IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, January 2024

# Role of Machine Learning Techniques in Advancing Drug Discovery

**A. Ramesh[1] and Dr. Pawan Kumar[2]**
Research Scholar, Department of Computer Science & Engineering[1]
Research Guide, Department of Computer Science & Engineering[2]
NIILM University, Kaithal, India

**Abstract:** *This study examines machine learning in drug development, including advantages, drawbacks, and future possibilities. The article's study of machine learning models for drug prediction, therapeutic target identification, and drug candidate development emphasizes machine learning's utility. The article discusses machine learning's ethical and regulatory issues, data quality requirements, and cooperation and data sharing requirements in drug development. The research also stresses the importance of transparent and responsible machine learning algorithm implementation and regulatory structures to ensure the safety and efficacy of innovative pharmaceuticals produced by these models. The article finishes with a discussion of future machine learning for drug discovery advances. Machine learning may be integrated with robots, automation, deep learning models, multi-task learning, and customized medicine. The authors advise tackling machine learning difficulties in drug development to speed up the process and guarantee patients have access to novel, effective drugs. They also recommend exploring these prospective research and development areas. This review article discusses machine learning in drug discovery, its pros and cons, and the important fields expected to drive future research and development. This essay will interest legislators, academics, and pharmaceutical producers that want machine learning to enhance patient outcomes and change drug development*

**Keywords:** Drug Discovery, Machine Learning

## I. INTRODUCTION

Drug development is challenging and develops new remedies for many diseases [1]. Health improvement and disease reduction need innovative medical research [2]. Drug development takes time, money, and skill [3]. Drug research typically encompasses target identification, lead discovery, lead optimization, preclinical testing, and clinical trials. Scientists identify biological targets linked to diseases [4]. This may be done by proteomics, genetics, and high-throughput chemical library screening. After finding a target, researchers may screen hundreds or millions of tiny compounds for therapy candidates that bind and change its activity [5]. The procedure called lead discovery. Lead compound chemical structure may be altered to improve efficacy, selectivity, and pharmacokinetics [6]. The practice called lead optimization. Animal models are used to test an enhanced lead chemical's safety and effectiveness [7]. If preclinical studies are promising, a chemical may undergo clinical trials for safety, effectiveness, and pharmacokinetics [8]. Drug research has evolved in recent decades, but numerous difficulties remain [9]. Complex biochemical processes make in vivo pharmacological effects challenging to predict. Another issue is pharmaceutical candidates' high clinical trial failure rate owing to inadequate pharmacokinetics, toxicity, or ineffectiveness. These difficulties are being addressed by machine learning and AI researchers. The AI subset "machine learning" finds data patterns and correlations using algorithms. Machine learning may find new drugs that treat particular diseases by analyzing massive volumes of data [10]. Machine learning's role in drug research and its many applications will be discussed next [11]. Future research and drug development issues using machine learning will be discussed. Drug development is improved by continually developing drug discovery technologies and methods [12]. Machine learning can quickly and accurately uncover new drug candidates by analyzing massive volumes of data [13].

Machine learning can examine complicated data sets and find patterns and correlations that conventional approaches overlook in drug development [14]. Genomic data may be searched for gene expression patterns or genetic variants

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, January 2024

linked to diseases using machine learning techniques [15]. Identifying pharmacological targets using this knowledge may lead to novel therapies. Machine learning can identify compounds that bind to targets and modify their activity using high-throughput screening data. This method drastically reduces lead discovery and optimization time and resources, allowing researchers to concentrate on the best treatments [16].

Machine learning may also change medication discovery in personalized medicine [17]. Big patient data sets may help machine learning algorithms identify patient subgroups most likely to respond to a pharmaceutical or therapy [18]. This may improve patient-specific, effective, and individualized treatment [19]. Machine learning in drug development has various obstacles. Poor data for machine learning algorithm training is a key concern [20]. Insufficient drug development data may reduce machine learning efficacy [21]. Finding interpretable models to explain how machine learning algorithms predict and detect data biases is another difficulty [22]. Drug research requires precise estimates to avoid disaster [23].

Despite these obstacles, machine learning for drug discovery offers numerous promising methods [24]. Deep learning algorithms may find new therapeutic targets in complex genomic data. Transfer learning trains algorithms on related data, whereas reinforcement learning enhances drug development. Machine learning might speed up and accurately find novel drug ideas by analyzing massive amounts of data [25]. Machine learning for drug discovery has drawbacks, but solutions are being explored. As medication research develops, machine learning will help create novel and effective treatments for many diseases [26].

**The Role of Machine Learning in Analyzing Vast Amounts of Data for Drug Discovery:**

To develop novel drugs, huge and sophisticated data sets must be evaluated. Machine learning (ML) helps researchers evaluate massive data sets rapidly and correctly [27]. Machine learning algorithms can uncover patterns and relationships in biological and chemical data to produce new drugs. ML can analyze genetic, protein, and chemical data. Combining data from several sources helps researchers comprehend disease and treatment biological processes [28]. A major feature of machine learning in drug development is its capacity to quickly and reliably examine large and complex data sets. This lets scientists identify the best drug candidates after additional research. ML can predict therapeutic candidate toxicity to prioritize drugs for further study. Machine learning can examine data in real time, which benefits medication development. By analyzing streaming data using ML algorithms, researchers may quickly uncover new patterns and relationships to develop new drugs [29].

Using ML in drug development has several drawbacks. For instance, ML models require high-quality data to forecast accurately. Additionally, interpretable models that illuminate pharmacological effect molecular pathways are needed [30]. Another issue is data bias and error, which may affect ML prediction accuracy.

Here Figure 1 depicts machine learning drug discovery. ML is utilized in drug development despite these challenges. Researchers are developing new algorithms that fully use ML for drug development [31]. Machine learning helps researchers quickly and consistently analyze large amounts of data to generate new health-improving drugs [32]. Supervised and unsupervised machine learning algorithms dominate [32]. Supervised learning trains a machine learning system on a labeled dataset containing outcomes or labels. Data patterns help the machine predict unlabeled data [33]. In drug research, supervised learning may predict effectiveness. Researchers can predict the effectiveness of new, untested drugs using a machine learning model trained on well-known pharmaceuticals and their efficacy data [34]. Instead, unsupervised learning trains ML systems on unlabeled datasets. Unlabeled data is analyzed for patterns and relationships. In large, complex datasets, unsupervised learning may uncover patterns and drug development targets. Machine learning is used in novel drug design [35]. ML algorithms may improve medication pharmacokinetics, toxicity, and effectiveness. ML drug development may be cheaper than trial-and-error [36]. ML optimizes drug doses and therapies. Based on patient and pharmaceutical response data, machine learning algorithms may predict therapy dose and timing, reducing side effects and improving patient outcomes [37]. Drug discovery is changing with ML. ML algorithms can quickly and correctly scan massive volumes of data to find therapeutic possibilities and create unique remedies. Machine learning (ML) in medication development has challenges, but researchers are developing new algorithms to overcome them and improve health [38].

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, January 2024

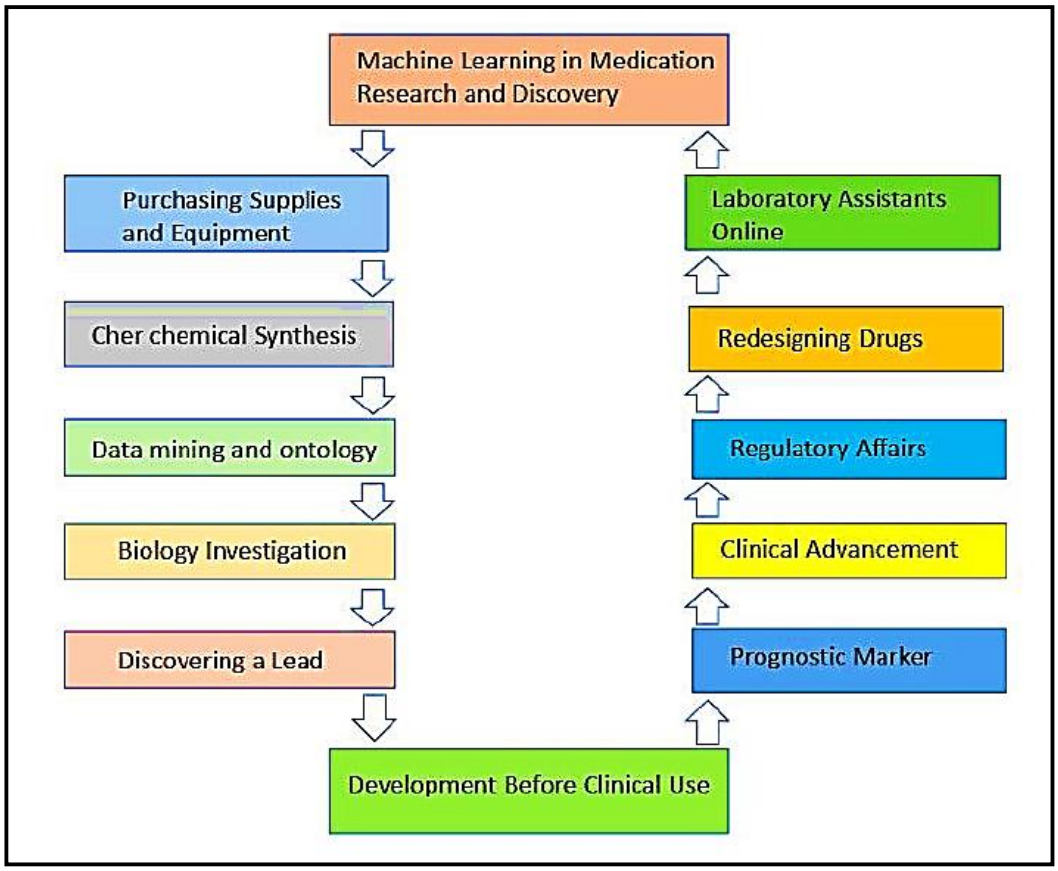


**Figure 1: Understanding Drug Discovery by machine learning through Flow Chat**

**Analyzing Vast Amounts of Data in Drug Discovery Using Machine Learning**

Historically, drug discovery has been time-consuming, expensive, and laborious [39]. Machine learning (ML), particularly for huge data processing, may revolutionize therapy discovery. Identification of chemical compounds that might be used to treat disorders is a crucial stage in drug development [40]. This technique identifies chemicals with biological activity and safety via several testing and analysis. Researchers may save expenses and speed up drug development by evaluating massive amounts of data using machine learning algorithms [41].

Machine learning's ability to swiftly and accurately analyze huge and complicated datasets is a major drug development benefit [42]. Machine learning algorithms may find patterns and connections in large data sets that conventional analytic approaches cannot [43]. This strategy may help researchers find medication candidates missed by typical screening methods. Large databases of chemical and biological data can train machine learning algorithms to predict new substance attributes [44]. Machine learning (ML) algorithms may be trained on enormous datasets of known pharmacological substances to discover the most promising drugs for treating particular illnesses [45]. Since researchers can rapidly discover promising pharmacological candidates, this strategy might minimize medication development time and cost [46].

Figure 2 illustrates AI drug development. Machine learning can predict new drug effectiveness and toxicity from large drug research data [47]. Large databases of well-known pharmaceutical substances and their efficacy and toxicity may train machine learning algorithms. Researchers can forecast new drugs' effectiveness and toxicity based on their chemical compositions and other features.

Machine learning algorithms can swiftly and reliably analyze enormous data sets, aiding drug development. Researchers can find medication candidates quicker and better. ML algorithms find data links and patterns that older approaches miss [48].

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, January 2024

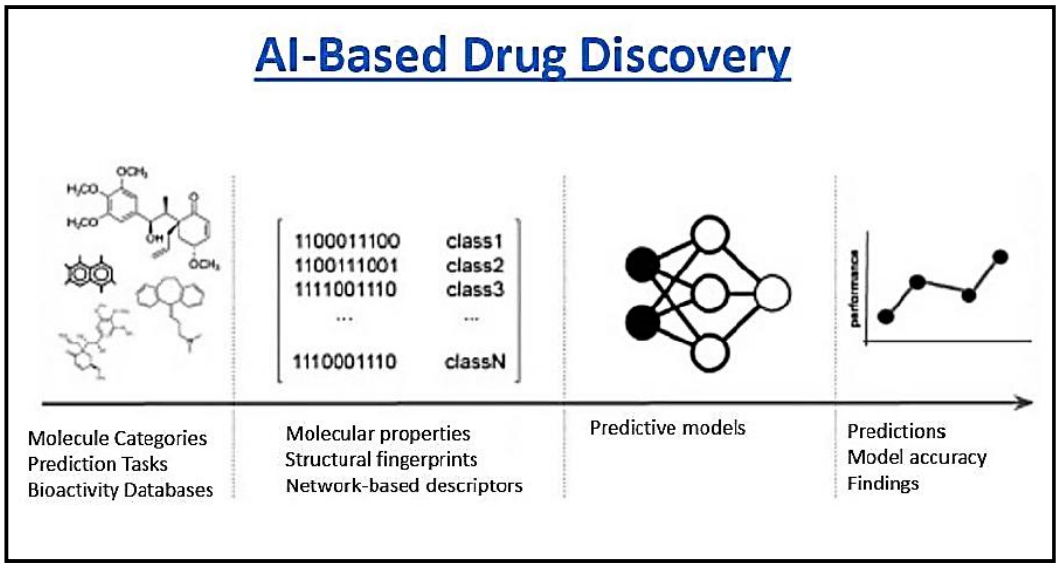


**Figure 2: AI based discovery of Drug through**

Using machine learning in medication development is hard. Needing plenty of high-quality data is difficult. Data quality influences prediction accuracy, and ML systems need large datasets to learn. Machine learning in drug development may need expertise not all companies have [49]. Machine learning in drug development may change the pharmaceutical industry despite these obstacles. ML algorithms may find therapy options quicker and cheaper than conventional approaches, decreasing drug development costs [50]. Machine learning algorithms can anticipate new chemical efficacy and toxicity, expediting medicine development and minimizing animal and human trials [51]. Analysis of vast volumes of data by machine learning may lead to novel drug discoveries [52]. Large biological and chemical compound databases may be examined by machine learning (ML) techniques to speed up medication candidate discovery [53]. Machine learning's potential advantages make it an interesting drug development research topic despite its limitations.

Deep learning algorithms can scan huge volumes of data and uncover drug candidates quicker and more accurately, expanding their drug development use [54]. Innovative drugs with great specificity and efficacy exist [55]. Deep learning can predict small molecule biological activity based on chemical structures, a prospective drug development application.

**Applications of Deep Learning in Drug Discovery**

Quantitative structure-activity relationship (QSAR) modeling uses deep neural networks to predict a compound's biological target behavior based on its chemical structure [56]. The neural network is trained on a vast dataset of known compounds and their biological activities to predict new chemical activity [57]. This method predicted the effects of hundreds of tiny compounds on many biological targets. One study employed a deep learning model to predict chemical reactions with cancer treatment target CHK1 [58]. The researchers used a vast dataset of well-known CHK1 inhibitors to build the model, which predicted the effect of novel drugs. The model discovered many drugs with significant CHK1 inhibitory action; one inhibited tumor growth in animal models [59]. Deep learning and generative models are used in medication research to create novel drugs with desired features. Large chemical datasets can train generative algorithms to find biological process patterns and structures [60]. These models may be used to create compounds with the desired activity after training. Researchers used a generative model to create chemicals that inhibited bromodomain-containing protein 4 (BRD4), a possible cancer treatment target. After training on a huge dataset of well-known BRD4 inhibitors, the model generated novel compounds. In vitro tests showed many drugs were very effective against BRD4. Deep learning can uncover novel therapies and predict drug toxicity and adverse effects. It may lower the number of dangerous medications that fail clinical trials. Using chemical structure, scientists constructed a deep learning algorithm

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, January 2024

to predict toxicity. After training on a huge dataset of chemicals with established toxicity profiles, the program could correctly evaluate novel compound toxicity.

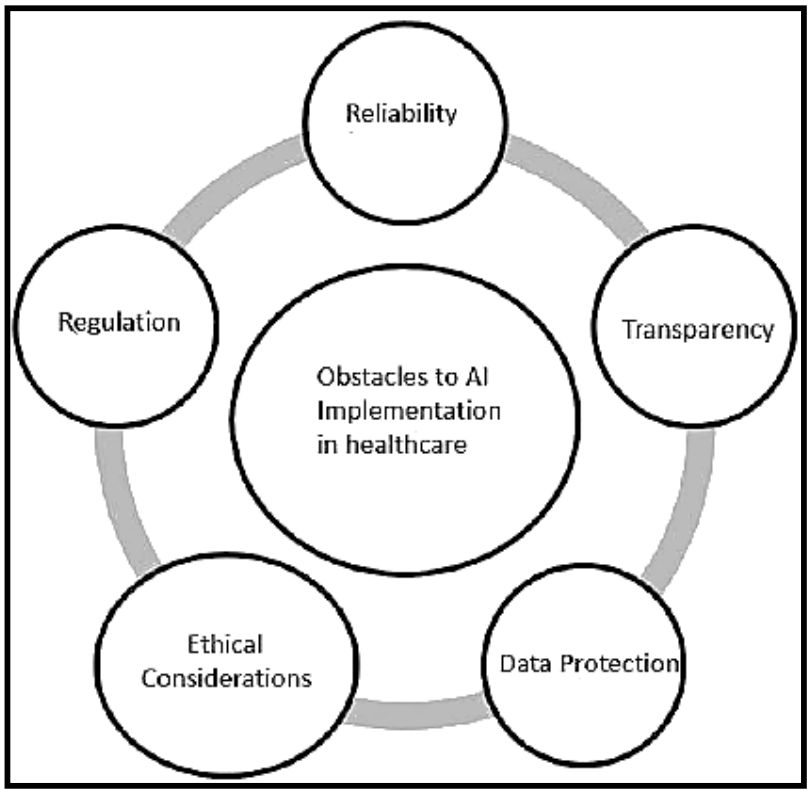


**Figure 3: Implication of AI in health care**

Figure 3 illustrates how AI is being used to healthcare. These examples demonstrate the huge potential of deep learning in drug discovery. Deep learning allows researchers to evaluate large amounts of data and create novel molecules with desirable qualities, which may expedite the drug development process and help provide new, effective treatments to patients more quickly [61]. However, there are several challenges with deep learning in drug development; these will be discussed in more depth in the section that follows. All things considered, deep learning in drug discovery may fundamentally alter the process of identifying and creating new drugs. Deep learning helps researchers analyze massive amounts of data and create novel molecules with desirable features, which may expedite the drug development process and help provide new, effective treatments to patients more quickly.

**Challenges and Limitations of using Machine Learning in Drug Discovery**

Machine learning has gained popularity because it helps accelerate medication development and identify new treatments faster and more accurately [62]. Machine learning must overcome several drug development obstacles to reach its full potential.

Machine learning drug development is hindered by data availability and quality. A large dataset of compounds with biological activity against specific targets is required to train machine learning algorithms to predict small molecule biological activity. However, this data is usually weak. Thus, model accuracy may decline due to biases. Machine learning models employ data from several sources, which may differ in quality and experimentation [63]. Data normalization and preprocessing issues may reduce model accuracy. Drug development using machine learning is challenged by model interpretability. Machine learning models are termed "black boxes" because their predictions are unclear. For medicine development, determining the biological processes and reasons behind tiny molecules' apparent action may be challenging. Training and inference need a lot of computer resources, making machine learning models complex. Small research teams or those without powerful computers may struggle. Medicine development via machine

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, January 2024

learning raises ethical, legal, and technical issues. Discriminatory machine learning algorithms may exclude certain patient populations or deliver less effective drugs for subgroups. Machine-learning drug development confronts regulatory difficulties. Machine learning-based medications may be difficult to approve since normal regulatory procedures were not designed to handle them [64]. Despite its shortcomings, machine learning has great promise for drug development. Machine learning can evaluate massive datasets and generate fresh medication candidates with desirable attributes, speeding drug development and improving patient outcomes. New methods and technologies must be developed to fully achieve this potential, along with resolving machine learning's drug development constraints.

Machine learning for drug development faces several more challenges. Development of drugs involves stakeholder input. Scientists, doctors, and business partners must work together to provide reliable data and models. Validating and testing machine learning models is crucial. Molecular activity can be predicted by machine learning algorithms, but experimental confirmation is needed. Important because new medications must be extensively examined for safety and efficacy before being allowed for human use. More transparency is needed for medication development using machine learning. As machine learning models get more complex, methods for assessing their conclusions and identifying data or model biases are required. Long-term impacts of machine learning in medical development must be recognized [65]. These technologies may accelerate medication research and improve patient outcomes, but they also present ethical and cultural issues and may increase treatment inequities. Machine learning-based drug discovery is tough and constrained, requiring ongoing research and development. By addressing these difficulties and creating new techniques and tools, machine learning may be fully used in medicine development to enhance patient outcomes safely and efficiently.

**Future Directions and Opportunities for Advancement in the Field**

Fast-growing machine learning in drug discovery offers several exciting paths and improvement chances. This article will discuss some of the important research and development fields expected to progress in the next years [66]. Deep learning models are interesting drug discovery machine learning research subjects. Artificial neural networks learn from data in deep learning. These models may improve prediction accuracy without requiring vast amounts of labeled data. Another area of investigation is multi-task learning models [67]. Multi-task learning uses a single model to learn several related tasks. This method predicts several chemical characteristics using a single model, possibly improving drug discovery speed and accuracy.

In addition to these technological advances, drug discovery using machine learning offers innovation. Personalized medicine using machine learning is becoming popular. The concept of "personalized medicine" tailors treatment to each patient's requirements. Machine learning may identify patient-specific biomarkers and predict therapy effects, improving patient outcomes. Innovation includes creating fresh data for machine learning algorithms. Imaging data for drug development is growing in popularity. Imaging data from MRI and CT scans may reveal tissue and organ composition and function, which can be utilized to predict micro molecule behavior. The use of machine learning to generate new drugs is also growing [68]. Generative modeling uses machine learning to produce molecules with desired properties. Generative modeling may reduce drug discovery time and expense by rapidly developing and evaluating many new drug candidates. Finally, machine learning may be used with automation and robots to accelerate medication development. Machine learning might be used for high-throughput screening technologies' enormous data processing and experiment design optimization. Machine learning for drug development has great promise, but it must overcome challenges to realize its full potential. A major issue is the need for better data. Machine learning may minimize drug development data, but successful model training still needs high-quality data.

Collaboration and data exchange among drug development players is another issue. Researchers, clinicians, and industry partners must share expertise to develop high-quality data and reliable models [69]. Additionally, drug development using machine learning needs more accountability and transparency. As machine learning models get more complex, strategies for understanding their results and identifying biases or limits in the data or models are essential.

Machine learning in drug development raises ethical and legal concerns. Discrimination and bias in machine learning algorithms might exclude some patient populations or provide less effective drugs for certain subgroups. The fast-growing area of machine learning for drug development has many exciting opportunities [70]. To maximize its potential, this technique must overcome its challenges. By addressing these difficulties and working together to create

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, January 2024

accurate and transparent models, machine learning can improve drug development and bring novel, effective medications to patients.

## II. CONCLUSION

To conclude, machine learning in drug development is a fascinating and rapidly expanding area that might speed up drug discovery and provide patients cutting-edge therapy. Machine learning algorithms may predict pharmacological properties, uncover novel therapeutic targets, and produce new drug candidates. Deep learning models, multi-task learning, tailored medicine, and merging machine learning with automation and robots are exciting research fields that will certainly improve in the future. Several challenges and limits must be handled before machine learning may be utilized to produce medications. These include the need for higher-quality data, cooperation and data sharing, accountability and openness in machine learning models, and ethical and legal issues. Machine learning in medication development has more room for growth. By working together to overcome these problems, we can employ machine learning to speed up drug development and provide patients better therapies.

## REFERENCES

**[1].** Utsha Sinha, Abhinav Singh, Deepak Kumar Sharma, Machine Learning in the Medical Industry, Handbook of Research on Emerging Trends and Applications of Machine Learning, 10.4018/978-1-5225-9643-1.ch019, (403- 424), (2020).

**[2].** Susanne Uusitalo, Jarno Tuominen, Valtteri Arstila, Mapping out the philosophical questions of AI and clinical practice in diagnosing and treating mental disorders, Journal of Evaluation in Clinical Practice, 10.1111/jep.13485, 27, 3, (478-484), (2020).

**[3].** Rashid, M. T., Zhang, D. Y., & Wang, D. (2019, December). Socialcar: A task allocation framework for social media driven vehicular network sensing systems. In 2019 15th International Conference on Mobile Ad-Hoc and Sensor Networks (MSN) (pp. 125-130). IEEE.

**[4].** Shams, A. T., & Akter, S. (2022). Eco-Centric Versus Anthropocentric Approach in Literary Pedagogy: Inclusion of Non-Human Narratives as Teaching Social Justice.

**[5].** Ozlem Erdas-Cicek, Ali Osman Atac, A. Selen Gurkan-Alp, Erdem Buyukbingol, Ferda Nur Alpaslan, Three- Dimensional Analysis of Binding Sites for Predicting Binding Affinities in Drug Design, Journal of Chemical Information and Modeling, 10.1021/acs.jcim.9b00206, 59, 11, (4654-4662), (2019).

**[6].** Sebastian Raschka, Automated discovery of GPCR bioactive ligands, Current Opinion in Structural Biology, 10.1016/j.sbi.2019.02.011, 55, (17-24), (2019).

**[7].** Javier Pérez-Sianes, Horacio Pérez-Sánchez, Fernando Díaz, Virtual Screening Meets Deep Learning, Current Computer-Aided Drug Design, 10.2174/1573409914666181018141602, 15, 1, (6-28), (2018).

**[8].** Rashid, M. T., Zhang, D. Y., & Wang, D. (2020). DASC: Towards a road Damage-Aware Social-media-driven Car sensing framework for disaster response applications. Pervasive and Mobile Computing, 67, 101207.

**[9].** Isidro Cortés-Ciriano, Nicholas C. Firth, Andreas Bender, Oliver Watson, Discovering Highly Potent Molecules from an Initial Set of Inactives Using Iterative Screening, Journal of Chemical Information and Modeling, 10.1021/acs.jcim.8b00376, 58, 9, (2000-2014), (2018).

**[10].** Dries Harnie, Mathijs Saey, Alexander E. Vapirev, Jörg Kurt Wegner, Andrey Gedich, Marvin Steijaert, Hugo Ceulemans, Roel Wuyts, Wolfgang De Meuter, Scaling machine learning for target prediction in drug discovery using Apache Spark, Future Generation Computer Systems, 10.1016/j.future.2016.04.023, 67, (409-417), (2017).

**[11].** Akhter, A., & Shams, A. T. (2022). Identity Economics in Emily Brontë's Wuthering Heights: An Empathetic Inquiry into Psychoanalysis. SCHOLARS: Journal of Arts & Humanities, 4(2), 74-80.

**[12].** Rashid, M. T., Chowdhury, P., & Rhaman, M. K. (2015, December). Espionage: A voice guided surveillance robot with DTMF control and web based control. In 2015 18th International Conference on Computer and Information Technology (ICCIT) (pp. 419-422). IEEE.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, January 2024

**[13].** Fenglei Li, Qiaoyu Hu, Xianglei Zhang, Renhong Sun, Zhuanghua Liu, Sanan Wu, Siyuan Tian, Xinyue Ma, Zhizhuo Dai, Xiaobao Yang, Shenghua Gao, Fang Bai, DeepPROTACs is a deep learning-based targeted degradation predictor for PROTACs, Nature Communications, 10.1038/s41467-022-34807-3, 13, 1, (2022).

**[14].** Muhammad Waqar Ashraf, Artificial Intelligence for Drug Development, Advances in Artificial Intelligence, Computation, and Data Science, 10.1007/978-3-030-69951-2_5, (127-132), (2021).

**[15].** Ling Hao, Tyler Greer, David Page, Yatao Shi, Chad M. Vezina, Jill A. Macoska, Paul C. Marker, Dale E. Bjorling, Wade Bushman, William A. Ricke, Lingjun Li, In-Depth Characterization and Validation of Human Urine Metabolomes Reveal Novel Metabolic Signatures of Lower Urinary Tract Symptoms, Scientific Reports, 10.1038/srep30869, 6, 1, (2016).

**[16].** Zhang, Y., Zong, R., Shang, L., Rashid, M. T., & Wang, D. (2021, June). Superclass: A deep duo-task learning approach to improving qos in image-driven smart urban sensing applications. In 2021 IEEE/ACM 29th International Symposium on Quality of Service (IWQOS) (pp. 1-6). IEEE.

**[17].** Harsh Chauhan, Jonathan Bernick, Dev Prasad, Vijay Masand, The Role of Artificial Neural Networks on Target Validation in Drug Discovery and Development, Artificial Neural Network for Drug Design, Delivery and Disposition, 10.1016/B978-0-12-801559-9.00002-8, (15-27), (2016).

**[18].** Sergio Ruiz-Carmona, Xavier Barril, Docking-undocking combination applied to the D3R Grand Challenge 2015, Journal of Computer-Aided Molecular Design, 10.1007/s10822-016-9979-z, 30, 9, (805-815), (2016).

**[19].** Dries Harnie, Alexander E. Vapirev, Jorg Kurt Wegner, Andrey Gedich, Marvin Steijaert, Roel Wuyts, Wolfgang De Meuter, undefined, 2015 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, 10.1109/CCGrid.2015.50, (871-879), (2015).

**[20].** Antonino Marvuglia, Mikhail Kanevski, Enrico Benetto, Machine learning for toxicity characterization of organic chemical emissions using USEtox database: Learning the structure of the input space, Environment International, 10.1016/j.envint.2015.05.011, 83, (72-85), (2015).

**[21].** Rashid, M. T., Zhang, D., & Wang, D. (2020, July). A Computational Model-Driven Hybrid Social Media and Drone-Based Wildfire Monitoring Framework. In IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS) (pp. 1362-1363). IEEE.

**[22].** Robert Wolfgang Rumpf, Samuel L. Wolock, William C. Ray, StickWRLD as an Interactive Visual Pre-Filter for Canceromics-Centric Expression Quantitative Trait Locus Data, Cancer Informatics, 10.4137/CIN.S14024, 13s3, (CIN.S14024), (2014).

**[23].** Bin Chen, Huijun Wang, Ying Ding, David Wild, Semantic Breakthrough in Drug Discovery, Synthesis Lectures on the Semantic Web: Theory and Technology, 10.2200/S00600ED1V01Y201409WEB009, 4, 2, (1-142), (2014).

**[24].** Donald Petrey, Barry Honig, Structural Bioinformatics of the Interactome, Annual Review of Biophysics, 10.1146/annurev-biophys-051013-022726, 43, 1, (193-210), (2014).

**[25].** Agarwal S, Dugar D, Sengupta S. 2010. Ranking chemical structures for drug discovery:A new machine learning approach. J Chem Info Model 50:716–731.

**[26].** Arodz T, Yuen DA, Dudek AZ. 2006. Ensemble of linear models for predicting drug properties. J Chem Info Model 46:416–423.

**[27].** Chowdhury, M. S. S., Nawal, M. F., Rashid, T., & Rhaman, K. (2015, December). Terminal analysis of the operations of a Rescue Robot constructed for assisting secondary disaster situations. In 2015 IEEE Region 10 Humanitarian Technology Conference (R10-HTC) (pp. 1-5). IEEE.

**[28].** Chen B, Harrison RF, Papadatos G, Willett P, Wood DJ, Lewell QX, Greenidge P, Stiefl N. 2007. Evaluation of machine learningmethods for ligand based virtual screening. J Comput-Aided MolDes 21:53–62.

**[29].** Deshpande M, Kuramochi M, Wale N, Karypis G. 2005. Frequentsubstructure based approaches for classifying chemical compounds. IEEE TKDE 17:1036–1050.

**[30].** Rashid, M. T., & Wang, D. (2021, October). Unravel: An anomalistic crowd investigation framework using social airborne sensing. In 2021 IEEE International Performance, Computing, and Communications Conference (IPCCC) (pp. 1-10). IEEE.

**[31].** Devore J, Peck R. 2004. Statistics: the exploration and analysis ofdata, 5th ed. Belmont, CA: Duxbury Press.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, January 2024

**[32].** Dix DJ, Houck KA, Martin MT, Richard AM, Woodrow Setzer R,Kavlock RJ. 2007. The toxcast program for prioritizing toxicitytesting of environmental chemicals. Toxicol Sci 95:5–12.

**[33].** Eom JH, Zhang BT. 2004. Pubminer: Machine learning based textmining system for biomedical information mining.

**[34].** Feldman HJ, Snyder KA, Ticoll A, Pintilie G, Hogue CWV. 2006.A complete small molecule dataset from the protein data bank.FEBS Lett 580:1649–1165.

**[35].** Geppert H, Horvath T, Gartner T, Wrobel S, Baorath J. 2008.Support vector machine based ranking significantly improves theeffectiveness of similarity searching using 2d fingerprints andmultiple reference compounds. J Chem Info Model 48:742–746.

**[36].** Guo J, Chen H, Sun Z, Lin Y. 2004. A novel method for proteinsecondary structure prediction using dual layer svm and profiles.Proteins 54:738–743.

**[37].** Helma C, Cramer T, Kramer S, Raedt LD. 2004. Data miningand machine learning techniques for the identification ofmutagenicity inducing substructures and structure activity relationships of noncongeneric compounds. J Chem Info Comp Sci44:1402–1411.

**[38].** Rashid, M. T., Zhang, D., & Wang, D. (2019, August). Edgestore: Towards an edge-based distributed storage system for emergency response. In 2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS) (pp. 2543-2550). IEEE.

**[39].** Hert J, Willet P, Wilton D. 2006. New methods for ligand basedvirtual screening: use of data fusion and machine learning toenhance the effectiveness of similarity searching. J Chem InfoModel 46:462–470.

**[40].** Hopkins AL. 2008. Network pharmacology: the next paradigm indrug discovery. Nat Chem Biol 4:682–690.

**[41].** Hopkins AL. 2009. Drug discovery: predicting promiscuity. Nature462:167–168.

**[42].** Jacob L, Vert JP. 2007. Kernel methods for in silico chemogenomics.In: Proceedings of the NIPS Workshop on Machine Learning inComputational Biology; Vancouver, Canada: Curran Associates,

**[43].** Jenkins JL, Bender A, Davies JW. 2006. In silico target fishing:Predicting biological targets from chemical structure. DrugDiscov Today 3:413–421.

**[44].** Jensen D, Neville J. 2002. Data mining in social networks. In:National Academy of Sciences Symposium on Dynamic SocialNetwork Modeling and Analysis; November 7–9, 2002. Washington, DC: The National Academies of Sciences.

**[45].** Rashid, M. T., Abir, I. K., Shourove, N. S., Muntaha, R., & Rhaman, M. K. (2016, May). Intelligent intrusion prevention system for households based on system-on-chip computer. In 2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE) (pp. 1-5). IEEE.

**[46].** Joachims T. 1998. Text categorization with support vector machines:Learning with many relevant features. In: Proc. of the EuropeanConference on Machine Learning. Chemnitz, Germany: Springer.p 137–142.

**[47].** Karypis G. 2006. Yasspp: better kernels and coding schemes lead toimprovements in protein secondary structure prediction. Proteins64:575–586.

**[48].** Kola I, Landis J. 2004. Can the pharmaceutical industry reduceattrition rates? Nature Rev Drug Discov 3:711–716.

**[49].** Yasar, M. S., & Rashid, M. (2015). Implementation of dynamic traffic light controllers using artificial neural networks to diminish traffic ordeals. In IEEE European Modelling Symposium.

**[50].** Kosala R. 2000. Web mining research: a survey. SIGKDD Explor2:1–15.

**[51].** Kubinyi H. 2006. Chemogenomics in drug discovery. Ernst ScheringRes Found Workshop 58:1–19.

**[52].** Lanckriet GR, Deng M, Cristianini N, Jordan MJ, Noble WS. 2004.Kernel based data fusion and its application to protein functionprediction in yeast. Proceedings, January 6–10, 2004, Hawaii.World Scientific: Pac Symp Biocomput. p 300–311.

**[53].** Menchetti S, Costa F, Frasconi P. 2005. Weighted decompositionkernels. Proceedings of the 22nd International Conference inMachine Learning; August 7–11, 2005; Bonn, Germany. NewYork: ACM 119:585–592.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, January 2024

**[54].** Michielan L, Stephanie F, Terfloth L, Hristozov D, Cacciari B, Klotz K, Spalluto G, Gasteiger J, Moro S. 2009. Exploring potency Mitchell TM. 1997.

**[55].** Muegge I, Oloff S. 2006. Advances in virtual screening. Drug DiscovToday 3:405–411.Palmer DS, O'Boyle NM, Glen RC, Mitchell JBO. 2007. Randomforest models to predict aqueous solubility. J Chem Info Model47:150–158.

**[56].** Paolini GV, Shapland RH, Van Hoorn WP, Mason JS, Hopkins AL.2006. Global mapping of pharmacological space. Nature Biotechnology 24:805–815.

**[57].** Ralaivola L, Swamidass SJ, Saigo H, Baldi P. 2005. Graph kernels forchemical informatics. Neural Netwk 18:1093– 1110.

**[58].** Rangwala H, Karypis G. 2006. Building multiclass classifiers forremote homology detection and fold recognition. BMC Bioinformatics 7:455.

**[59].** Rangwala H, Karypis G. 2007. frmsdpred: predicting local rmsdbetween structural fragments using sequence information.Comput Syst Bioinformat Conf 6:311–322.

**[60].** Rangwala H, Karypis G. 2008. frmsdalign: frmsdalign: Proteinsequence alignment using predicted local structure information.In: Proceedings of the 6th Asia Pacific Bioinformatics Conference;January 17–19, 2008. London: Imperial College.

**[61].** Rangwala H, Kauffman C, Karypis G. 2007. A generalized framework for protein sequence annotation. In: Proceedings of the NIPS Workshop on Machine Learning in Computational Biology; December 10, 2008. Vancouver, Canada: CurranAssociates.

**[62].** Raymond JW, Cardiner EJ, Willet P. 2002. Heuristic for similaritysearching of chemical graphs using a maximum common edge subgraph algorithm. J Chem Info Comp Sci 42:305–316.

**[63].** Rognan D. 2007. Chemogenomic approaches to rational drugdesign. Br J Pharmacol 152:38–52.

**[64].** Root DE, Kelley BP, Stockwell BR. 2002. Global analysis oflargescale chemical and biological experiments. Curr Opin DrugDiscov Dev 5:355–360.

**[65].** Russ AP, Lampel S. 2005. The druggable genome: an update. DrugDiscov Today 10:1607–1610.

**[66].** Sakiyama Y. 2009. The use of machine learning and nonlinearstatistical tools for adme prediction. Expert Opin Drug MetabToxicol 5:149–169.

**[67].** Salim N, Holliday JD, Willett P. 2003. Combination of fingerprintbased similarity coefficients using data fusion. J Chem InfoComput Sci 43:435–442.

**[68].** Schreiber SL. 1998. Chemical genetics resulting from a passion for synthetic organic chemistry. Bioorg Med Chem 6:1127–1152.

**[69].** Schroeter TS, Schwaighofer A, Mika S, Laak AT, Suelzle D, Ganzer U, Heinrich N, Muller KR. 2007. Estimating the domainof applicability for machine learning qsar rmodels: a study onaqueous solubility of drug discovery molecules. J Comput-Aided21:485–498.