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

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).

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

[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.

[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).

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

[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.

[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.

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

[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.

[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.