulation, pharmacokinetics, and medication development are influenced by the needs of the patient and the condition of the illness (Vora K.L., et al., 2023).

7. ARTIFICIAL INTELLIGENCE IN MEDICINE DISTRIBUTION

7.1. ARTIFICIAL INTELLIGENCE TO DEVELOP ORAL SOLID DOSAGE FORM

Since solid dosage forms are the most convenient to use and promote disease compliance, individuals choose to take them in the form of tablets, granules, and powdered medications (Jiang J. et al., 2022). In the pharmaceutical industry, tablets are one of the most popular formats. Preparing tablets for use entails a number of aspects. The formulator has established these characteristics to fulfill the unique demands of the target patient population. A variety of excipients are put into tablets to manage the intended product outcome, such as tablet disintegration, dissolution, and drug release. Artificial intelligence can be used to forecast drug release in the setting of systemic drug administration and assist in examining the desirable relevant aspects of improved medication formulations. For the purpose of developing solid dosage forms, artificial neural networks and their subfields, such as neural networks and genetic algorithms, are used to improve comprehension of inputs and outputs. Genetic algorithms are employed to forecast outcomes from the usage of input parameters, however artificial neural networks offer superior prediction skills for solid dosage forms (Galata D.L. et al., 2021, Ghourichay M.P. et al., 2021; Navya K. et al., 2022).

7.1.1. Drug Release Prediction Through Formulations

The release of drugs from oral solid dosage forms advances our knowledge of important material characteristics and processing variables. Compression parameters, such as the pressure applied to regulate tablet hardness, the geometric orientation of the tablets, and drug loading properties, are factors that influence drug release. In the formulation of drugs, artificial intelligence is used to predict drug release. As a result, only a small number of runs are needed to optimize the batch, which further reduces labor and expenses during the manufacturing and pilot batch scale-up processes. Artificial intelligence can also be used to predict drug release profiles and dissolution profiles, as well as investigate disintegration time to effectively select the best batch for subsequent scale-up (Vora K.L. et al., 2023).

7.1.2. Applications of Artificial Intelligence for Formulation of Tablet Defects

Tablet photos are analyzed using artificial intelligence algorithms and computer vision techniques, which makes it possible to automatically and effectively detect flaws like cracks, discolorations, or variations in size and shape. The method gains a high degree of accuracy by accurately classifying and identifying various sorts of errors through the training of AI models on massive datasets of annotated photos. The interior structure of tablets has been studied using conventional techniques like Xray computed tomography, however these techniques still take time and interfere with the need for quick tablet production. To find tablet flaws, deep learning and X-ray tomography are combined. Not only does this AI-powered detection increase problem identification speed and accuracy, but it also minimizes human mistake and subjective judgment by reducing the need for manual inspection. AI systems' real-time monitoring capabilities allow for the prompt identification of flaws, which allows for prompt response and can stop faulty tablets from being sold. In the end, incorporating AI into tablet defect detection raises productivity and enhances product quality while guaranteeing the security and effectiveness of pharmaceuticals (Vora K.L. et al., 2023).

7.1.3. Artificial Intelligence for Prediction of Physicochemical Stability

AI can predict the stability of oral formulations by analyzing and interpreting large datasets containing drug properties, formulation parameters, and environmental conditions. AI models can assess factors like drug degradation, interactions with excipients, and environmental influences on formulation stability. These capabilities are achieved by utilizing machine learning algorithms and computational models. With the use of AI's predictive skills, researchers may improve formulation designs, spot any stability problems early in the development process, and make wise decisions that will extend the shelf life and effectiveness of oral dosage forms. Artificial intelligence (AI) integration in stability prediction leads to more economical and effective drug development procedures, which in turn provides patients with safe and effective medications (Vora K.L. et al., 2023).

7.1.4. Contribution of Artificial Intelligence to Dissolution Rate Predictions

The term "dissolution rate" describes how quickly a medicine dissolves in a biological fluid. The drug's bioavailability and therapeutic efficacy are determined by this feature. Because artificial intelligence models can identify important physicochemical properties and molecular characteristics that influence the dissolution process through the analysis of large amounts of experimental data, they have greatly aided in the optimization of drug formulations and dosage forms and helped predict dissolution rates. These models achieve great prediction accuracy by using machine learning algorithms to identify intricate patterns and correlations between drug characteristics and dissolution rates. Artificial intelligence offers valuable insights into the dissolving behavior of various drug formulations. These insights can be utilized to build more efficient drug delivery systems and pick the best formulation techniques for enhanced drug absorption and solubility. Scientists now have useful tools to expedite medication development, improve formulation techniques, and ultimately enhance patient outcomes thanks to artificial intelligence's help for dissolution rate prediction advancements (Mukhamediev R.L. et al., 2022).

CONCLUSION AND DISCUSSION

Technology known as artificial intelligence has been incorporated into pharmaceutical R&D to expedite and lower the cost of the medication development and discovery processes. Owing to the advancement of machine learning theory and the synthesis of pharmacological data, artificial intelligence technology now functions as a potent data mining instrument in several drug design domains, including activity prediction, virtual screening, QSAR analysis, and in silico assessment of absorption, distribution, metabolism, excretion, and toxicity (ADME/T) properties (Çelik İ.N. et al. 2021). It can forecast proteomes, genomes, and patient-specific dosage formulations in addition to enhancing currently available medications. The development of novel compounds with target binding qualities that improve therapeutic efficacy and decrease adverse effects is made possible by systems created in partnership with scientists and artificial intelligence specialists. In order to improve compliance, AI-enabled systems will continuously gather data from wearables, sensors, and remote patient monitoring. They will also use genetic profiles, biomarkers, and electronic health records to identify eligible patients, lower the cost of trials, and expedite approval. However, this presents ethical questions regarding patient consent. It has a number of benefits over conventional experimental techniques, including lower clinical trial attrition rates, fewer animal studies due to less frequent use of in vivo assays, process and expense control, and labor cost savings. Artificial intelligence (AI) is at the core of cutting-edge technologies because it has the unmatched ability to find novel candidate therapies that can be swiftly made available for clinical trials and, if authorized, integrated into healthcare. Accordingly, AI has promise for the creation of new medications and the repurposing of those already in use to treat human diseases, particularly those that are emerging like Coronavirus Disease 2019 (COVID-19) (Zhou Y., et al. 2016). Despite all of these benefits, artificial intelligence is still viewed as a mystery because it cannot be explained. Features are not well defined throughout the training phase, and the network designer might not know what is being looked at in the intermediate steps or why the model has reached a certain conclusion. Because of this, a lot of work has been done to speed up the drug discovery process and integrate AI tools into the system. However, before the full potential of AI in drug discovery and development can be realized, more successful applications of these tools will be needed (Chan H.C.S., et al. 2019).

REFERENCES

1. Adamson A.S., Smith A. Machine learning and health care disparities in dermatology. JAMA Dermatol. 2018;154(11):1247-1248. https://doi.org/10.1001/jamadermatol.2018.2348
2. Andrade E.L., Bento A.F., Cavalli J. et al. Non-clinical studies required for new drug development - Part I: early in silico and in vitro studies, new target discovery and validation, proof of principles and robustness of animal studies, 2016; 49(11): e5644.
3. Ardila D, Kiralyap, Bharadwaj S, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. 2019;25(6):954-961.
4. Askr H., Elgeldawi E., Ella H.A. et al. Deep learning in drug discovery: an integrative review and future challenges, 2023;56(7): 5975–6037.
5. Azuaje F. Artificial intelligence for precision oncology: beyond patient stratification. 2019;3(6):6. https://doi.org/10.1038/s41698-019-0078-1
6. Baskin I.I., Winkler D.D., Tetko I.V., A renaissance of neural networks in drug discovery, 2016;11(8):785-95.
7. Bhhatarai, B.; Walters, W.P.; Hop, C.E.C.A. et al. Opportunities and Challenges Using Artificial Intelligence in ADME/Tox. 2019;18(5):418-422.
8. Bielza C., Larrañaga P. Discrete Bayesian Network Classifiers: A Survey, 2014;47(1):1-43. https://doi.org/10.1145/2576868
9. Bonini CP. Simulation of Information and Decision System in the Firm. Englewood Cliffs, NJ: Prentice Hall; 1963.
10. Breiman L., Random Forests, 2001;45:5-32, http://dx.doi.org/10.1023/A:1010933404324
11. Cano G., Rodriguez J.G., Garcia A.G. et al. Automatic selection of molecular descriptors using random forest: Application to drug discovery, 2017;72:151-159.
12. Chan H.C.S., Shan H., Dahoun T. et al. Advancing Drug Discovery via Artificial Intelligence. 2019;40(8):592-604.
13. Chapelle O, Schölkopf B, Zein A. Semi-supervised Learning. Cambridge, MA: MIT Press; 2006.
14. Chauhan R., Ghanshala K.K., Joshi R.C. Convolutional Neural Network (CNN) for Image Detection and Recognition, 2018. First International Conference on Secure Cyber Computing and Communication (ICSCCC). doi:10.1109/icsccc.2018.8703316
15. Chavda, V.P. Chapter 1 - Nanotherapeutics and

Nanobiotechnology, 2019;1-13. https://doi.org/10.1016/B978-0-12-814029-1.00001-6
16. Chen R., Liu X., Jin S. et al. Machine Learning for Drug-Target Interaction Prediction. 2018;23(9):2208.
17. Choy G, Khalilzadeh O, Michalski M, et al. Current applications an future impact of machine learning in radiology. 2018;288(2):318-328.
18. Chui M, Loeffler M, Roberts R. The Internet of Things. McKinsey Quarterly. March 2010. https://www.mckinsey.com/industries/high-tech/our-insights/the-internet-of-things. Accessed June 21, 2019.
19. Colombo, S. Chapter 4 - Applications of artificial intelligence in drug delivery and pharmaceutical development. 2020;85-116. https://doi.org/10.1016/B978-0-12-818438-7.00004-6
20. Çelik İ.N., Arslan F.K., Tunç R. et al. ARTIFICIAL INTELLIGENCE ON DRUG DISCOVERY AND DEVELOPMENT, 2021;45(2): 400-427.
21. D'Souza S., Prema K.V., Balaji S., Machine learning models for drug–target interactions: current knowledge and future directions, 2020;25(4):748-756.
22. Dana D., Gadhiya S.V., Surin L.G.ST. et al. Deep Learning in Drug Discovery and Medicine; Scratching the Surface. 2018;23(9):2384.
23. Das P.J., Preuss C., Mazumder B. Artificial Neural Network as Helping Tool for Drug Formulation and Drug Administration Strategies. 2016;263-276.
24. Deo R.C. Machine learning in medicine. 2015;132(20):1920-1930. https://doi.org/10.1161/CIRCULATIONAHA.115.001593
25. Dodd F.S., Target-based drug discovery: is something wrong?, 2005;10(2):139-147. https://doi.org/10.1016/S1359-6446(04)03316-1
26. Dreyfus S., The numerical solution of variational problems. 1962;5(1):30-45. https://doi.org/10.1016/0022-247X(62)90004-5
27. Bejnordi B.E., Veta M. et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. 2017;318(22):2199-2210.
28. Ekins S., The Next Era: Deep Learning in Pharmaceutical Research. 2016;33(11):2594-603. DOI: 10.1007/s11095-016-2029-7
29. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. 2017;542(7639):115-118.
30. Fleming N., How artificial intelligence is changing drug discovery. 2018;557(7707):S55-S57. DOI: 10.1038/d41586-018-05267-x
31. Fukushima K., Neocognitron: a hierarchical neural network capable of visual pattern recognition. 1988;1(2):119–130. https://doi.org/10.1016/0893-6080(88)90014-7
32. Galata D.L., Könyves Z., Nagy B. et al., Real-time release testing of dissolution based on surrogate models developed by machine learning algorithms using NIR spectra, compression force and particle size distribution as input data. 2021;597:120338.
33. Ghasemi F., Mehridehnavi A., Garrido A.P. et al. Neural network and deep-learning algorithms used in QSAR studies: merits and drawbacks. 2018;23(10):1784-1790.
34. Ghourichay M.P., Kiaie S.H., Nokhodchi A. Formulation and Quality Control of Orally Disintegrating Tablets (ODTs): Recent Advances and Perspectives. 2021;24:2021:6618934.
35. Gilboa E., Saatçi Y., Cunningham J.P. Scaling Multidimensional Inference for Structured Gaussian Processes. 2015;37(2):424-36.
36. Goel A.K., Davies J. Artificial intelligence. In: Cambridge Handbook of Intelligence Cambridge. 2019.
37. Goodfellow I, Bengio Y, Courville A. Deep Learning. Cambridge, MA: MIT Press; 2016.
38. Gulsen V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. 2016;316(22):2401-2410.
39. Gupta R., Srivastava D., Sahu M. et al. Artificial intelligence to deep learning: machine intelligence approach for drug discovery. 2021;25(3):1315-1360.
40. Gurung A.B., Ali M.A., Lee J. et al. An Updated Review of Computer-Aided Drug Design and Its Application to COVID-19. 2021; 24:2021:8853056.
41. Hamet P, Tremblay J., Artificial intelligence in medicine, 2017;69:S36-S40. https://doi.org/10.1016/j.metabol.2017.01.011
42. Hassanzadeh P., Atyabi F., Dinarvand R. The significance of artificial intelligence in drug delivery system design. 2019;151-159:169-190.
43. Haugeland J., Artificial Intelligence: The Very Idea. Cambridge, MA: MIT Press; 1985.
44. Heikamp K., Bajorath J. Support vector machines for drug discovery. 2014;9(1):93-104. doi: 10.1517/17460441.2014.866943.
45. Hessler G., Baringhaus K.H., Artificial Intelligence in Drug Design, 2018; 23(10):2520. https://doi.org/10.3390/molecules23102520
46. Hinselmann G., Rosenbaum L., Jahn A. et al. Large-Scale Learning of Structure−Activity Relationships Using a Linear Support Vector Machine and Problem-Specific Metrics. 2011;51(2):203-213.
47. Hinton G, Deng L, Yu D, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. 2012;29(6):82-97.
48. Hinton G., Sejnowski T.J., Unsupervised Learning: Foundations of Neural Computation. Cambridge, MA: MIT Press; 1999.
49. Hirschberg J, Manning C.D. Advances in natural language processing. 2015;349(6245):261-266. https://doi.org/10.1126/science.aaa8685
50. Howard J. Artificial intelligence:Implications for the future of work. 2019; 62(11):917-926. https://doi.org/10.1002/ajim23037
51. Hughes J.P., Rees S., Kalindjian S.B. et al. Principles of early drug discovery. 2011;162(6):1239-49.
52. James G, Witten D, Hastic T, Tibshirani R. An Introduction to Statistical Learning. 2013;87.
53. Jeon J., Nim S., Teyra J. et al. A systematic approach to identify novel cancer drug targets using machine learning, inhibitor design and high-throughput screening. 2014; 6(7):57.
54. Jiang J., Ma X., Ouyang D.et al. , Emerging Artificial Intelligence (AI) Technologies Used in the Development of Solid Dosage Forms. 2022;14(11):2257.
55. Jordan A.M., Artificial Intelligence in Drug Design—The Storm Before the Calm? 2018; 9(12):1150-1152. doi: 10.1021/acsmedchemlett.8b00500.
56. Jordan M.I., Mitchell T.M., Machine learning: trends, perspectives, and prospects. 2015;349(6245):255-260. https://doi.org/10.1126/science.aaa8415
57. Jorgensen W.L., The Many Roles of Computation in Drug Discovery. 2004; 303(5665):1813-8. DOI: 10.1126/science.1096361
58. Kanagasingam Y., Xiao D., Vignarajan J. et al. Mehrotra A., Evaluation of artificial intelligence-based grading of diabetic retinopathy in primary care. 2018;1(5):e182665-e182671.
59. Kapetanovic I.M., Computer-aided drug discovery and development (CADDD): In silico-chemico-biological approach. 2008; 171(2):165-76. https://doi.org/10.1016/j.cbi.2006.12.006.
60. Katsila T., Spyroulias G.A., Patrinos G.P. et al. Computational approaches in target identification and drug discovery, 2016; 14: 177–184.
61. Kawato T., Mizohata T., Shimizu Y. et al. Structure-based design of a streptavidin mutant specific for an artificial biotin analogue. 2015;157(6):467-475.
62. Kehoe B, Patil S, Abbeel P. et al. A survey of research on cloud robotics and automation. 2015;12(2):398-409.
63. Kelley H.J., Gradient Theory of Optimal Flight Paths, 1960;30(10):947-954. https://doi.org/10.2514/8.5282
64. Kim S.B., Han K.S., Rim H.C. et al. Some Effective Techniques for Naive Bayes Text Classification. 2006; 18(11):1457-1466.
65. Knight W. Reinforcement learning: ten breakthrough technologies 2017. MIT Technol Rev. 2017. https://www.technologyreview.com/s/603501/10-breakthrough-technologies-2017-reinforcement-learning/.
66. Kononenko I. Machine learning for medical diagnosis: history, state of the art and perspective. 2001;23(1):89-109. doi: 10.1016/

s0933-3657(01)00077-x.
67. Korotcov A., Tkachenko V., Russo D.P. et al. Comparison of Deep Learning With Multiple Machine Learning Methods and Metrics Using Diverse Drug Discovery Data Sets. 2017;14(12): 4462–4475.
68. Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional networks. Proceedings of Advances in Neural Information Processing Systems. 2012;25:1090-1098.
69. Lakhani P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. 2017;284(2):574-582. doi: 10.1148/radiol.2017162326.
70. Lavecchia A., Deep learning in drug discovery: opportunities, challenges and future prospects, 2019;24(10): 2017-2032, https://doi.org/10.1016/j.drudis.2019.07.006
71. Lee J., Bogyo M., Target deconvolution techniques in modern phenotypic profiling, 2013;17(1):118-126, https://doi.org/10.1016/j.cbpa.2012.12.022
72. Lee I., Keum J., Nam H., DeepConv-DTI: Prediction of drug-target interactions via deep learning with convolution on protein sequences. 2019; 15(6): e1007129.
73. Lengauer T., Rarey M., Computational methods for biomolecular docking, 1996;6(3):402-406.https://doi.org/10.1016/S0959-440X(96)80061-3
74. Li X., Xu Y., Cui H. et al. Prediction of synergistic anti-cancer drug combinations based on drug target network and drug induced gene expression profiles. 2017;83:35-43.
75. Lipinski C.F., Maltarollo V.G., Oliveira P.R. et al. Advances and Perspectives in Applying Deep Learning for Drug Design and Discovery. 2019;5:6:108.
76. Lv B.M., Quan Y., Zhang H.Y. Causal Inference in Microbiome Medicine: Principles and Applications. 2021;29(8):736-746.
77. Mayr A., Klambauer G., Unterthiner T. et al. DeepTox: Toxicity Prediction using Deep Learning. Fronti-ers. 2016.
78. McFedries A., Schwaid A., Saghatelian A. Methods for the Elucidation of Protein-Small Molecule Interactions. 2013; 20(5):667-73.
79. Mencattini A., Giuseppe D.D., Comes M.C. et al. Discovering the hidden messages within cell trajectories using a deep learning approach for in vitro evaluation of cancer drug treatments. 2020; 10(1):7653.
80. Mohamed C., Nsiri B., Abdelmajid S. et al. Deep Convolutional Networks for Image Segmentation: Application to Optic Disc detection. FLORIDA INTERNATIONAL UNIVERSITY, July 24,2020.
81. Muegge I., Bergner A., Kriegl J.M., Computer-aided drug design at Boehringer Ingelheim, 2017;31(3): 275-285.
82. Mukhamediev R.I., Popova Y., Kuchin Y. et al. Review of Artificial Intelligence and Machine Learning Technologies: Classification, Restrictions, Opportunities and Challenges. 2022;10(15):2552.
83. Nag A, Mukhopadhyay SC, Kosel J. Wearable flexible sensors: a review. 2017;17(13):3949-3960.
84. Navya K., Kamaraj R., Bharathi M., The Trending Role of Artificial Intelligence and Its Applications in Formulation of Solid Dosage Forms: A Review. 2022;107(1):20049.
85. Nicolaou C.A., Brown N. Multi-objective optimization methods in drug design. 2013;10(3):e427-e435, https://doi.org/10.1016/j.ddtec.2013.02.001
86. Notodiputro K.A., Sartono B., Zubedi F. Implementation of Winsorizing and random oversampling on data containing outliers and unbalanced data with the random forest classification method. 2022;22(2),108-116.
87. Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. 2016;375(13):1216-1219. https://doi.org/10.1056/NEJMp1606181
88. Öztürk H., Özgür A., Ozkirimli E. DeepDTA: deep drug–target binding affinity prediction. 2018; 34(17):i821-i829.
89. Park H., Chien P.N., Ryu S.E. Discovery of potent inhibitors of receptor protein tyrosine phosphatase sigma through the structure-based virtual screening. 2012;22(20):6333-6337.
90. Patel L., Shukla T., Huang X. et al. Machine Learning Methods in Drug Discovery. 2020; 25(22): 5277.
91. Paul D., Sanap G., Shenoy S. et al. Artificial intelligence in drug discovery and development. 2021;26(1):80-93.
92. Pervan G., Willcocks L.,. Introduction to the special issue on decision support systems. 2005;20:65-66.
93. Phillips-Wren G. AI tools in decision support systems: a review. 2012;21(2):1240005.
94. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. 2018;2(3):158-164.
95. Pratt G.A. Is a Cambrian explosion coming for robotics? 2015;29(3):51-60. https://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.29.3.51.
96. Pun F.W., Ozerov I.V., Zhavoronkov A., AI-powered therapeutic target discovery. 2023; 44(9):561-572.
97. Qureshi R., Zou B., Alam T. et al. Computational Methods for the Analysis and Prediction of EGFR-Mutated Lung Cancer Drug Resistance: Recent Advances in Drug Design, Challenges and Future Prospects. 2023; 20(1):238-255.
98. Ragoza M., Hochuli J., Idrobo E. et al. Protein–Ligand Scoring with Convolutional Neural Networks. 2017; 57(4):942-957.
99. Rahman R., Dhruba S.R., Ghosh S., Pal R. Functional random forest with applications in dose-response predictions, Scientific Reports. 2019;9 (1):1628.
100. Ratanamahatana C., Gunopulos D. Feature selection for the naive bayesian classifier using decision trees. 2003;17:475-487.
101. Rathi P.C., Ludlow R.F., Verdonk M.L. Practical High-Quality Electrostatic Potential Surfaces for Drug Discovery Using a Graph-Convolutional Deep Neural Network. 2020;63(16):8778–8790,
102. Reddy A.S., Zhang S. Polypharmacology: drug discovery for the future. 2013;6(1):10. doi: 10.1586/ecp.12.74
103. Reher R., Kim H.W., Zhang C. et al. , A Convolutional Neural Network-Based Approach for the Rapid Annotation of Molecularly Diverse Natural Products. 2020;142(9):4114-4120.
104. Rifaioglu A.S., Atas H., Martin M.J. et al. , Recent applications of deep learning and machine intelligence on in silico drug discovery: methods, tools and databases. 2019;20(5):1878-1912.
105. Rosenblatt F (1957) The Perceptron: A Perceiving and Recognizing Automaton, Report 85–60–1.
106. Russel S.J., Norvig P., Artificial Intelligence A Modern Approach Third Edition, 2003.
107. Sarica A., Ceresa A., Quattrone A. Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer's Disease: A Systematic Review. 2017; 6:9:329.
108. Sarkar C., Das B., Rawat V.S. et al. Artificial Intelligence and Machine Learning Technology Driven Modern Drug Discovery and Development. 2023; 24(3):2026.
109. Schenone M., Dančík V., Wagner B.K. et al. Target identification mechanism of action in chemical biology and drug discovery. 2013; 9(4):232-40.
110. Siepmann, J., Siepmann, F. Modeling of Diffusion Controlled Drug Delivery. 2012;161(2):351–362. https://doi.org/10.1016/j.jconrel.2011.10.006
111. Sliwoski G., Kothiwale S., Meiler J. et al. Computational Methods in Drug Discovery. 2014;66(1):334-395.
112. Sun H. A Naive Bayes Classifier for Prediction of Multidrug Resistance Reversal Activity on the Basis of Atom Typing. 2005;48(12):4031–4039. https://doi.org/10.1021/jm050180t
113. Sutton R.S., Barto A.G., Reinforcement Learning: An Introduction. 2nd ed. Cambridge, MA: MIT Press; 2018.
114. Swinney D.C. Phenotypic vs. Target-Based Drug Discovery for First-in-Class Medicines. 2013; 93(4):299-301. https://doi.org/10.1038/clpt.2012.236
115. Swinney D.C., Anthony J. How were new medicines discovered? 2011;10(7):507–519. DOI: 10.1038/nrd3480
116. Tajbakhsh N., Shin J.Y., Gurudu S.R. et al. Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? 2016;35(5):1299-1312.

117. Tan M. Prediction of anticancer drug response by kernelized multitask learning. Artif Intell Med. 2016;73:70-77. https://doi.org/10.1016/j.artmed.2016.09.004
118. Terstappen G.C., Reggiani A. In silico research in drug discovery. 2001;22(1):23-26. DOI:https://doi.org/10.1016/S0165-6147(00)01584-4
119. Terstappen G.C., Schlüpen C., Raggiaschi R., Gaviraghi G. Target deconvolution strategies in drug discovery. 2007;6(11):891–903. DOI: 10.1038/nrd2410
120. Thorndike E. The Fundamentals of Learning. New York, NY: Teachers College Bureau of Publications; 1932.
121. Topol E. Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again. New York, NY: Basic Books; 2019.
122. Vamathevan J., Clark D., Czodrowski P. et al. Applications of machine learning in drug discovery and development. 2019;18(6):463-477.
123. Varian H. Artificial Intelligence, Economics, and Industrial Organizations. In: Agrawal A, Gans J, Goldfarb A, eds. The Economics of Artificial Intelligence: An Agenda. Chicago, IL: University of Chicago Press; 2019:399-422.
124. Vemula D., Jayasurya P., Sushmitha V. et al. CADD, AI and ML in drug discovery: A comprehensive review. 2023:181:106324.
125. Vora K.L., Gholap A.D., Jetha K. et al. Artificial Intelligence in Pharmaceutical Technology and Drug Delivery Design. 2023;15(7), 1916.
126. Wang S., Sim T.B., Kim Y.S. et al. Tools for target identification and validation. 2004; 8(4):371-7.
127. Wann F., Zeng J.M., 2016, Deep learning with feature embedding for compound-protein interaction prediction, doi: https://doi.org/10.1101/086033
128. Wann F., Zeng J.M. Deep learning with feature embedding for compound-protein interaction prediction. 2016. doi: https://doi.org/10.1101/086033
129. Waring M.J., Arrowsmith J., Leach A.R. et al. An analysis of the attrition of drug candidates from four major pharmaceutical companies. Nature Reviews Drug Discovery,2015;14(7):475–486.
130. Wassermann A.M., Geppert H., Bajorath J. Application of Support Vector Machine-Based Ranking Strategies to Search for Target-Selective Compounds. Methods Mol Biol. 2011:672:517-30.
131. West DM, Allen JR. How Artificial Intelligence is Transforming the World. Access date: April 24, 2018. Access Adress: https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/.
132. Yamashita R., Nishio M., Do R.K.G. et al. Convolutional neural networks: an overview and application in radiology. 2018; 9(4):611-629.
133. Yang X., Wang Y., Byrne R. et al. Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery. 2019;119(18): 10520–10594.
134. Yu W., MacKerell A.D., Computer-Aided Drug Design Methods. 2017;1520:85-106. doi: 10.1007/978-1-4939-6634-9_5
135. Zang Q., Mansouri K., Williams A.J. et al. , In Silico Prediction of Physicochemical Properties of Environmental Chemicals Using Molecular Fingerprints and Machine Learning. 2017;57(1):36-49.
136. Zhang L., Tan J., Dan H. et al. From machine learning to deep learning: progress in machine intelligence for rational drug discovery. 2017;22(11):1680-1685.
137. Zhou, Y., Wang, F., Tang, J. et al. Artificial intelligence in COVID19 drug repurposing. The Lancet Digital Health, 2020;2(12): e667–e676.