IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

THE ROLE OF AI AND ML IN DRUG DISCOVERY

Vidhi Pandya[1], Dhara Patel[2], Tejas Patel[3], Grishma Patel[4], Dhanajay Meshram[5]

[1]M.Pharm Student, Pioneer Pharmacy College, Vadodara,
[2]Professor, Pioneer Pharmacy College, Vadodara,
[3]Associate Professor, Pioneer Pharmacy College, Vadodara,
[4]Associate Professor, Pioneer Pharmacy College, Vadodara,
[5]Principal, Pioneer Pharmacy College, Vadodara,

Corresponding author
Ms. Vidhi Pandya
Email-vidhipandya98252@gmail.com
Mobile No: 9825294714

ABSTRACT
Artificial Intelligence (AI) and Machine Learning (ML) are transforming drug discovery by streamlining processes and enhancing decision-making. These technologies analyze vast datasets, enabling the identification of novel drug targets and predicting the interactions of compounds with greater accuracy. AI-driven algorithms can model complex biological systems, accelerating the assessment of how drugs will perform in clinical settings. Additionally, ML techniques optimize high-throughput screening, allowing researchers to rapidly evaluate thousands of compounds, thereby increasing efficiency. AI also plays a crucial role in drug design, predicting toxicity, and personalizing treatments, which reduces development time and costs while improving success rates. By integrating AI and ML into drug discovery pipelines, the pharmaceutical industry is poised to develop more effective therapies tailored to individual patient needs, ultimately leading to better health outcomes and a more efficient drug development process. As these technologies advance, their impact on the future of medicine is expected to grow significantly.

Key words: Artificial Intelligence (AI), Machine Learning (ML), Pharmaceutical industry, Drug discovery

INTRODUCTION
The process of finding and developing new drugs is a complex, diverse endeavor. Preclinical research and clinical trials are the final stages of this complex process, which starts with target discovery and validation and continues via compound screening and lead optimization. The initial stage in this approach is to identify the biological targets and pathophysiological variables. Deciphering cellular and genetic targets requires study in proteomics, genetics, and bioinformatics. First, the hit, or first molecule, is identified that shows activity against the designated target. Chemical libraries or the process of isolating natural compounds from bacteria, plants, and fungi can be used to accomplish this. The next stage involves identifying the lead chemical with the best chance of becoming a medication. The practice of further tweaking a chosen lead to improve its specificity and efficacy even at lower dosages is known as lead optimization. Through an iterative cycle combining cellular tests and structure–activity relationships, the functional qualities of freshly synthesized therapeutic candidates are improved. Following that, the animal models are employed in in vivo investigations, such as toxicity evaluations and pharmacokinetic analyses. After thorough preclinical testing, the medication candidate is finally given to patients in clinical trials. A medication candidate's ability to provide the desired medical benefits while maintaining the patients' well-being is largely determined by the results of clinical trials.[1,2] The process is time-consuming and cumbersome.

For this reason, pharmaceutical businesses look for ways to cut costs and expedite their projects. Artificial intelligence (AI) refers to a machine's capacity to simulate cognitive processes such as learning and problem-solving, just like a human brain.[3] Artificial intelligence (AI) systems based on technology can mimic human intelligence using a variety of advanced tools and networks. In order to save time and increase profitability, AI-based technologies are being deployed more and more at various phases of the drug discovery process. These comprise a range of activities such as real-time cell sorting, cell classification, quantum mechanics (QM)-based compound attribute calculation, computational organic synthesis, compound generation, and more. [4]

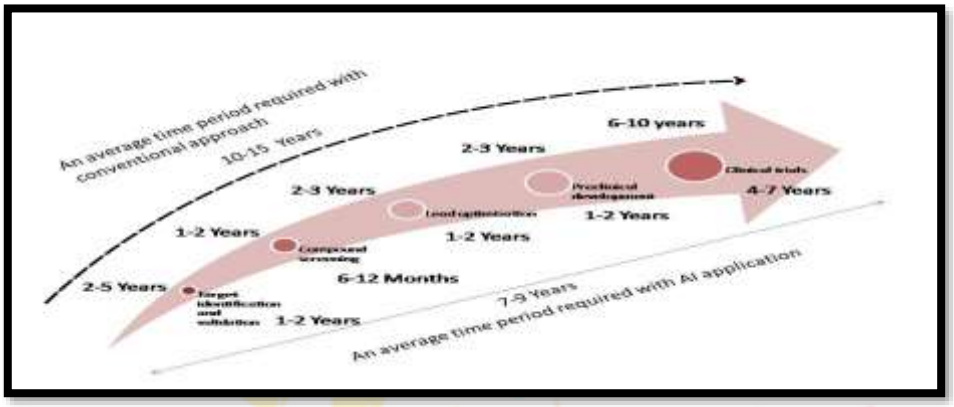

Figure 1: An examination of the relative times needed for important steps in the drug discovery process in relation to artificial intelligence.

It shows a comparison of the timescales for traditional and AI-driven drug discovery, giving an indication of the substantial speedups that AI technologies have brought about in this important area. Machine learning (ML) tools and algorithms are used in all stages of drug research and development to speed up the procedure. According to the McKinsey Global Institute, the way society functions will fundamentally change due to the rapid advancements in AI-driven automation.[5]Artificial Intelligence (AI) has gained prominence in medication development recently since it lowers the costs and duration of research as well as the failure rate of clinical trials. Because huge data for the life sciences is readily available and machine learning techniques are developing quickly, many artificial intelligence businesses are concentrating on medication development. These days, novel target identification, understanding disease-target associations, small-molecule design and optimization, disease mechanisms, comprehension of disease and non-disease phenotypes, identification of unknown prognostic biomarkers, and many other areas can all be seen to benefit from the use of machine learning.[6] The pharmaceutical sector has experienced a major data digitalization in recent years, but obtaining, validating, and applying this knowledge is challenging. AI's adoption is aided by its ability to process vast amounts of data through improved automation. To close the gap between random and rational drug development, a variety of machine learning (ML) methods, including Random Forest (RF) and Support Vector Machine (SVM), have been used in the drug design process. As a result, more and more medicinal chemists are tackling the fundamental issue of assessing and forecasting the biological effects of substances by utilizing a variety of AI methodologies. The approach known as pattern recognition has garnered significant attention due to its ability to elucidate and investigate common patterns among chemical substances. Its basic premise is that substances with similar structural formulas are probably going to have similar physicochemical characteristics and in vitro biological actions.

DRUG DISCOVERY PROCESS

Since there are many diseases or clinical conditions for which there is no viable medication on the market, drug research programs must continue to target these unmet clinical requirements. Proteins or pathways with potential therapeutic benefits are frequently found through early academic research; nevertheless, additional validation of the target choice may be necessary before drug development may move forward. Extensive research is carried out to identify a small-molecule or biological therapeutic candidate if hints are found. Before it may be sold as a prescription medication, preclinical and clinical testing are required. The pricey drug development process takes many years, from discovery to approval. Note that there may be variations in the related expenses and approval times. Drug research and discovery have historically benefited greatly from the application of computational techniques. However, there are still certain issues with traditional computing techniques, such as dependability, cost, and time. In drug discovery, protein–ligand interactions and binding stability are assessed using molecular dynamics (MD) simulation. Its potential for being very time-consuming and labor-intensive is one of its main drawbacks. Nowadays, medicinal chemistry makes extensive use of randomized procedures and large-scale high-throughput testing methodologies. These techniques look at a wide range of potential molecules to find ones that

have the necessary characteristics. But these methods can be costly, time-consuming, and frequently result in erroneous data. In fact, ligand-based virtual screening (LBVS) and other similarity-based techniques exhibit a high degree of "selectivity" when it comes to finding different active components. Its propensity for false-positive detections is a drawback in "real" applications. Drug discovery relies too heavily on trial-and-error testing, which leads to imprecise predictions about possible novel bioactive chemicals. Thus, the restricted old methods of pharmaceutical testing have worked well in the past. Industrial research and development (R&D) continuously assesses new technologies from a variety of research fields in an effort to speed up the discovery process. Investments in extremely complicated technology, enhanced production techniques, and creative research methods are made possible by the process's on going delays and rising costs. These investments will lead to real successes. These issues could be resolved with the use of a variety of AI-based algorithms, including rule-based, scalable, reinforcement learning, and supervised and unsupervised methods. The drug development industry is at a turning point in its history as medical innovation and technology convergence bring about paradigm shifts. Artificial intelligence (AI) and machine learning (ML), two extremely potent technologies that are transforming the drug research and discovery processes, are at the vanguard of this revolution. In a new era of customized medicine, the smooth integration of AI and ML has the potential to speed up research and improve efficiency. The FDA recognizes that AI/ML is increasingly being used in a variety of therapeutic domains and at different phases of the drug development process. In recent years, there has been a discernible increase in the incorporation of AI/ML elements in the submission of pharmacological and biologic applications. Furthermore, the submissions include a wide range of drug development operations, from post-market safety monitoring and sophisticated pharmaceutical manufacturing to original drug discovery and clinical trials. The European Medicine Agency recognized in a recent reflection paper that artificial intelligence is developing at a rapid pace and that safe and efficient research, regulation, and use of human and veterinary pharmaceuticals require a regulatory framework. Throughout the course of a pharmaceutical product's lifecycle, AI and ML tools can effectively assist with data collecting, transformation, analysis, and interpretation. Their application in preclinical development is multifaceted, utilizing AI/ML modeling techniques to reduce, eliminate, and enhance the usage of animal models. AI/ML systems can help select individuals in clinical trials based on particular clinical variables or disease traits. They can also enable data collecting and analysis that will be sent to regulatory agencies for marketing permission. With the use of these tools, scientists can now examine enormous volumes of data, spot trends and patterns, and forecast which substances will work best. They are also capable of forecasting the toxicity and effectiveness of these substances by simulating how they behave in the body. With the aid of these tools, researchers may now more quickly find possible medication candidates, estimate their efficacy, and shorten the time needed for development. [7]

BASICS OF AL AND ML

For every authorized drug, the conventional drug discovery process might take up to 15 years and cost between $1 and $2 billion. It is an extremely difficult and demanding process. This is mostly because clinical trials are taking longer to complete and attrition rates are increasing. Even after making it to the phase-I clinical trial, about 90% of possible treatment proposals fail, despite significant resource expenditure. For academic institutions and pharmaceutical corporations alike, moving a therapeutic candidate from rigorous preclinical optimization to a phase-I clinical trial is a major accomplishment. The success rate of lead compounds in clinical trials has been increased by the use of large-scale computer screening and docking. Nevertheless, these techniques have drawbacks like inaccuracy and inefficiency. Deep learning (DL) and machine learning (ML) algorithms—which are subsets of artificial intelligence—have been suggested as viable remedies to these problems. These artificial intelligence technologies have cheap processing costs and can accurately anticipate macrosystem features. As a result, chemical and biological experts are using AI algorithms more and more in the drug discovery process.[8]

1. Artificial Intelligence

Artificial Intelligence (AI) teaches machines and/or computers to carry out human jobs and make decisions when addressing an issue. The ability to learn through memory and adaptability, as well as the generalization ability gained from teaching machines to handle new difficulties, are the foundations of problem solving. The Turing machine, first shown by Alan Turing in 1963, is thought to have been the earliest example of artificial intelligence. Artificial Intelligence (AI) during the 1950s and 1980s was restricted to symbolic AI, which deals with logical problems like playing chess. AI developed to a higher level in the 1990s, using algorithms to predict new properties, analyze data, learn from it, and make judgments. In a word, machine learning (ML), reasoning, pattern recognition, natural language processing, and planning are all included in AI today.[9]

AI is committed to understanding intelligence computationally and creating things that can exhibit these behaviors. These things, or machines, are also capable of performing tasks that require human intelligence. The

procedure entails obtaining information, formulating application rules, drawing broad or detailed conclusions, and finally doing self-correction**.** An AI-based strategy's main advantage is that it can build a model and learn from examples even if we don't fully comprehend the underlying mechanism. AI is used in healthcare to diagnose and cure a variety of illnesses. Tempus for cancer treatment, PathAI for pathology diagnosis, IBM Watson for oncology, and Insilico Medicine for drug discovery are a few examples. The digitalization of data in the pharmaceutical industry has increased dramatically in the last several years. However, the emergence of digitalization has made it more challenging to get, assess, and use this knowledge when dealing with complex clinical problems. AI, then, is made up of a variety of state-of-the-art tools and networks that may mimic human intelligence.

2. Machine Learning

Machine learning (ML) is the process of teaching computers to interpret data and draw conclusions from examples and patterns. This could be unsupervised learning, supervised learning, or reinforcement learning. ML may be used to create prediction models from any type of data. The dataset needs to be cleaned, outliers removed, and missing values imputed before a model can be constructed. Important phases include selecting data,extracting features, choosing essential attributes and algorithms, developing models, and evaluating them. With machine learning (ML), a task may be taught to a machine by observing the links between the raw data input and the conclusions it draws. This is done without the need to write explicit programs that require expert knowledge. That is to say, machine learning (ML) is the process of teaching computers to learn from parsed data, such as clinical data, medical imaging, electronic health records (which include patient, laboratory, and billing information), and "omics."It is best to conduct the training over a number of iterations in order to enable precise identification of the desired attributes and trustworthy forecasts. Computational intelligence techniques including statistical pattern recognition, clustering,neural networks, probability theories, statistical learning, and pattern classification are all used in machine learning. Computational intelligence (CI) differs from artificial intelligence (AI) in that the latter uses genetic algorithms and neurons to convey information, drawing inspiration from biological processes like evolution. Good machine learning algorithms and large data sets are needed to develop models with high accuracy and precision. "Big data" refers to data that is large in volume (TB to ZB), very diversified (graphs, vectors, phrases, symbols, photos, etc.), and changes rapidly over brief periods of time. An effective machine learning model in AI is one that can, without over fitting. Predictions using extrapolated data are plausible, and interpolations—predictions inside the dataset range—are typically easy to produce. In other words, they generalize well to new situations that have not been integrated in the training database. Weight regularization, or keeping the weights in the neural network modest, is one technique to prevent over fitting (L 1 or L 2 variants). For gradient-based learning approaches, there is an additional meansof mitigating the vanishing gradient problem. Repaired linear activation or the pretraining approach, which trains each layer separately, can both prevent the vanishing gradient issue. The dropout strategy, which involves setting some randomly selected neurons to zero during training, is another way to reduce over fitting. Data augmentation, which involves applying database changes (such as rotations, cropping, noise addition, translations, scaling, etc.) to increase the amount of data points, might prevent over fitting, especially for photos. It is standard procedure in machine learning applications in the sciences to reserve a portion of the training set for validation. Model performance can be evaluated using metrics like kappa, normalized accuracy, precision, F1 scores, AUC, correlation, and accuracy. Drug development makes substantial use of machine learning (ML), which uses methods like support vector machines (SVM), clustering, random forests (RF), Bayesian networks (BN), and DL. Fig. 3 shows the wide classification of machine learning. Large volumes of data are processed and analyzed using DL models for applications like clinical imaging, virtual screening (VS and bioactivity predictions.[10]

Generally speaking, there are three main areas of machine learning: supervised, unsupervised, and reinforcement learning, each of which has multiple subclasses. These two techniques are used in supervised learning: regression, which predicts continuous outcomes from input features, and classification, which uses algorithms to classify input data into predetermined classes. In addition, related data points are grouped using unsupervised approaches like clustering. The architecture of machine learning algorithms is based on these subcategories, which have a major impact on a variety of drug discovery applications.[11]

A. Supervised Machine Learning

Supervised machine learning (ML) is the process of using labelled datasets to train algorithms for data categorization or precise result predicting. The model updates the weights using reinforcement learning (RL) after input is received to make sure the model fits correctly. Supervised learning is used by a multitude of sectors and organizations to solve a broad range of real-world problems. There are two forms of supervised learning: regression and classification.

B. Unsupervised Machine Learning

Unsupervised learning is applied when an input variable exists but an output variable does not. Its main goal is to gain a better understanding by comprehending the data distribution. It can be separated further into clustering and association.

C. Reinforcement Learning

RL is a fascinating subfield of machine learning that has garneredsignificant attention in both the academic and business domains. RL is an autonomous continuous learning strategy as opposed to supervised and unsupervised learning. Because RL is responsive to changing conditions, it is used in robots, gaming, and trading.

D. Deep Learning

Deep learning (DL) is a subclass of machine learning (ML) in which the system learns from unlabeled and unstructured data on its own. In the late 20th century, Igor Aizenberg et al. coined the term "deep learning" to refer to an artificial neural network (ANN).

AI IN DRUG TARGET IDENTIFICATION

To bring a new drug to market, the lengthy, costly, and hazardous process of drug development must take more than ten years and two billion dollars. By 2022, fewer than 500 viable drug targets had been identified—a small fraction of the anticipated total number of human druggable targets.[12] One of the most crucial steps in identifying a disease's biological origin and developing successful treatments is target identification. It is the process of deciding which biological molecules or cellular pathways to target with medications in order to produce desired therapeutic effects. Biomedical data, from clinical studies to basic research into the origins of disease, has become more widely available in recent years. But there are problems with data analysis because of this volume of information: noise, scalability, integration, quality, interpretability, validation, and computing complexity. AI is capable of handling and analyzing such intricate biological data networks. Recently, a promising technique that integrates AI algorithms with multi-omics data for target identification has been created. In order to prioritize treatable genes and find possible therapeutic targets in amyotrophic lateral sclerosis (ALS), Puna et al. used a variety of bioinformatics and DL-based models trained using disease-specific text and multi-omics data. This resulted in the identification of 18 prospective ALS therapeutic targets. PandaOmics evaluates targets using data on druggability, developmental stage, and tissue specificity in addition to their correlations with certain diseases. It does this by utilizing more than 20 AI and bioinformatics models. PandaOmics could predict the target genes associated with a specific disease by combining omics AI scores, text-based AI scores, finance scores, and key opinion leaders (KOL) ratings. This was accomplished by utilizing advanced DL models and AI methodologies. Using a drosophila model that mimics ALS, suggested AI therapeutic targets for ALS were validated. This discovered eight previously unreported targets, the removal of which dramatically reverses ocular neurodegeneration. In the same therapeutic field, Zhang et al. also created an ML-based method to identify KANK1 as a novel ALS-related gene and verified the neurotoxic effects of KANK1 mutations reproduced by CRISPR-Cas9 in human neurons.

AI has drawn a lot of interest, and ML-based algorithms—particularly AI techniques—have yielded impressive outcomes in the pharmaceutical sector. Compared to earlier ML techniques, contemporary DL designs like generative adversarial networks (GANs), recurrent neural networks (RNNs), and transfer learning approaches are garnering greater attention and are employed in a greater number of healthcare applications. Large Modules (LMs) with AI capabilities aid in streamlining searches and expediting target identification. For the purpose of recognizing substances, diseases, and genes buried in free-text materials, automatic biomedical named entity recognition(BioNER) is a helpful technique in addition to utilizing a multitask learning neural network with shared character and word layers (MTM-CW). Potential benefits of AI-driven learning machines (LMs) include data analysis and target identification and prioritizing support.

Furthermore, Fabris et al. created a DL technique with a new modular design to uncover human genes linked to a number of age-related disorders utilizing genetic or protein features such gene ontology words, protein–protein interactions, and biological pathways.

Data that have been purposefully produced to mimic patterns and traits found in the actual world are known as synthetic data. Scientists may investigate and evaluate a greater range of potential outcomes by using artificial intelligence (AI) algorithms to create synthetic data that replicates various biological conditions. Furthermore, synthetic data that offers a better level of assurance for the target identification process can be used to validate the predictions made by AI systems. Algorithms for bioanalysis based on machine learning (ML) and bioanalysis based on networks comprise two categories of AI algorithms. Drug development and the discovery of predicted cancer targets are two typical uses for these. Nevertheless, three major challenges remain in the way of cancer medication development and target discovery, even with the application of ML algorithms. The challenges are

integrating diverse information and lacking trustworthy data for validation. and DL models that are challenging to comprehend.

COMPOUND SCREENING WITH AI

Finding tiny compounds that can change the target protein's function and the phenotypic features of a disease is the goal of drug discovery. Additionally, it is necessary to look for tiny compounds with low toxicity and good pharmacokinetic qualities. Preclinical toxicity evaluation, pharmacokinetics, validation, and drug candidate discovery are challenging, costly, and time-consuming procedures. A medicine must be developed for ten to twelve years on average before it can be sold, and each one costs between USD 800 and USD 1.8 billion. Pharmaceutical research entered a breakthrough age when AI was used to drug screening, which resulted in a 50% reduction in R&D expenses and an increase in accuracy and efficiency. AI tackles a number of drug screening-related issues, such as physicochemical property prediction and bioactivity and toxicity assessment.[13]

APPLICATIONS OF ARTIFICIAL INTELLIGENCE

I. AI in Pharmaceutical Science

The scientific methods used in medication development and discovery are widely varied and comprise the pharmaceutical sciences. The growth of health care services is going to need a lot of work. AI offers the most effective strategy for improving the healthcare system. In the pharmaceutical industry, artificial intelligence (AI) and machine learning (ML) are already being used for a variety of tasks, including radiotherapy, clinical trials, disease diagnostics, drug discovery, and active chemical evaluation.

II. AI in Pharmacology

The pharmaceutical sciences are greatly impacted by AI/ML. When compared to conventional technologies, the use of several analytical techniques, such as magnetic resonance imaging (MRI), X-ray, electrocardiogram (ECG), and histopathological imaging, yields more refined results with the use of variable sensor and data acquisition systems. Furthermore, AI offers medical services related to animal research, including the identification of appropriate medications for particular illnesses as well as animal behavior, movement, and physiological and pathological changes.

III. AL in Medical Diagnosis

Numerous publications on a wide range of AI applications show that AI is effectively used in medical diagnosis, including those for skin cancer, neurological conditions, strokes, Alzheimer's disease, acute ischemic stroke, etc. Infectious infections are also treated with natural language processing (NLP), the first reading aid that gave physicians exceptional flexibility in analysing the characteristics of chest X-rays and the AI-based least square support vector machine (LSSVM), which is used to diagnose cancer. Two years later, SVM was applied to find imaging biomarkers for both psychiatric and neurological diseases.

AI is used in veterinary and agricultural sciences, including disease identification, chewing patterns, hybrid prediction, and cow behaviour, in addition to human-based databases. ANN models are developed for rumen fermentation patterns using dairy cattle.

IV. AI in De Novo Drug Design

The goal of de novo drug design (DNDD), which is based on artificial intelligence (AI) and uses a variety of approaches, such as auto encoders (AE), graph neural networks (GNNs), recurrent neural networks (RNNs), and GAN, is to create novel compounds with desirable qualities that have never been seen before. Typically, the algorithms consisted of two steps: first, the model automatically generated new molecules based on rules (SMILES, molecular graph) from the worthy databases (CHEMBL, ZINC, PubChem); second, reinforcement learning methods accelerated the exploration of novel regions to design structures with promising activities. The advantages of DNDD are enormous, including increased chemical space exploration, reduced costs, expedited design of structures with intellectual properties, and so on. Nevertheless, there are still difficulties, such as the need for training datasets, regulatory acceptance and standardization of models, analysis, sharing platforms, and synthetic formula procedures.

RNN-based generative models can be used for multi-objective evolutionary DNDD and are appropriate for sequential data (SMILES). The RNN model is a multi-objective strategy that targets neuraminidase, acetyl cholinesterase, and the new SARS-CoV-2 major protease. It consists of three layers, each having 512 gated recurrent units. This approach worked well for the phases of lead generation and optimization; the compounds produced had pertinent physicochemical characteristics (MW, logP, HBA, and HBD). Additionally, the training set for the recently developed bidirectional generative RNNs for SMILES-based molecule design (BIMODAL) included 30,000 unique and novel SMILES samples. This method was found to be suitable for scaffold diversity and chemical-biological relevance, as measured by the FCD values, which were lower than those of other models

with 1024 hidden units and FCD = 1.59 ± 0.03 when the starting point was fixed and FCD = 1.62 ± 0.04 when the starting point was random.[14]

V. AI in ADMET Prediction

The development of increasingly realistic models has coincided with the rapid advancement of artificial intelligence (AI) techniques in the field of in silico absorption, distribution, metabolism, excretion, and tolerated toxicity (ADMET) assessments. In this discipline, techniques like DNNs, ANNs, RFs, SVMs, and k-NN have opened up a lot of space with good performance; nevertheless, a number of criteria, including encoding functions and the quantity and quality of input data, should be taken into account. One of the biggest factors influencing bioavailability of drugs is intestinal absorption, as multiple studies have shown to have an acceptable range of accuracy. To predict the colon cancer cell layer (Caco-2) using a dataset of 104 and 26 molecules (training and test set, respectively), an inventive hierarchical SVM scheme was created. The model had good qualitative performance and revealed great accuracy. Previously, cellular permeability prediction in Caco-2 cell lines of substances was achieved with good discriminant power by employing DNN architecture, which was proposed utilizing 209 molecular descriptors. Four methods were used to predict permeability in a bigger dataset (1272 compounds): MLR, partial least-squares PLS, SVM, and boosting. The boosting model was merged with the other methods to determine which was the most suitable, achieving the highest Q2, RMSE, CV, and R2 values.

Overall, strong advancement and excellent reliability have come from ADMET evaluation using machine learning techniques. As of right now, in vitro and in vivo measures are still being used in addition to AI-based ADMET prediction algorithms.The drug hit-to-lead and lead-optimization processes now take less time and money because to this growing method.

VI. AI in Adverse Drug Reaction

AI is essential to pharmacovigilance.In addition to helping to shorten the time and burden associated with data processing, it first uniquely identifies the incoming data and the kind of adverse medication occurrences. Additionally, it improves the information's quality and assesses the case studies impartially.But given this arrangement, the economic component is still debatable.[15] AI uses cutting-edge approaches from data gathered both before and after the medication candidate is marketed, such as ML and DL algorithms.

Mixed elements are gathered by the computerized reporting system from different healthcare facilities.DL has an enlightening effect since it combines elements like picture and speech recognition and natural language processing to produce meaningful data with the highest accuracy.Additionally, a number of research indicate that neural networks and their extensive web work have enhanced the analytical application of deep learning.The DL models that are now in use greatly alter the raw data and accurately identify the clinical results. In contrast, machine learning (ML) is an algorithmic method that establishes parameters between variables and creates a model using the provided data in order to produce precise predictions.

An organization called Individual Case Safety Reports (ICSRs) is responsible for providing information on adverse events, product flaws, and consumer complaints while adhering to FDA standards. Many machine learning approaches are used to decrease labor costs and improve work efficiency. Initially, all of the unprocessed data—whether organized or not—is entered. The typically raw ICSR content is then extracted using machine learning and natural learning techniques. Currently, AI is crucial in enumerating the incidents, categorizing the medications according to the specifications, and performing the appropriate correlations.

VII. AI in Drug Repurposing

The process of developing new drugs requires a great deal of work, money, and time. Because a known candidate can have many target sites, drug repurposing offers a fundamental possibility to use the current prospects for new therapeutic reasons. Numerous computational techniques, such molecular docking, which has produced a sizable database for the assessment of the effects of drugs on numerous targets, have made this easier. The connection map (CMap), which is utilized for mRNA and gene expression in GWAS (genome-wide association studies), is one such instance.The pharmaceutical business has been approached by excellent AI platforms, as PREDICT, Netlap RLS, and DTINet, for drug repurposing for heterogeneous data sources. The majority of the studies that have been published to date have used learning algorithms to generate precise predictions based on the establishment of a strong association between the medications, the targets, and the illness.The diverse data sources include the biological pathways, disease-dependent phenotypes, and the small compounds being studied. The relative significance of each component is still unknown, though.

Three separate kernels are used for varying information levels in order to implement this strategy.Beginning with a structure based on a kernel that collects data on the similarities between the chemical configurations.The other kernel is based on transcriptional data, which provides information about gene expression based on drug-drug similarities. The kernel also provides information about targets, including target distance and protein interaction.When this data is further combined, output predictions are produced; this process is known as

supervised learning. However, in some circumstances where high-quality sample data is unavailable, specific unsupervised and semi-supervised learning techniques are used.

For training, the unsupervised computational method makes use of the previously established predictors regarding the drug labels. These rely on algorithms for clustering. For clustering methods, this unsupervised algorithm and the topographical pharmacophore descriptor (CATS) are employed. However, unsupervised learning only achieves a moderate level of prediction accuracy. A significant model for a small number of labelled training sets and enormous amounts of unlabelled input information is the semi-supervised learning paradigm. LapRLS is one such instance, which creates algorithms for drug-target interactions while being aware that target predictors alone were used for FDA approval. On the other hand, this approach receives high marks because of its simultaneous prediction feature. Other methods are LPMIHN, BLM-NII, and Net CBP.

Still, the use of AI-based medication repurposing is still in its infancy. In order to guarantee its wide range of applications in the field, the system must first surpass the manual forecast accuracy achieved by the experts.[16]

CONCLUSION

The entire world is on the way of digitalization and for the purpose the artificial intelligence and machine learning concepts play important role. Our article is totally based upon, how the intelligence and new machine technologies get invented in our day to day life. Today's machine are ready to give the knowledge based education and are responsible for improving the intelligence. In future, we don't think and imagine about the progress of world due to only AI and ML.

REFERENCE

1. Gentile, F., Yaacoub, J. C., Gleave, J., Fernandez, M., Ton, A. T., Ban, F, & Cherkasov, A. (2022). Artificial intelligence–enabled virtual screening of ultra-large chemical libraries with deep docking. *Nature Protocols*, *17*(3), 672-697.

2. Bule, M., Jalalimanesh, N., Bayrami, Z., Baeeri, M., & Abdollahi, M. (2021). The rise of deep learning and transformations in bioactivity prediction power of molecular modeling tools. *Chemical Biology & Drug Design*, *98*(5), 954-967.

3. Kapsiani, S., & Howlin, B. J. (2021). Random forest classification for predicting lifespan-extending chemical compounds. *Scientific reports*, *11*(1), 13812.

4. Manne, R. (2021). Machine learning techniques in drug discovery and development. *International Journal of Applied Research*, *7*(4), 21-28.

5. Birhane, A., Kasirzadeh, A., Leslie, D., & Wachter, S. (2023). Science in the age of large language models. *Nature Reviews Physics*, *5*(5), 277-280.

6. Alballa, N., & Al-Turaiki, I. (2021). Machine learning approaches in COVID-19 diagnosis, mortality, and severity risk prediction: A review. *Informatics in medicine unlocked*, *24*, 100564.

7. Yu, L., Shi, X., Liu, X., Jin, W., Jia, X., Xi, S., ...& Sun, D. (2021). Artificial intelligence systems for diagnosis and clinical classification of COVID-19. *Frontiers in Microbiology*, *12*, 729455.

8. Castro, A. A., Antonio, T. D., Martínez, E. C., Gallardo, M. G., Gascón, M. B., & Pinos, D. D. (2021). Usefulness of chest X-rays for evaluating prognosis in patients with COVID-19. *Radiologia (English Edition)*, *63*(6), 476-483.

9. Xie, W., Wang, F., Li, Y., Lai, L., & Pei, J. (2022). Advances and challenges in de novo drug design using three-dimensional deep generative models. *Journal of Chemical Information and Modeling*, *62*(10), 2269-2279.

10. Li, W. T., Ma, J., Shende, N., Castaneda, G., Chakladar, J., Tsai, J. C., ...& Ongkeko, W. M. (2020). Using machine learning of clinical data to diagnose COVID-19: a systematic review and meta-analysis. *BMC medical informatics and decision making*, *20*, 1-13.

11. Chatzigoulas, A., & Cournia, Z. (2021). Rational design of allosteric modulators: Challenges and successes. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *11*(6), e1529.

12. Zhang, Q., Chen, Y., Ni, D., Huang, Z., Wei, J., Feng, L., ...& Zhang, J. (2022). Targeting a cryptic allosteric site of SIRT6 with small-molecule inhibitors that inhibit the migration of pancreatic cancer cells. *Acta Pharmaceutica Sinica B*, *12*(2), 876-889.

13. Shiri, I., Maleki, H., Hajianfar, G., Abdollahi, H., Ashrafinia, S., Hatt, M., & Rahmim, A. (2020). Next-generation radiogenomics sequencing for prediction of EGFR and KRAS mutation status in NSCLC patients using multimodal imaging and machine learning algorithms. *Molecular imaging and biology*, *22*, 1132-1148.

14. Srinidhi, C. L., Ciga, O., & Martel, A. L. (2021). Deep neural network models for computational histopathology: A survey. *Medical image analysis*, *67*, 101813.

15. Zhou, Y., Zhang, Y., Lian, X., Li, F., Wang, C., Zhu, F & Chen, Y. (2022). Therapeutic target database update 2022: facilitating drug discovery with enriched comparative data of targeted agents. *Nucleic acids research*, *50*(D1), D1398-D1407.
16. Pun, F. W., Ozerov, I. V., & Zhavoronkov, A. (2023). AI-powered therapeutic target discovery. *Trends in pharmacological sciences*.