ARTIFICIAL INTELLIGENCE IN DRUG DISCOVERY

Project Report submitted in a partial fulfillment

For the award of the degree of

BACHELOR OF PHARMACY

Submitted by

KUNWAR ASIF

17SMAS102044

IN

BRANCH OF STUDY

SCHOOL OF MEDICAL AND ALLIED SCIENCE

Under the supervision of

SONALI SUNDRAM

Greater Noida

March,21

SCHOOL OF MEDICAL AND APPLIED SCIENCE

BONAFIDE CERTIFICATE

Certified that this project report "**ARTIFICIAL INTELLIGENCE IN DRUG "** Is the Bonafide work of "**KUNWAR ASIF"** who is carried out the project work under my Supervision.

SIGNATURE.
Prof. Parmod kumar sharma
Dean of school
School of Medical and Allied Science.

SIGNATURE
SONALI SUNDRAM
SUPERVISOR
Assistant professor
School of Medical and Allied Science

ABSTRACT

Good amount of knowledge leads to promise for drug discovery, additional with high level image analysis, different types of molecular structure and function, generation of chemical entities with number of different properties. Achieving the number of successful application results, the complex mathematical models are always exclusive for the human brain. That why there is a need for questionable deep learning methods to address the demand for exclusive machine languages for the study of structural science. This review summaries the vary methods and concept how artificial intelligence works and the future opportunities, working application as well as multiple challenges faced. As we know the most difficult task in the pharmaceutical companies is to find out the new drug molecules, lot of money and time used but at the end result are not satisfying hundreds of trails are held on a particular project but the success rate is low to tackle with these types of problems now Pharmaceutical companies approaches toward the artificial intelligence. They joined hands with different world leading software companies and by working together as a team they form a deep learning program which helps in the findings the new drug molecules .In this machine learning process they put the data of every molecules and there mix up reaction when there is need to trail for drug now it is been used and it is reliable and efficient and providing promising results even at faster rate and with the help of that the success rate is to 60 to 70 percent. This review provide the various machine learning concept and in which disease they are used to find there treatment and medicines. Companies like Microsoft have made there artificial intelligence software and use it for researching different drugs for a particular disease recently research for SARS-CoV-2 in this time of pandemic where there is a need for fast searching of vaccine AI helps in many ways .

TABLE OF CONTENT

LIST OF TABLES

TABLE NO. **TITLE.** **PAGE NO.**

LIST OF FIGURES

FIGURE NO. **TITLE.** **PAGE NO**

CHAPTER 1

INTRODUCTION

Artificial intelligence (AI) has recently been developed into a hot topic in area of medicine industry. The Pharmaceutical industries are putting efforts to approach AI to enhance drug discovery process, reduce research and development budget, low down failure rates in clinical trials and ult-imately produce reliable medicines. Computational methods play a key role in the design of thera-petically important molecules for modern drug development. It could be interpreted in many ways. Traditionally, computers have not been creative, that is, they can only do what humans tell them to do . Perhaps, by saying “they only give answers “. Modern artificial intelligence (AI) has the ability to significantly enhance the role of computers and computational methods in science and engineering. The World Economic Forum refers to a combination of big data and AI as both the fourth paradigm of science and the fourth industrial revolution.

1.1 What is artificial intelligence?

It is intelligence demonstrated by machines, unlike the natural intelligence displayed by humans and animals, which involves consciousness and emotionality. Artificial intelligence was founded as an academic discipline in 1955 and in the years since has experienced several waves of optimism. The productivity of the pharmaceutical industry is on the decline. Failure rates of clinical trials exceeds 85% after therapies are tested in model organisms, the cost to develop a new drug exceed $2.5 billion. Recent advances in artificial intelligence (AI) may help to reverse this trend and accelerate and improve pharmaceutical R&D. While the term AI and the concept of deep learning are not knew, recent advances in high performance computing, availability of large annotated datasets required for training and new framework For implementing deep neural network resulted in an unprecedented acceleration of the field.

1.1.1. Machine learning

Machine learning is the study of computer algorithms that improves automatically through experience and by the use of data. It seen as a part of artificial intelligence. Machine learning algorithms build a model based on sample data, known as training data,in order to make decisions without being explicitly programmed to do so. A subset of machine learning is closely related to computational statistics, which focus on making prediction using computers; but not at all machine learning is statistical Learning.

1.2. Deep Learning

It is an artificial intelligence function that imitates the working of the human brain in processing data and creating patterns for use in decision making .Deep learning is a subset of machine learning in artificial intelligence that has network available of understanding unsupervised from data that is very complicated and hard to understand. Also known as a deep neural learning or deep neural network. Deep learning is an AI function that mimics the working of the human brain in processing data for use in detecting objects, recognizing , speech, translating languages and making decisions. It helps in learn without human supervision, drawings from data that is both unstructured and unlabelled evolved day by day with the digital era, which has brought about an explosion of data in all forms and from every country in the world .

Figure 1.1 Depiction of AI in pharma Drug discovery

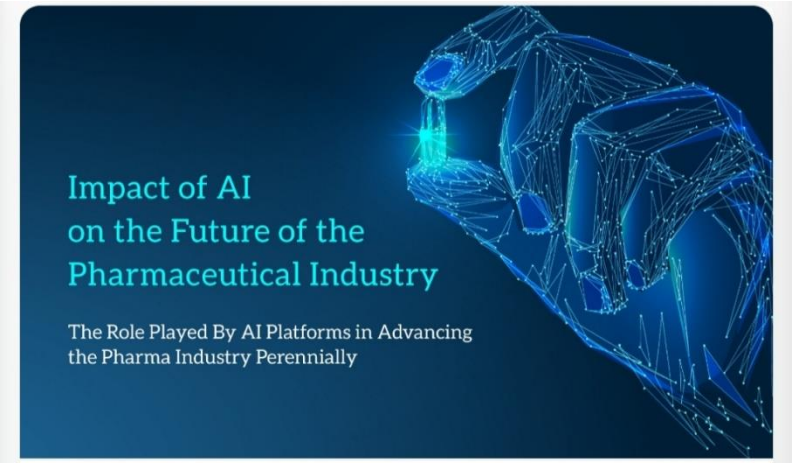


Figure 1.2 How AI can change the future depiction

1.2.1. ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Difference between the two is shown in the fig.1.2 . AI as a computer application process on input which humans runs on it to give same results for multiple runs. It is a software which have a tendency to behave like humans [1]. It has a other name like computer learning or we can call it machine learning ,which uses different types of method with having ability to self learn with or without being programmed.2] .

Machine learning have subgroups supervised, unsupervised and reinforcement learning. Supervised learning consist classification and regression model in which the data from the input and output sources are provided to give the resulting model .Output from this type of machine learning finds disease cure and how we can tackle it under the subgroup classification; and drug efficacy under the sub group regression. [3].

Clustering and feature finding methods by grouping and analyzing data based solely on input data are used in unsupervised learning. Disease Type uncovering from bunching and disease goal unearthing from attributes finding methods can both be done using unsupervised machine learning outputs[4]. Reinforcement learning is mainly powered by decision-making in a given conditions and

it's implementation to escalate its performance. DE novo drug design under decision-making and practical designs under enactment are two examples of outputs from this form of machine learning, both of which can be masterly using simulation and quantum chemistry[3].

Medicinal chemistry has used AI to design compounds in different forms and with varying degrees of success since the 1960s. It is widely used supervised learning, in which labelled training datasets are used to train models. An example is the quantitative structure–activity relationship (QSAR) approach, which is widely used to predict properties, such as log p, solubility and bioactivity, for given chemical structures[5]. Unsupervised learning, on the other hand, is common in medicinal chemistry, with techniques like hierarchical clustering, algorithms, and principal components analysis being used to analyze and break down large molecular libraries into smaller groups of related compounds.

In 2014 and 2015, early presentations to the pharmaceutical industry on deep learning developments were met with scepticism and were discarded[6][7]. Many pharmaceutical companies began collaborating with AI startups and academics in 2017 or developed internal R&D programmes . Deep learning techniques quickly spread into many areas of biomedical science, from training DNNs on transcriptional response data to predicting the pharmacological properties of small molecules and biomarker production to the generation of novel chemistry[8].

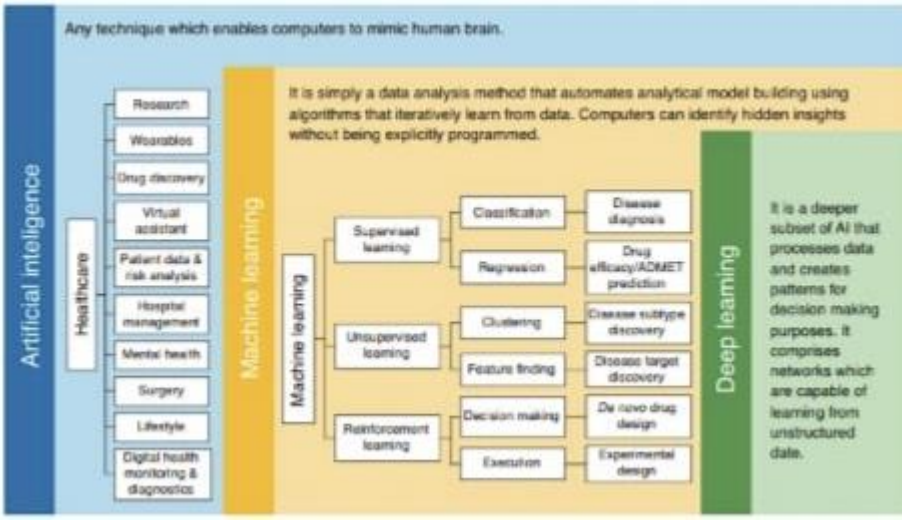


Figure 1.3 Shows difference between AI, ML and DL

1.2.2 WHAT To BE EXPECTED (AI and ML)

The ultimate goals of applying AI and ML to drug development challenges are the same as they have always been to bring the best medicines to the clinic to fulfil unmet patient needs. This includes activities such as identifying drug targets, identifying lead compounds, optimising their designs against several property profiles of interest, and identifying synthetic routes to realize the composition of matter in the case of drug discovery and medicinal chemistry.

AI is often portrayed as a magical button that can be pressed at any time to achieve the ideal output, regardless of the input. If the AI challenge is to produce the ideal picture of a cat from a model trained on cat photos, a vehicle that can drive itself without making a single error, or a medicine that can safely and effectively treat a disease.. While AI is not the answer to every challenge, it is a useful tool that if we use it correctly it can give us enormous and unimaginable results. The best AI in medicinal chemistry and drug discovery isn't necessarily a single AI that can design a new drug on its own, but one or more AIs that allow better understanding and design of new inputs at every stage of the drug discovery process, from target selection to hit detection to lead optimization to preclinical studies and clinical trials .

1.3 DRUG PROCESSING THROUGH AI

The feedback-driven drug development process begins with existing results obtained from a variety of sources, including high-throughput compound and fragment screening, computational modelling, and literature knowledge. Induction and deduction alternate in this phase[9].

This inducible-deducible loop ultimately open on to hit and lead compounds that are optimized. The artificial intelligence of particular segment of the cycle eliminates uncertainty and mistakes while increasing drug production efficiency .

For virtual screening models and In silico compounds that serve as replacement for biological effectiveness, poisonous and biochemical studies DE novo design methods necessitate understanding of chemistry . Finally, operative learning algorithms allow the verification of new or novel compounds with promising anti-disease activity.

The discovery of novel chemical compounds with biological activity is the first step in drug production. The association of the chemical compound with a particular enzyme or having an entire organism will result in biological activity[11]. A ‘hit' is the new compound that displays action in opposition to a specific biological target. Hits are often discovered when screening chemical repositories, computer simulations, or naturally pure materials like fungi, bacteria and plants[10]. The second stage in drug production is to recognize a lead molecule. If a lead compound has been recognized, its structure is used to guide chemical adaptation with the goal of finding compounds with the greatest therapeutic profit and the least possibilities for damage. A lead is a chemical substance that has the ability to guide to the production of a new drug for a disease treatment. To characterize the effectiveness of the compound and its likely safety profile, recognize hits are televise in cell based assays predictive of disease state and in animal models of disease.Hit molecules are routinely modified during the lead generation process to enhance their behavior and selectivity against particular Targets biological while lowering down the poisonous level and undesirable affect. Analogues are similar compounds extracted from a hit, and the procedure is known as hit expansion. Medicinal producers use well-established organic chemistry techniques to extend hits[12]. Chemists rely on a single reaction or collection of reactions to rapidly collect building blocks to make a sequence of analogues to improve synthetic throughput.

A 'building block' is a compound with a reactive functional group and atoms that interact with a biological target's active site. The active site in a biological target is a particular region to which the compound (or substrate) binds through interaction forces. Models depicting the binding of a substrate to an active site are known as "lock and key" or "induced-fit" models[12].

CHAPTER 2

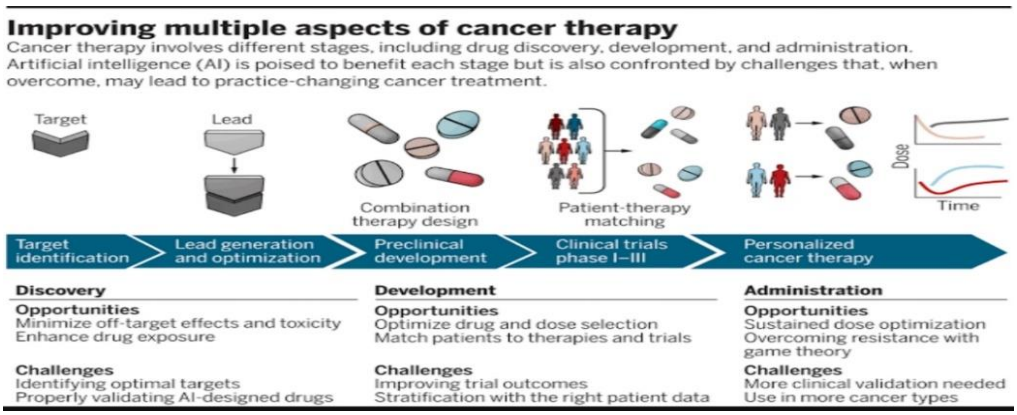
2.1 AI IN CANCER TREATMENT

Despite the fact that more than half of cancer patients today live for more than ten years, many forms of cancer have poor results and are poorly served by drugs (www.icr.ac.uk/about-us/our-cancerdrug-access-report).

Furthermore, oncology physicians and drug developers must now deal with cancer's potential to adapt, grow, and become immune to therapy. Meanwhile, the Big Data revolution is having an increasingly large impact across various areas, impacting many facets of our everyday lives as well as a wide variety of research fields. This is due to the improved accessibility and lower costs of technology for generating, storing, and analyzing large and complex data sets. Some aspects of multidisciplinary drug discovery, especially medicinal chemistry, have long embraced computational methods and the processing and analysis of Big Data to aid decision-making. The use of large-scale data mining to find target hypotheses is now commonplace. Furthermore, the US FDA's launch of programmes in oncology aimed at "going beyond the reductionist approach to drug production" indicates the potential direction of travel (A holistic approach (for example, combination therapies targeting complex multiunit signatures and real-world evidence) is preferable to a single drug targeting a driver mutation and standard clinical trials.'. Despite advancement, the drug discovery community is still a long way from understanding the full potential of Big Data analytics and Artificial Intelligence (AI) [15].

The major public databases and tools that provide access to Big Data that can guide modern cancer drug research are reviewed in this article. We look at how combining Big Data from different information domains can help with decision-making in the drug discovery process. We address how AI and Machine Learning (ML) are transforming the way we discover new drugs, highlighting areas where increased use of Big Data analytics will greatly support drug discovery. We mainly concentrate

on cancer research, specifically small-molecule drug development. We do assume, however, that many of the messages are applicable[13][14].**Above figure 1.4 shows the cancer treatment through AI**



2.2 TREATMENT OF DIABETES THROUGH AI

Deep learning-based AI grading of DR from retinal photographs now has a sensitivity and specific[16]. (FDA) recently approved the first medical device to use artificial intelligence (AI) to screen diabetic patients for retinopathy. The computer, known as IDx-DR (IDx LLC, Coralville, IA), is a software programmed that analyses images of the eye taken with a Topcon NW400 retinal camera using an AI algorithm (Topcon Medical Systems, Inc., Oakland, NJ). Digital photographs of the patient's retinas are transferred to a cloud server where the IDx-DR software is installed [17][18]. If the photos are of good enough quality, the programme gives the doctor one of two options: (1) ‘‘Detection of more than mild diabetic retinopathy: consult an eye doctor" or (2) ‘‘negative for more than mild diabetic retinopathy; rescreen in 12 months". IDx-DR is the first device authorized for marketing that provides a screening decision without the need for a clinician to also interpret the image or results[21].. These automated systems enable non-eye health professionals in primary care physician offices to conduct on-site retinal screening and provide on-the-spot normal results or direct referrals to an eye specialist without the need for eye specialists, resulting in substantially higher patient satisfaction[19][20].

Today, AI-driven predictive modelling proactively identifies diabetes populations at high risk of avoidable complications, such as excessive ER visits, admissions, and readmissions. Larger

physician associations, health care systems, and health plans use AI to proactively classify and define diabetes populations, locate patients at risk for diabetic comorbidities, identify patients for special diabetes disease management services, and discover important proteins and genes correlated with and predictive of diabetic[22][23].

Today, AI provides doctors and other health practitioners caring for people with disabilities with practice decision-support resources. Machine learning techniques assist doctors in tailoring diabetes treatments to improve adherence and health outcomes. AI-enabled technologies assist clinicians in diagnosing diabetes noninvasively and assessing and tracking the intensity of diabetic neuropathy

Category	*Number of articles*	*Most common clinical AI applications*
Automated Retinal Screening	96	Detection of diabetic retinopathy, maculopathy, exudates, and other abnormalities from normal findings
Clinical Decision Support	126	Detection and monitoring of diabetes and comorbidities such as neuropathy, nephropathy and wounds
Predictive Population Risk Stratification	135	Identification of diabetes subpopulations at higher risk for complications, hospitalization, and readmissions
Patient Self-Management Tools	94	AI-improved glucose sensors, artificial pancreas, activity and dietary tracking devices
TOTAL	450	

and diabetic wounds more accurately[24][25].

Figure 1.4 AI application used in Diabetes

2.3 TRIALS SELECTION THROUGH ARTIFICIAL INTELLIGENCE

In the processing of a population for clinical trials, artificial intelligence is used. An absolute AI tool for clinical trials will recognize patients' disease, identify DNA Target and forecast the impact of the molecule engineered, including on- and off-target effects. In a Phase II study of schizophrenia patients, a novel AI platform called AI cure was created as a software application to assess medication constancy , with results showing that it improved adherence by 25% when compared to the conventional ‘modified explicitly observed therapy '[26][27] . The method of selecting patients for a clinical trial is critical. Examining the relationship between human-relevant biomarkers and in vitro phenotypes allows for a more predictable and quantifiable evaluation of therapeutic response uncertainty in a particular patient. The advancement of artificial intelligence (AI) methods to classify and predict human-relevant disease biomarkers allows for the recruitment of a particular patient

group in Phase II and III clinical trials[28]. The use of AI predictive modelling in the selection of a patient group would boost clinical trial success rates.

2.4 PRECISE MEDICINE DISCOVERY THROUGH AI

The introduction of large-scale, high-dimensional data, which includes scientific literature, omics data (genomic, metabolomics, proteomic, and exposomic, to name a few), and other sources. with the help of physiological and behavioural data gathered by monitoring Wearable technology has paved the way for the simultaneous use of multiple devices. Many new molecules discovery and personalised medicine are provided with different specific functions. By offering information about a disease's genetic architecture in individual patients, as well as the processes that underpin Pathogenesis of disease, the fundamental concept of drug development, and the focus of growth is moving from symptom relief to more long-term goals. If we can low down the R&D cost on clinical trials with the increase in success rates to find out the potential drug for the simple and complicated disease with validated identification. The person's who read science is forwarded to a past overview of the latest big data try to make accurately precise medicine and their coming utility in new molecule Discovery and their different personalized usage. In the time, making accurately precise medicine, a rigorous and systematic findings of these big , multi mixing , and constantly modified tools is mostly important to find out potential associations and use "the collection of data . "However, Clarification of vague "non-genomic" (exposomic and phenotypic) data, as well as the construction of many ways to derive the inner molecular structure, are both difficult tasks in the curation workflow. In the learnings of basic biomedical data and about their inner molecular structure , various machine learning programme or artificial intelligence system , especially deep learning methods, have a lot of possible ways in advancing the combination of different drug discovery and accurately precise medicine forward . To fully realize AI's potential in tasking the making of new molecules, as well as to make it easier to adapt the therapeutic agent chosen for a particular patient to optimize benefit, analysis and a shared concept of data sharing are needed, with molecule-related data made possible in a structured format on an open website. Therefore, managing large number of data, especially those present in not having structure data , becoming a significant challenge from a teaching, studying and doing prospective towards findings, despite significant efforts to encourage the advancement of information sharing technology and whatever important for searching.

2.5 AI ASSISTED MOLECULAR STRUCTURE

As a descriptive process, computational chemistry is a helpful tool for illuminating structural and biochemical aspects of relevance to the modeller in a variety of contexts.. In recent years, molecular dynamics, in combination with multistage atomic physics/particle physics approaches for observing the particle level time progression of bimolecular structures, has received a lot of attention. [29][30] Techniques, however, are also too computationally intensive for systematic analyses of broader subsets of the chemicals under review in current research. In theory, the combination of AI and computational chemistry gives a far more efficient solutions.[31][32][33] The BehlerParinello symmetry function is a well-known example of the development of neural network potential for large dimensional structures of thousands of atoms. Functional evolution of machine-learned density[31] 34) [34] [34] [35] Schrodinger generalised solutions, [36] molecular characteristics predictions of excited electronic states, [37] The first of them [37] [38] chemical trajectory data classification,830 multi-body expansions[32] and molecular dynamics acceleration,[39][40] part of different gap prediction, and other applications have recently been investigated.[41]. In the field of pharmaceutical design, applications using AI and computational chemistry confront several hurdles due to the significant complexities of bio-structures and the large number of atoms in thousand-degree biomacromolecules.

Table 1 | Computational approaches towards explainable AI in drug discovery and related disciplines, categorized according to the respective methodological concept

Family	Aim	Methods	Reported applications in drug discovery
Feature attribution	Determine local feature importance towards a prediction	• Gradient based • Surrogate models • Perturbation based	Ligand pharmacophore identification, structural alerts for adverse effect, protein–ligand interaction profiling
Instance based	Compute a subset of features that need to be present or absent to guarantee or change a prediction	• Anchors • Counterfactual instances • Contrastive explanations	Not reported
Graph convolution based	Interpret models within the message-passing framework	• Subgraph approaches • Attention based	Retrosynthesis elucidation, toxicophore and pharmacophore identification, ADMET, reactivity prediction
Self-explaining	Develop models that are explainable by design	• Prototype based • Self-explaining neural networks • Concept learning • Natural language explanations	Not reported
Uncertainty estimation	Quantify the reliability of a prediction	• Ensemble based • Probabilistic • Other approaches	Reaction prediction, active learning, molecular activity prediction

For each family of approaches, a brief description of its aim is provided, along with specific methods and reported applications in drug discovery. 'Not reported' refers to families of methods that, to the best of our knowledge, have not been yet applied in drug discovery. Potential applications of these are discussed in the main text. *ADMET: absorption, distribution, metabolism, excretion and toxicity.

In order to avoid the direct measurements of energy by means of quantum mechanics or molecular mechanics techniques, identifying acceptable local chemistry descriptors informatively encoding atomic circumstances to provide adequately predictive results for out-of-sample molecules.. Another source of concern is the planning of reference records, which must be precise and reliable, ensuring sufficiently low intrinsic errors. For tiny organic compounds (for example, tiny pharmaceutical compounds), the first principles-based approaches [for example CCSD-(T)] are thought to suit these needs; however, due to the computational cost, These approaches cannot be used directly to macromolecular targets (for example, proteins); here best practices remain an open-ended question. A promised approach was to include reweighting correction to anticipate the results at a requested stage of theory with high accuracy (e.g. quantum chemicals), based on results achieved at a low-cost base theory level.. (e.g., semi empirical particle chemistry), Verified for thermochemical characteristics of molecules[42]and more recently in free energy variation estimates for chemical reactions.[43]. The needed extrapolation Theory levels are less obvious in biological systems, and the minimal overlap in sample areas at different theory levels contributes to the error. In addition, though scientists have been keen to quantify the ambiguity of functional choices in protein link interactions energy estimates for a long time, there is still uncertainty surrounding the proper depiction of complicated interactions between ligands and targets (e.g. van des Waals forces). Despite these challenges, AI-inspired quantum mechanics/molecular mechanics would be most frequently employed to speed up the discovery of chemical space and the detection of new pharmaceutical candidates by several orders of magnitude, while assuring near mechanical accuracy through constant progression in computer chemistry and the development of AI algorithms. This intimate connection makes AI technology an authentic standard platform for future pharmaceutical applications.

2.6 DIFFERENT MOLECULAR REPRESENTATION THROUGH AI

Approaches to machine learning may leverage chemical signatures, numbers, ASCII sequence and molecules as inputs in drug conception. The molecular attributes are encoded as a sequence of binary bits in the molecular fingerprints (“1” means that the molecular attribute occurs, and “0”

indicates that it does not). Since it is an easy and accurate way of representing molecules, molecular fingerprints are often used in the parts of different designing of molecules to forecast structural properties and measure molecular similarity.. The inputs to neural networks are presently employed for structure-based, molecular, 2D fingerprinting systems, such as the Molecular ACCess System(54), the Extension Connectivity Fingerprint (ECFP) (56), the Functional Class Fingerprint (FCFP), and the Molprint2D (57). MACCS has been used to develop an AAE model, for example, to locate compounds to combat cancer (58).

Chemists have utilised 2D molecular diagrams for long periods to depict molecule structures and analyse molecular attributes qualitatively. Surprisingly, the progress of AI has made it possible to quantify this mechanism. CNN is a flexible tool for automated molecular characteristically extraction, which is suitable in bioactivity molecular representation (59), physiochemical characteristics (55), toxicity (60), and the prediction of protein and ligand affinity (Jimenezet al., 2018). Graphs are more adaptable than the ECFP because the design of the graph may be adjusted based on the tasks. Furthermore, neural networks can be used for the graph convolution architection to mimic molecular properties so that all molecular characteristics may be prepared, extracts and models can be built in simultaneously. The Duvenaud graph convolutionary fingerprints on the basis of atomic irradiation (55), Kearnes' atom-based graph convolutionary fingerprints (61), and Coley's molecular graph CNN fingerprint graph convolutionary fingerprints (55) are included. Duvenaud's core theology of graph convolutionary fingerprints is close to the ECFP, and all of them are finally enlarged by atomic radiation procedures to molecular substructure. Duvenaud et al. specifically first encoded atomic characteristics into vectors and then utilised the vectors for atomic and bonding functions to build the original molecular feature vectors. Duvenaud et al. CNN may be used to obtain characteristics from the initial function sectors above at every iteration, and values are then summarised as molecular fingerprints.. The basic atomic and bond characteristics are expertly created, rather than learning from the molecular graph. Duvenaud's graph CNN benefits from the ability to generate molecular fingerprints that are suitable for a certain function and may be interpreted as molecular fragments connected with certain molecular characteristics may be traced back by neural network nodes. The plot CNN is presented in the DeepChem Werkzeugbox, which indicates that it may learn valuable chemical characteristics and occasionally outperforms other versions by finding out from molecules benchmarks. (64). (65). In addition to the CNN, recurrent neural networks may be utilised to represent molecules. For example, for molecular graph representation, Gregor Urban et al. built

internal and external recursive neural networkes (64). This strategy generally delivers superior prevision than Kearnes' technique for public data sets from molecular net benchmark research (63). Small molecules are represented by the Wiswess line formula (WLN) (65), by the SYBYL line notation (SLN)(66), by SMILES (67) and the International Chemical Identifier (InChI) (68). SMILES has various programmes (for example ChemDraw, Cheopy, and RDKit) and data stores which support SMILES (e.g., PubChem and ZINC). The SMILES (62) coding syntax may be used for studying RNN,, which can then be transformed into a molecular graph. SMILES may also be used directly to determine chemical characteristics as an input RNN function (Goh et al., 2017). Molecular descriptors are often used for describing the structural or physical chemical characteristics of a molecule and can be generated by conventional or molecular encoding. (55). The comprehensive definition of this descriptor was also examined elsewhere (61). For machine learning, the gathering of descriptors is vital, because it may simplify calculations, increase model generalization and enhance model output and interpretability (63). Dragon is a popular molecular descriptor calculation tool.

2.7 POLY PHARMACOLOGY DATA DRIVEN

New techniques are being provided by the scientific community to gather insights into the overall topology, complexity and medicinal product and target-disease interactions, due to the resurgence of AI and the digital health revolution (proteomics, clinical, and molecular investigations of patients and disease states, etc.). This strategy might contribute to data-led polypharmacology scanning by discovering novel therapeutic goals and disclosing the key goals for a certain condition, becoming a mainstream paradigm to possible drug discovery and manufacture. Polypharmacology discovery[44][45] applies in order to deliver the required therapeutic outcomes to the synthesis of one pharmacological molecule which interacts with several targets within a molecular network linked to disease. It is considered to be a possible tool for generating better and less harmful medicines to heal complicated conditions such as inflammation, Parkinson's disease, neurodegenerative illness, diabetes, Alzheimer's disease, cancer, etc. Of course, the successful creation of polypharmacology profiles are frequently important, as explains in the drug recycling section, for identifying possible goals for current medicines..

Multiple target combinations are frequently accessible in order to manage a disease network and an effective combination should be easy to regulate with drugs while retaining excellent network control. The complex interconnections between medications, targets and disorders in the disease network are at now not fully understood. However, this sort of culprit analysis implies that AI can play a key role due of its great data collecting and mining abilities. I can help with another important difficulty in research on polypharmacology: the rational design of multi-market drugs, which has proven challenging as the molecules need to co-ordinate their activity against many goals, particularly when the goals are from distinct families of proteins. Virtual screening is one option using AI-assisted systems. A database of virtual compounds is simultaneously or sequentially screened to identify therapeutic candidates that could bind to all the objectives concerned. This strategy requires a well constructed library with enough drug applicants to cover a wide variety of chemical areas. Continuous study on enhanced molecular images would also aid to achieve logical design. In addition, using AI-assisted de novo medication creation, compounds with the required polypharmacological profiles might be produced to tackle the question of rational design. Directly. Directly. Typically, AI-assisted de novo architecture creates a larger variety of structure compared to a virtual screening method that involves existing compound libraries which leads to a better overall chance of identifying multimarket compounds in especially with targets with different binding packs. This enables a more concentrated approach. Recent developments in multitasking and enhancement methods of in-depth learning have boosted expectations of multi-target AI-aideted exploration of polypharmacology, such that both primary (on and off-target ties of the small molecule to targets) and secondary design constraints are taken into account simultaneously (stability, solubility, lipophilicity, synthetic accessibility, ADMET properties, etc.). One worry is that AI models function well if each task has sufficient related data, but have demonstrated over-confidence if the data are insufficient for trustworthy prediction.. One option to alleviate the harmful impacts of data sparsity is to incorporate computer chemistry methodologies into the model development process. New data points may be constructed around an unknown area where no structural property data have been included in a dataset, for example by selecting a particular sub-set from the presently formed structures.. It is still a work in progress to develop a technology, which can smoothly integrate molecule creation, fine tuning and adaptive data production. Pharma companies know the immense revenue prospects of AI and big data in Liguria ovary polypharmaceutical. For instance, Sanofi has a deal worth $270 million. In 2016, it is important to identify bispecific small compounds for curing diabetes and associated co-morbidities. In addition, Essential committed to work together with the

German pharmaceutical company Evotec in order to create particular immunotherapy for cancer.. Importantly, while artificial intelligence and big data have generated a little of excitement throughout polypharmacology research, the road to success will require substantial further technique advancement.

2.8 DISCOVERY (AI) FOR COVID-19

AI is currently being used extensively in drug testing for Coronavirus disease (COVID19) because AI platforms can be more effective for detecting possible existing drugs with inhibitory human coronavirus (HCoV) activities by using various learning datasets such as [46].

Reported or proven potential active compounds for combating SARS-CoV, SARSCoV-2, influenza virus, human immunodeficiency virus (HIV) [1]).Inhibitors for known 3C-like protease (main protease of SARS-CoV-2) and other important protein targets translated by the SARS-CoV-2 viral genome [47]

The fighting feline coronavirus may be used in an in-vitro cell-based assay for all drugs expected using AI-based drug discovery, and the findings from these assays would provide input to the AI system, allowing it to relearn [48]. This could lead to the creation of a new AI-based paradigm for the re-search of existing drugs. This advantageous technology is used in gradually speeding up drug testing, where normal testing takes a long time, and therefore aids in essentially speeding up this process, which could be impossible for a person to do [49]. It has also proven to be an invaluable tool for diagnostic testing and the production of drugs and vaccines at a much faster pace than anticipated, as well as for clinical trials during vaccine development. This work's hypothetical perspective explains designing an anti-coronavirus method based on artificial intelligence to classify anti-coronavirus targets in the form of peptide drugs. We can search anti-coronavirus target protein sequences against anti-coronavirus target datasets using AI-based tools to provide precision-based

anti-coronavirus peptides/drugs based on the anti-coronavirus target protein sequences identified from the current outbreak.

Artificial Intelligence in Computational Drug Designing seeks high-quality studies on artificial intelligence approaches to leveraging the capacity of Computational Drug Designing by integrating Artificial Intelligence and core chemistry [84]. Computational Drug Designing is a growing field of study that focuses on the design and testing of molecular properties, interactions, and behavior in order to build better materials, processes, and systems for specific purposes. COVID-19 drug design approaches are progressing at the same time as computational artificial intelligence and molecular chemistry. This technique is proving to be a useful method in medicinal chemistry for identifying the starting points for COVID-19 hit molecules. This method cuts down on the time and money spent on drug research and development. The applications that use an AI-based approach for drug design are specifically concerned with the molecular structure of the drugs. AI-based applications are important for identifying new drug candidates and optimizing drug repurposing by extracting data and information from engines. The COVID-19 is becoming the benchmark for “Artificial Intelligence and Computational Drug Designing” strategies, opening up new avenues for drug development [86]. As the cost of drug production increases while investment returns decline, new approaches are needed.

Several companies are now using AI to classify and screen existing drugs that could be repurposed for the treatment of COVID-19, as well as to help clinical validation, sift through trial results, and scour patient electronic medical records (EMRs). TCS0 Innovation Lab in India, for example, found 31 possible hits that could act as COVID19 inhibitors [50] by a team of TCS scientists. Similarly, Benevolent AI, a company that raised $292 million to implement AI-based COVID-19 drug discovery, came up with an already approved drug for COVID-19 as an effective treatment using an AIbased drug discovery method [51].UK-based company Exscientia already team-up with Diamond Light Source (UK’s national synchrotron science facility) to utilize its AI drug discovery platform for identifying potential compounds against COVID-19 [8]. In an attempt to help researchers quickly synthesise and evaluate possible candidate molecules against COVID-19, The Molecule. one, a European AI-centered startup, has made its proprietary syntheses preparation tool available to the

scientific community for free. IBM has also utilized its AI generative frame- works to three COVID-19 drug targets and has generated 3000 novel potential hits or molecules[54].

CONCLUSION

In order to extract pharmacological information from enormous volumes of data, artificial intelligence (AI) technologies, notably deep learning, can be applied (e.g., QSAR and chemical structure). The experience acquired will then be utilized to identify and develop the required features of a molecule in order to maximize molecular attributes and increase the success rate of the clinical approval. AI technology has brought fresh life to computer-aided medication creation due to its great data extraction capabilities. However, several difficult situations may also arise: (1) The number of accessible data as a data mining technique significantly affects the output of comparable deep learning models as the high level of neural network training is largely dependent on massive amounts of data. The progress of technology for transfer learning could be a possible answer to this issue. (2) A black box and the mechanics of the model are not known. To elucidate the process of the deep learning system, the new hypothesis was introduced, and the counter factual sample (for instance, LIME) was utilized to open the black box in the deep learning processes. On the other hand, the research on deep learning models is barely in its beginning.. (3) Training neural network models means changing many parameters, but only a small number of functional rules are available and comprehensive theoretical foundation for optimization of these models is not available. All parts of current medication research and development are anticipated to be included in AI technology in the near future. We anticipate that a breakthrough in current research will emerge as an advanced network of drug creations that includes theoretical computational results (e.g. molecular docking, modelling of molecular dynamics and quantum chemistry), omics data, chemical data, and biological data..

REFERENCE

1. lee,J -G .et.al.(2017)Deep learning in medical imaging: general overview.korean j.radiol.18,570-584
2. Guncer,G.et al.(2018) An application of machine learning to haematology diagnosis.sci. Rep. 8, 411
3. Koohy, H. (2017) The rise and fall of machine learning methods in biomedical research.F1000 Res. 6 http://dx.doi.org/10.12688/f1000research.13016.2
4. Butina D. Unsupervised database clustering based on daylight's fingerprint and tanimoto similarity: a fast and automated way to cluster small and large datasets. J. Chem. Inf. Comput. Sci. 39(4), 747–750 (1999).
5. Aliper, A.; Plis, S.; Artemov, A.; Ulloa, A.; Moshing, P.;Zhavoronkov, A. Deep Learning Applications for Predicting Pharmacological Properties of Drugs and Drug Repurposing Using Transcriptomic Data. Mol. Pharmaceutics 2016, 13, 2524.
6. Putin, E.; Mamoshina, P.; Aliper, A.; Korzinkin, M.; Moskalev,A.; Kolosov, A.; Ostrovskiy, A.; Cantor, C.; Vijg, J.; Zhavoronkov, A.Deep biomarkers of human aging: Application of deep neural networks to biomarker development. Aging 2017, 8, 1021−1033.
7. Mamoshina, P.; Vieira, P.; Putin, E.; Zhavoronkov, A.Applications of Deep Learning in Biomedicine. Mol. Pharmaceutics 2016, 13 (5), 1445−1454.
8. Aliper, A.; Plis, S.; Artemov, A.; Ulloa, A.; Moshing, P.;Zhavoronkov, A. Deep Learning Applications for Predicting Pharmacological Properties of Drugs and Drug Repurposing Using Transcriptomic Data. Mol. Pharmaceutics 2016, 13, 2524.
9. Putin, E.; Mamoshina, P.; Aliper, A.; Korzinkin, M.; Moskalev,A.; Kolosov, A.; Ostrovskiy, A.; Cantor, C.; Vijg, J.; Zhavoronkov, A.Deep biomarkers of human aging: Application of deep neural networks to biomarker development. Aging 2017, 8, 1021−1033.
10. Mamoshina, P.; Vieira, P.; Putin, E.; Zhavoronkov, A.Applications of Deep Learning in Biomedicine. Mol. Pharmaceutics 2016, 13 (5), 1445−1454.
11. Schneider, G. and Clark, D.E. (2019) Automated de novo drug design: are we nearly there yet? Angewandte Chemie 131, 10906–10917
12. Oakes K. Here's what . A lead is a chemical compound that shows promising potential that can lead to the development of a new drug as a treatment for a disease. Identified hits are screened in cell-based assays predictive of the disease state and in animal models of disease to characterise the efficacy of the compound and its probable safety profile scientists searched for in 2018: AI is up, stress is down. Nature. 2019. Available from: https://www.nature.com/ articles/d41586-018-07879-9
13. Sellwood MA, Ahmed M, Segler MH, et al. Artificial intelligence and drug discovery. Futur Med Chem. 2018;10:2025–2028. • Commentary on the application of AI in drug discovery
14. Khozin S, Kim G, Pazdur R. Regulatory watch: from big data to smart data: FDA's INFORMED initiative. Nat Rev Drug Discov. 2017;16:

15. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016;316:2402– 2410.
16. Ref: Tariq A, Akram MU, Shaukat A, Khan SA. Automated detection and grading of diabetic maculopathy in digital retinal images. J Digit Imaging 2013;26:803–812.
17. Ref: Rahim SS, Palade V, Shuttleworth J, Jayne C. Automatic screening and classification of diabetic retinopathy fundus images. In: Mladenov V, Jayne C, Iliadis L, eds. Engineering applications of neural networks. Communications in computer and information science. Vol 459. Basel, Switzerland: Springer Nature Switzerland AG, 2014:113–122.
18. Ref: Lam C, Yu C, Huang L, Rubin D. Retinal lesion detection with deep learning using image patches. Invest Ophthalmol Vis Sci 2018;59:590–596.
19. Ref: US Food and Drug Administration. FDA permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems. https://www.fda.gov/ NewsEvents/Newsroom/PressAnnouncements/ucm604357 .htm Accessed July 18, 2018.
20. Ref: Keel S, Lee PY, Scheetz J, et al. Feasibility and patient acceptability of a novel artificial intelligence-based screening model for diabetic retinopathy at endocrinology outpatient services: a pilot study. Sci Rep 2018;8:4330.
21. Ref: Matimba A, Woodward R, Tambo E, Ramsay M, Gwan-zura L, Guramatunhu S. Teleophthalmology: opportunities for improving diabetes eye care in resource- and specialist- limited Sub-Saharan African countries. J Telemed Telecare 2016;22:311–316.
22. Ref :. Han L, Luo S, Yu J, Pan L, Chen S. Rule extraction from support vector machines using ensemble learning approach: an application for diagnosis of diabetes. IEEE J Biomed Health Inform 2015;19:728–734.
23. Ref: Shankaracharya, Odedra D, Samanta S, Vidyarthi AS. Computational intelligencebased diagnosis tool for the detection of prediabetes and type 2 diabetes in India. Rev Diabetes Stud 2012;9:55–62.
24. Ref : Wei W-Q, Tao C, Jiang G, Chute CG. A high throughput semantic concept frequency based approach for patient identification: a case study using type 2 diabetes mellitus clinical notes. AMIA Annu Symp proc 2010;2010:857–861.
25. Bain, E.E. et al. (2017) Use of a novel artificial intelligence platform on mobile devices to assess dosing compliance in a Phase 2 clinical trial in subjects with schizophrenia. JMIR Health Unhealthy 5, e18
26. Ref : Perez-Garcia, J.L. et al. (2017) Strategies to design clinical studies to identify predictive biomarkers in cancer research. Cancer Treat. Rev. 53, 79–97
27. Ref: Deliberato, R.O. et al. (2017) Clinical note creation, binning, and artificial intelligence. JMIR Med. Inf. 5, e24
28. Li, L.; Snyder, J. C.; Pelaschier, I. M.; Huang, J.; Niranjan, U.N.; Duncan, P.; Rupp, M.; Müller, K. R.; Burke, K. UnderstandingM Learned Density Functionals. Int. J. Quantum Chem. 2016,116, 819−833.
29. Rupp, M. Machine Learning for Quantum Mechanics in Nutshell. Int. J. Quantum Chem. 2015, 115, 1058−1073.
30. Behler, J. Representing Potential Energy Surfaces by High-Dimensional Neural Network Potentials. J. Phys.: Condens. Matter2014, 26, 183001.

31. Behler, J. Perspective: Machine Learning Potentials forAtomistic Simulations. J. Chem. Phys. 2016, 145, 170901.
32. Behler, J. Constructing High-Dimensional Neural NetworkPotentials: A Tutorial Review. Int. J. Quantum Chem. 2015, 115,1032−1050.
33. Behler, J. First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems.Angew. Chem., Int. Ed. 2017, 56, 12828−12840.
34. Snyder, J. C.; Rupp, M.; Hansen, K.; Müller, K.-R.; Burke, K.Finding Density Functionals with Machine Learning. Phys. Rev. Lett.2012, 108, 253002.
35. Brockherde, F.; Vogt, L.; Li, L.; Tuckerman, M. E.; Burke, K.;Müller, K.-R. Bypassing the Kohn-Sham Equations with MachineLearning. Nat. Commun. 2017, 8, 872.
36. Yao, K.; Parkhill, J. Kinetic Energy of Hydrocarbons as aFunction of Electron Density and Convolutional Neural Networks. JChem. Theory Comput. 2016, 12, 1139−1147.
37. Snyder, J. C.; Rupp, M.; Hansen, K.; Blooston, L.; Müller, K.R.; Burke, K. Orbital-Free Bond Breaking via Machine Learning. J.Chem. Phys. 2013, 139, 224104.
38. Mills, K.; Spanner, M.; Tamblyn, I. Deep Learning and theSchrodinger Equation. Phys. Rev. A: At., Mol., Opt. Phys. 2017, 96,042113.
39. Liu, F.; Du, L.; Zhang, D.; Gao, J. Direct Learning HiddenExcited State Interaction Patterns from ab initio Dynamics and ItsImplication as Alternative Molecular Mechanism Models. Sci. Rep2017, 7, 8737.
40. Hase, F.; Valleau, S.; Pyzer-Knapp, E.; Aspuru-Guzik, A.˙Machine Learning Exciton Dynamics. Chem. Sci. 2016, 7, 5139−5147.(830) Carpenter, B. K.; Ezra, G. S.; Farantos, S. C.; Kramer, Z. C.;Wiggins, S. Empirical Classification of Trajectory Data: AnOpportunity for
the Use of Machine Learning in MolecularDynamics. J. Phys. Chem. B 2018, 122, 3230−3241.
41. Yao, K.; Herr, J. E.; Parkhill, J. The Many-Body Expansion Combined with Neural Networks. J. Chem. Phys. 2017, 146, 014106.
42. Li, Z.; Kermode, J. R.; De Vita, A. Molecular Dynamics with on-the-Fly Machine Learning of Quantum-Mechanical Forces. Phys.Rev. Lett. 2015, 114, 096405.
43. Botu, V.; Ramprasad, R. Adaptive Machine LearningFramework to Accelerate Ab Initio Molecular Dynamics. Int. J.Quantum Chem. 2015, 115, 1074−1083.
44. Pilania, G.; Gubernatis, J. E.; Lookman, T. Multi-FidelityMachine Learning Models for Accurate Bandgap Predictions ofS Comput. Mater. Sci. 2017, 129, 156−163.
45. Pilania, G.; Mannodi-Kanakkithodi, A.; Uberuaga, B.;Ramprasad, R.; Gubernatis, J.; Lookman, T. Machine LearningBandgaps of Double Perovskites. Sci. Rep. 2016, 6, 19375.
46. Mannodi-Kanakkithodi, A.; Pilania, G.; Huan, T. D.;Lookman, T.; Ramprasad, R. Machine Learning Strategy forAccelerated Design of Polymer Dielectrics. Sci. Rep. 2016, 6, 20952.
47. Huan, T. D.; Mannodi-Kanakkithodi, A.; Ramprasad, R.Accelerated Materials Property Predictions and Design Using Motif-Based Fingerprints. Phys. Rev. B: Condens. Matter Mater. Phys. 2015,92, 014106.
48. Pilania, G.; Wang, C.; Jiang, X.; Rajasekaran, S.; Ramprasad, R.Accelerating Materials Property Predictions Using Machine Learning.Sci. Rep. 2013, 3, 2810.

49. Lee, J.; Seko, A.; Shitara, K.; Nakayama, K.; Tanaka, I.Prediction Model of Band Gap for Inorganic Compounds byCombination of Density Functional Theory Calculations andM Learning Techniques. Phys. Rev. B: Condens. Matter Mater.Phys. 2016, 93, 115104.
50. Pyzer-Knapp, E. O.; Li, K.; Aspuru-Guzik, A. Learning fromt Harvard Clean Energy Project: The Use of Neural Networks toAccelerate Materials Discovery. Adv. Funct. Mater. 2015, 25,
6495−6502.
51. Gomez-Bombarelli, R.; Aguilera-Iparraguirre, J.; Hirzel, T. D.;´Duvenaud, D.; Maclaurin, D.; Blood-Forsythe, M. A.; Chae, H. S.Einzinger, M.; Ha, D.-G.; Wu, T.; et al. Design of Efficient MolecularOrganic Light-Emitting Diodes by a High-Throughput VirtualScreening and
Experimental Approach. Nat. Mater. 2016, 15,1120−1127
52. .Corey, E. and Wipke, W.T. (1969) Computer-assisted design of complex organic syntheses. Science 166, 178–192 Grzybowski, B.A. et al. (2018)
53. Chematica: a story of computer code that started to think like a chemist. Chem 4, 390–398
54. Durant, J.L., Leland, B.A., Henry, D.R., and Nourse, J.G. (2003). Reoptimization of Mkeys for use in drug discovery. J Chem Inf ComputSci 34, 1273–1280.

55. Duvenaud, D., Maclaurin, D., Aguilera-Iparraguirre, J., Hirzel, T., and Adams, R.P. (2015). Convolutional networks on graphs for learning molecular fingerprints. In Proceedings of the 28th InternationalConference on Neural Information Processing Systems, pp. 2224-2232.
56. Rogers, D., and Hahn, M. (2010). Extended-connectivity fingerprints. JChem Inf Model 50, 742–754.
57. Bender, A., And, H.Y.M., Glen, R.C., and Reiling, S. (2004). Similaritysearching of chemical databases using atom environment descriptors(MOLPRINT 2D): evaluation of performance. J Chem Inf Comput Sci44, 1708–1718.
58. Kadurin, A., Aliper, A., Kazennov, A., Mamoshina, P., Vanhaelen, Q.,Khrabrov, K., and Zhavoronkov, A. (2017a). The cornucopia of meaNingful leads: Applying deep adversarial autoencoders for new molecUle development in oncology. Oncotarget 8, 10883.
59. Wallach, I., Dzamba, M., and Heifets, A. (2015). AtomNet: A deep conVolutional neural network for bioactivity prediction in structure-basedDrug discovery. Mathematische Zeitschrift 47, 34–46.
60. Xu, Y., Pei, J., and Lai, L. (2017). Deep learning based regression andMulticlass models for acute oral toxicity prediction with automatic c-Hemical feature extraction. J Chem Inf Model 57, 2672–2685.
61. Kearnes, S., Goldman, B., and Pande, V. (2016a). Modeling industrial A-DMET data with multitask networks. arXiv:1606.08793v3.
62. Kearnes, S., Mccloskey, K., Berndl, M., Pande, V., and Riley, P. (2016b).Molecular graph convolutions: moving beyond fingerprints. J CompuAid Mol Design 30, 1–14
63. Wu, Z., Ramsundar, B., Feinberg, E.N., Gomes, J., Geniesse, C., Pappu, A.S., Leswing, K., and Pande, V. (2017). MoleculeNet: A benchmark for Molecular machine learning. arXiv:1703.00564v2.

64. Urban, G., Subrahmanya, N., and Baldi, P. (2018). Inner and outer recursiveNeural networks for chemoinformatics applications. J Chem Inf Model58, 207–211.
65. Smith, E.G., and Wiswesser, W.J. (1975). The Wiswesser Line-FormulaChemical Notation (New York: McGraw-Hill).
66. Ash, S., Cline, M.A., Homer, R.W., Hurst, T., and Smith, G.B. (1997).ChemInform abstract: SYBYL line notation (SLN): a versatile languageFor chemical structure representation. ChemInform 28, no.Ashburn, T.T., and Thor, K.B. (2004). Drug repositioning: identifying and Developing new uses for existing drugs.
67. Weininger, D. (2011). Simplified Molecular Input Line Entry Specification.Willett, P. (2006). Similarity-based virtual screening using 2D fingerprint Drug Discovery Today 11, 1046–1053.
68. Heller, S.R., McNaught, A., Pletnev, I., Stein, S., and Tchekhovskoi, D (2015). InChI, the IUPAC international chemical identifier. J Chemiform 7, 23.