Review Article



Available online at: http://www.iajps.com

A COMPREHENSIVE REVIEW ON "ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND BIG DATA" IN PHARMACEUTICAL WORLD

Krishna Prasad Davarasingi^{1*}, Lakshmi Prameela Devi Katari²

^{1*}. Department of Pharmaceutical Quality Assurance, Shri Vishnu College of Pharmacy,

Vishnupur, Bhimavaram- 534202, Andhra Pradesh, India.

². Jr. Executive, Auropeptides Ltd, Hyderabad-502307, Telangana, India.

Abstract:

Artificial Intelligence(AI), Machine Learning(ML) and Big Data are facilitating the present society as front runner beneficiary, This review highlights the imactful use of AI, ML and Big Data in diverse areas of pharmaceutical fields. Drug Discovery and development with collaborative inputs of technology reducing the human workload as achieving the target in a short period. This paper surveys big data with highlighting the big data analytics. Big data analytics covers integration and analysis of large amount of complex heterogeneous data such as various – omics data (genomics, epigenomics, transcriptomics, proteomics, metabolomics, interactomics, pharmacogenomics, diseasomics), biomedical data and electronic health records data. We underline the challenging issues about big data privacy and security.

Keywords: Artificial Intelligence(AI), Machine Learning(ML) and Big Data, Technology, Big Data Analytics, Data Mining, Health Informatics, Healthcare Information

Corresponding author:

Krishna Prasad Davarasingi,

Department of Pharmaceutical Quality Assurance,

Shri Vishnu College of Pharmacy,

Vishnupur, Bhimavaram, 534202, Andhra Pradesh, India.

E-mail address: krishnaprasad2728@gmail.com

Tel: + 91 8121299369.

Please cite this article in press Krishna Prasad Davarasingi et al, A Comprehensive Review On "Artificial Intelligence, Machine Learning And Big Data" In Pharmaceutical World., Indo Am. J. P. Sci, 2023; 10 (05).



INTRODUCTION:

ARTIFICIAL INTELLEGENCE

Artificial Intelligence has recently started to gear up its application in various sectors of society with pharmaceutical field. Artificial Intelligence (AI) is a stream of science related to intelligent machine learning, mainly intelligent computer programs, which provides results in the similar way to human attention process. This process generally comprises obtaining data, developing efficient systems for the uses of obtained data, illustrating definite or conclusions approximate and selfcorrections/adjustments. In general, AI is used for analyzing the machine learning to imitate the cognitive tasks of individuals.2,3 AI technology is exercised to perform more accurate analyses as well as to attain useful interpretation. In this perspective, various useful statistical models as well as computational intelligence are combined in the AI technology. The progress and innovation of AI applications are often associated to the fear of unemployment threat. However, almost all advancements in the applications of AI technology are being celebrated on account of the confidence, which enormously contributes its efficacy to the industry.

THE ADOPTION AI INCLUDES

- It allows for learning from real time data
- Identify the right candidates for clinical trails
- Processing real time patient feed back
- Integrating data exchanges with partners
- Reducing costs
- Increasing productivity
- It perform analysis faster and more accurately
- It is capable of seeing patterns that even trained professional might miss

APPLICATIONS

- AI: networks and tools AI involves several method domains, such as reasoning, knowledge representation, solution search, and, among them, a fundamental paradigm of machine learning (ML). ML uses algorithms that can recognize patterns within a set of data that has been further classified. A subfield of the ML is deep learning (DL), which engages artificial neural networks (ANNs).
- AI in Clinical trail design Clinicaltrials are directedtoward establishing the safety and efficacy of a drug product in humans for a particular disease condition and require 6–7 years along with a substantial financial investment. However, only one out of ten

molecules entering these trials gain successful clearance, which is a massive loss for the industry . These failures can result from inappropriate patient selection, shortage of technical requirements, and poor infrastructure. However, with the vast digital medical data available, these failures can be reduced with the implementation of AI

- AI in Quality Control and Quality Assurance - Manufacturing of the desired product from the raw materials includes a balance of various parameters. Quality control tests on the products, as well as maintenance of batch-to-batch consistency, require manual interference. This might not be the best approach in each case, showcasing the need for AI implementation at this stage. The FDA amended the Current Good Manufacturing Practices (cGMP) by introducing a 'Quality by Design' approach to understand the critical operation and specific criteria that govern the final quality of the pharmaceutical product
- AI in Pharmaceutical Manufacturing Tools, such as CFD, uses Reynolds-Averaged NavierStokes solvers technology that studies the impact of agitation and stress levels in different equipment (e.g., stirred tanks), exploiting the automation of many pharmaceutical operations. Similar systems, such as direct numerical simulations and large eddy simulations, involve advanced approaches to solve complicated flow problems in manufacturing.
- AI in Advancing pharmaceutical product development – silico absorption, distribution metabolism, excretion(ADMET), 3D Imaging, Blockchain (type of data structure based collection)
- **Prediction of toxicity** The prediction of the toxicity of any drug molecule is vital to avoid toxic effects. Several web-based tools, such as LimTox, pkCSM, admetSAR, and Toxtree, are available to help reduce the cost. Advanced AI-based approaches look for similarities among compounds or project the toxicity of the compound based on input features. The Tox21 Data Challenge organized by the National Institutes of Health, Environmental Protection Agency (EPA), and US Food and Drug Administration (FDA) was an initiative to evaluate several computational techniques to forecast the toxicity of 12 707 environmental compounds and drugs
- **Prediction of the Physicochemical properties** Physicochemical properties, such as solubility,

partition coefficient (logP), degree of ionization, and intrinsic permeability of the drug, indirectly affect its pharmacokinetics properties and its target receptor family and, hence, must be considered when designing a new drug. Different AI-based tools can be used to predict physicochemical properties. For example, ML uses large data sets produced during compound optimization done previously to train the program.

- **Prediction of Bioactivity** Efficacy of the drug molecules depend on their affinity for the target protein or receptor. Drug target binding affinity (DTBA) is vital to predict drug-target interactions. AI-based methods can measure the binding affinity of a drug by considering either the features or similarities of the drug and its target. Web applications, such as ChemMapper and the similarity ensemble approach (SEA), are available for predicting drug-target interactions. Many strategies involving ML and DL have been used to determine DTBA, such as KronRLS, SimBoost, DeepDTA and PADME.
- **Drug Screening** Nearest-Neighbour classifiers, SVMs, Deep Neural Networks are used for drug screening because nine out of ten fail in trails, Bayer, Roche and Pfizer have teamed up for developing a platform for therapies

TOOLS OF AI

- i. Robot Pharmacy
- ii. MEDi Robot
- iii. Erica Robot
- iv. TUG Robot
- v. Automated Control Process System (ACPS)
- vi. Berg
- vii. Nano Robots

ADVANTAGES OF AI TECHNOLOGY

- i. **Error minimization:** AI assists to decrease the errors and increase the accuracy with more precision. Intelligent robots are made of resistant metal bodies and capable of tolerating the aggressive atmospheric space, therefore, they are sent to explore space.
- ii. **Difficult exploration:** AI exhibits its usefulness in the mining sector. It is also used in the fuel exploration sector. AI systems are capable of investigating the ocean by defeating the errors caused by humans.
- Daily application: AI is very useful for our daily acts and deeds. For examples, GPS system is broadly used in long drives. Installation of AI in Androids helps to predict

what an individual is going to type. It also helps in correction of spelling mistakes.

- iv. Digital assistants: Now-a-days, the advanced organizations are using AI systems like 'avatar' (models of digital assistants) for the reduction of human needs. The 'avatar' can follow the right logical decisions as these are totally emotionless. Human emotions and moods disturb the efficiency of judgement and this problem can be overcome by the uses of machine intelligence.
- v. **Repetitive tasks:** In general, human beings can perform single task at a time. In contrast to the human beings, machines are capable of performing multi-tasking jobs and can analyze more rapidly in comparison to the human beings. Various machine parameters, i.e., speed and time can be adjusted according to their requirements.
- vi. **Medical uses:** In general, the physicians can assess the condition of patients and analyze the adverse effects and other health risks associated with the medication with the help of AI program. Trainee surgeons can gather knowledge by the applications of AI programs like various artificial surgery simulators (for examples, gastrointestinal simulation, heart simulation, brain simulation, etc.
- vii. **No breaks:** Unlike human beings who have the capacity of working for 8 h/day with breaks, the machines are programmed in such a way that these are capable of performing the work in a continuous manner for long hours devoid of any kinds of confusions and boredom.
- viii. **Increase technological growth rate:** AI technology is widely used in most of the advanced technological innovations worldwide. It is capable of producing different computational modelling programs and aims for the invention of the newer molecules. AI technology is also being used in the development of drug delivery formulations.
- ix. **No risk:** In case of working at the risky zone like fire stations, there are huge chances of causing harm to the personnel engaged. For the machine learning programs, if some mishap happens then broken parts can be repairable. x) Acts as aids: AI technology has played a different function by serving children as well as elders on a 24x7 basis. It can perform as teaching and learning sources for all.
- x. **Limitless functions:** Machines are not restricted to any boundaries. The emotionless machines can do everything more efficiently

and, also produce more accurately than the human beings

DISADVANTAGES OF AI TECHNOLOGY

- i. **Expensive:** The launch of AI causes huge money consumption. Complex designing of machine, maintenance and repairing are highly cost effective. For the designing of one AI machine, a long period of time is required by the R&D division. AI machine needs updating the software programmes, regularly. The reinstallations as well as recovery of the machine consume longer time and huge money.
- ii. **No replicating humans:** Robots with the AI technology are associated with the power of thinking like human and being emotionless as these add some advantages to perform the given task more accurately without any judgement. If unfamiliar problems arise, robots

cannot take the decision and provide false report.

- iii. No improvement with experience: Human resource can be improved with experiences. In contrast, machines with AI technology cannot be enhanced with experience. They are unable to identify which individual is hard working and which one is nonworking.
- iv. **No original creativity:** Machines with AI technology have neither sensitivity nor the emotional intelligence. Humans have the ability to hear, see, feel and think. They can use their creativity as well as thoughts. These features are not achievable by the uses of machines.
- v. **Unemployment:** The widespread uses of AI technology in all the sectors may cause large scale unemployment. As because of the undesirable unemployment, human workers may lose their working habits and creativity.

Technique	Application	Method
Deep Learning	Drug Screening	Interaction between Protiens and
		corresponding ligand
Neural Networks	Drug Designing	Predict the structure of different
		proteins
Reinforcement Learning	Screening of Drug	Finding an inhibitor molecule for
		specific protien

Table 1: AI associated techniques for drug discovery

Pharmaceutical Organisation	AI Organisation	Collobarative Work
Roche	X OWKIN	Drug Discovery and Development, Clinical trails on ML network
Pfizer	XtalPi	3D structures of the molecules, depicting their mechanical and chemical properties and how they bind with protein receptors
U NOVARTIS	Microsoft	Cell and gene-based therapies, generative chemistry, image segmentation and analysis for smart and personalized delivery for therapies
THERAPEUTICS	P ointR [™] DATA	Treatment of late stage melanoma, gliomas and pancreatic cancer
AstraZeneca	Benevolent	Neural network based platforms for discovery and development of new treatments of chronic kidney disease and idiopathic pulmonary fibrosis
Lilly	Atomwise Better medicines faster.	Develop drugs on novel protein targets
Takeda	Numerate	Finding drug molecules for oncology, gastroenterology and central nervous system disorders
BAYER E R	Exscientia	Discover and develop bispecific small molecules for diabetes and its comorbidities
Boehringer Ingelheim	healx	Find therapies for rare neurological diseases

Table 2: Artificial Intelligence Collaborations With Pharmaceutical Organisation

MACHINE LAERNING:

Machine Learning (ML) approaches provide a set of tools that can improve discovery and decision making for well specified quesions with abundant and high quality data. Stages of drug discovery that includes with Machine Learning are Target Validation, Identification of progbnostic biomarkers and analysis of digital pathology data in clinical trails.

APPLICATIONS OF ML IN DRUG DISCOVERY

- 1. Target Identification and Validation -The pre-eminent approach in drug discovery is to develop drugs (small molecules, peptides, antibodies or newer modalities including short RNAs or cell therapies) that will alter the disease state by modulating of the activity of a molecular target
- 2. Small molecule design and optimization -The discovery of drug candidates that can block or activate the target protein of interest involves extensive virtual and experimental high-throughput screening of large compound libraries. Candidate structures are then further refined and modified to improve target specificity and selectivity, along with optimized pharmacodynamic, pharmacokinetic and toxicological properties.
- **3.** Computational pathology Pathology is a descriptive field, as a pathologist interprets what is seen on a glass slide by visual inspection. Analysis of these glass slides provides a vast amount of information, such as the type of cell present in the tissue and their spatial context. The interplay between tumour and immune cells within the tumour microenvironment is increasingly important in the study of immuno-oncology and is not captured by other technologies.
- 4. Disease Identification and Diagnosis – Personalised Treatment and Behavioral Modification - Somatix - a data-analytics B2B2C software platform company whose MLbased app uses "recognition of hand-to-mouth gestures in order to help people better understand their behavior and make lifeaffirming changes", specifically in smoking cessation. SkinVision - the self-described "skin cancer risk app" makes its claim as "the first and only CE certified online assessment." Interestingly, we couldn't find SkinVision in the app store. The first that app that came up under a "SkinVision" Search was DermCheck, in which images are submitted to dermatologists (people, not machines) by phone in exchange for a

personalized treatment plan—perhaps a testament to some of the kinks in machine learning-based accuracy at scale that still need to be ironed out.

- 5. Radiology and Radiotherapy hey might look more like cyborgs: supervising algorithms reading thousands of studies per minute." Until that day comes, Google's <u>DeepMind Health</u> is working with <u>University College London</u> <u>Hospital (UCLH)</u> to develop machine learning algorithms capable of detecting differences in healthy and cancerous tissues to help improve radiation treatments.
- 6. Smart Electronic Health Records Document classification (sorting patient queries via email, for example) using support vector machines, and optical character recognition (transforming cursive or other sketched handwriting digitized into characters), are both essential ML-based technologies in helping advance the collection and digitization of electronic health information. MATLAB's ML handwriting recognition technologies and Google's Cloud Vision API for optical character recognition. Artificial Neural Network using MATLAB -Handwritten Character Recognition
- Epidemic 7. Outbreak Prediction technologies are also being applied to monitoring and predicting epidemic outbreaks around the world, based on data collected from satellites, for example, to predict malaria outbreaks, taking into account data such as temperature, average monthly rainfall, total number of positive cases, and other data points. Predicting outbreak severity is particularly pressing in third-world countries, which often lack medical infrastructure, educational avenues, and access to treatments. ProMED-mail is an internet-based reporting program for monitoring emerging diseases and providing outbreak reports in real-time
- 8. **Predictive biomarkers** ML-based biomarker discovery and drug sensitivity predictive models are demonstrated approaches to help improve clinical success rates, to better understand the mechanism of action of a drug and to identify the right drug for the right patients56–58. Late-stage clinical trials take many years and millions of dollars to conduct, so it will be most beneficial to build, validate and apply predictive models earlier, using preclinical and/or early-stage clinical trail data.

9. **Bioactivity** – KRONECKER(KronRLS) least square - evaluate similarity between drug and protein molecule to determine DTBA

Simboost- used regression trees to predict DTBA

Techniques	Application	Method
Traditional Reinforcement Learning	New Drug Development	Create new molecules
Transfer Learning	Emulating Biological Process	Model response to anticancer
		medication
Multitask Learning	Drug Development and Testing	Monitor signals between pathways
Multitask Analysis	Drug Target Interaction and Post	Analyse Interaction between drug
	manufacturing drug reviews	targets

 Table 3: Techniques associated with machine learning

BIG DATA

Big Data represents large amounts of data that is unmanageable using traditional software or internet based platforms. It traditionally used amount of storage, processing and analytical power.Big data is available for the classification and understanding of disease process followed by target identification and lead compound discovery, charectarize all levels of molecular changes. This Concept has its worth in Health Claims, Clinical Aspects, Sales and Dispensing, Clinical research-Safety Pharmacovigilence, Patient generated data etc., Big data has the following charecterstics representing the 6Vs i.e., Volume, Variety, Velocity, Veracity, Value, Varieability.

APPLICATIONS:

- _Electronic health records
- Clinical Trail Applications Registries, Mining Registry Data, Data Quality and trasnsparency
- Data integration
- Digitalization of healthcare and big data
- Big data in biomedical research
- Big data from omics studies

- Internet of Things (IOT)
- Mobile Computing and mobile health (mhealth)
- Information Extraction, Chemoinformatics, Bioinformatics, Pharmacovigilance
- Genomics and Genetics Pharmacogenomics
- Drug Bank
- Predictions Molecular Properties, Molecular Similarities
- Polypharmacology
- Multiparameter Optimisation, Synthesis Planning

CHALLENGES ASSOCIATED WITH BIGDATA

• Storage

- Cleaning
- Unified format
- Accuracy
- Image pre-processing
- Security
- Meta data
- Querying
- Visualization
- Data Sharing

	Company	Description
MedeAnalytics	Mede/Analytics [*]	Provides performance management solutions health systems and plans and health analytics
IBM Watson	IBM Watson	Services on sharing clinical and health related data among hospital and researchers
Roam Analytics	ROÅM	Natural Language processing infrastructure for modern healthcare systems
Optum Health		Healthcare analytics and comprehensive and innovative solutions for healthcare industry
<u>Lumiata</u>	lumiata	Services for analytics and risk management for efficient noutcomes in healthcare
<u>Apixio</u>		Cognitive computing platform for analyzing clinical data
Linguamatics	Linguamatics	Text mining platform for digging important information for unstructured healthcare data
Health Fidelity		Management solutions for risk assessment in workflows of healthcare organization and methods for optimization and adjustments.
Flatiron Health	flatiron.	Applications for organizing and improving oncology data for better cancer treatment
Enlitic	© enlitic	Deep learning using large scale data sets from clinical tests for health diagnosis

Table4: Big companies which provides services on big data analysis

CONCLUSION:

Artificial intelligence and the machine learning algorithms. It is clearly described how the artificial intelligence can be used for saving the time as well as the resources. The retail pharmacy must have to adopt this for the better performance. Instead of giving the known medicine, the pharmacist can provide the effective medicine to the patient with the help of machine learning prediction algorithms. These algorithms can be used for predicting the disease of a patient and can also be used for predicting the medicine for that disease. The future work can be the implementation of these machine learning algorithms so that the correct medicine can be provided to the patients. This will save the time and will provide the more accuracy.

CONFLICT OF INTEREST:

The authors have no conflicts of interest regarding this investigation.

REFERENCES:

 Cai, Y., Dai, D., & Hua, S. (2016). Using machine learning algorithms to improve the prediction accuracy in disease identification: An empirical exampleRetrieved from https://search.proquest.com/docview/180642900 9?accountid=35493 Cassel, C. K. J. J. (2012). Retail clinics and drugstore medicine. 307(20), 2151-2152. Chen, M., Hao, Y., Hwang, K., Wang, L., & Wang, L. J. I. A. (2017). Disease prediction by machine learning over big data from healthcare communities. 5, 8869-8879.

- 2. Donepudi, P. K. (2017b). Machine Learning and Artificial Intelligence in Banking. Engineering International, 5(2), 83-86. https://doi.org/10.18034/ei.v5i2.490
- Iyawa, G. E., Herselman, M., & Botha, A. (2017). A scoping review of digital health innovation ecosystems in developed and developing countries. Piscataway: The Institute of Electrical and Electronics Engineers, Inc. (IEEE). Retrieved from https://search.proquest.com/docview/196231666
 <u>4?accountid=35493</u>
- Michalski, R.S., Carbonell, J.G., Mitchell, T.M. (1983). Machine Learning: An Artificial Intelligence Approach. Springer, <u>https://www.springer.com/gp/book/97836621240</u> 79
- Vyas, M., Thakur, S., Riyaz, B., Bansal, K. K., Tomar, B., & Mishra, V. J. A. J. P. (2018). Artificial intelligence: the beginning of a new era in pharmacy profession. 12(2), 72-76.
- 6. 1] Yang C, Li C, Wang Q, Chung D, Zhao H. Implications of pleiotropy: challenges and opportunities for mining big data in biomedicine. Front Genet 2015;6:229.
- Viceconti M, Hunter P, Hose R. Big data, big knowledge: big data for personalized healthcare. IEEE J Biomed Health Inform 2015;19:1209– 15.
- Kankanhalli A, Hahn J, Tan S, Gao G. Big data and analytics in healthcare: introduction to the special section. Inform Syst Front 2016;18:233– 5.
- 9. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inform Sci Syst 2014;2:3.
- Wu PY, Cheng CW, Kaddi CD, Venugopalan J, Hoffman R, Wang MD. –Omic and Electronic Health Record Big Data Analytics for Precision Medicine. IEEE Trans Biomed Eng 2017;64:263–73.
- 11. Wang Y, Kung LA, Wang WY, Cegielski CG. An integrated big data analytics-enabled transformation model: application to health care. Inf Manag 2017;55:64–79.
- El-Gayar O, Timsina P. Opportunities for business intelligence and big data analytics in evidence based medicine. In: System Sciences (HICSS), 2014 47th Hawaii international conference on 2014:749–57.
- 13. Gu D, Li J, Li X, Liang C. Visualizing the knowledge structure and evolution of big data

research in healthcare informatics. Int J Med Inform 2017;98:22–32.

- Gligorijević V, Malod-Dognin N, Pržulj N. Integrative methods for analyzing big data in precision medicine. Proteomics 2016;16:741–58.
- 15. Luo J, Wu M, Gopukumar D, Zhao Y. Big data application in biomedical research and health care: a literature review. Biomed Inform Insights 2016;8:1
- Mak KK, Pichika MR. Artificial intelligence in drug development: Present status and future prospects. Drug Discov Today. 2019;24(3):773-80.
- Hassanzadeh P, Atyabi F, Dinarvand R. The significance of artificial intelligence in drug delivery system design. Adv Drug Deliv Rev. 2019;151:169-90.
- Russel S, Dewey D, Tegmark M. Research priorities for robust and beneficial artificial intelligence. AI Mag. 2015;36(4):105-14.
- 19. Duch W, Setiono R, Zurada JM. Computational intelligence methods for rulebased data understanding. Proc IEEE. 2004;92(5):771-805.
- 20. Dasta JF. Application of artificial intelligence to pharmacy and medicine. Hosp Pharm. 1992;27(4):319-22.
- 21. Jiang F, Jiang Y, Zhi H. Artificial intelligence in healthcare: Past, present and future. Stroke Vasc Neurol. 2017;2(4):230-43.
- 22. Gobburu JV, Chen EP. Artificial neural networks as a novel approach to integrated pharmacokinetic-pharmacodynamic analysis. J Pharm Sci. 1996;85(5):505-10.
- 23. Mnih V, Kavukcuoglu K, Silver D et al. Humanlevel control through deep reinforcement learning. Nature 518(7540), 529–533 (2015).
- 24. Butina D. Unsupervised database clustering based on daylight's fingerprint and tanimoto similarity: a fast and automated way to cluster small and large datasets. J. Chem. Inf. Comput. Sci. 39(4), 747–750 (1999).
- 25. Brown N. In Silico Medicinal Chemistry: Computational Methods to Support Drug Design. Royal Society of Chemistry, Cambridge, UK (2016)
- 26. Firth NC, Atrash B, Brown N, Blagg J. MOARF, an integrated workflow for multiobjective optimization: implementation, synthesis, and biological evaluation. J. Chem. Inf. Model. 55(6), 1169–1180 (2015)
- 1 Ramesh, A. et al. (2004) Artificial intelligence in medicine. Ann. R. Coll. Surg. Engl. 86, 334– 338
- Miles, J. and Walker, A. (2006) The potential application of artificial intelligence in transport. IEE Proc.-Intell. Transport Syst. 153, 183–198

- 29. Yang, Y. and Siau, K. (2018) A Qualitative Research on Marketing and Sales in the Artificial Intelligence Age. MWAIS 4 Wirtz, B.W. et al. (2019) Artificial intelligence and the public sector—applications and challenges. Int. J. Public Adm. 42, 596–615
- 30. Smith, R.G. and Farquhar, A. (2000) The road ahead for knowledge management: an AI perspective. AI Mag. 21 17–17
- 31. Lamberti, M.J. et al. (2019) A study on the application and use of artificial intelligence to support drug development. Clin. Ther. 41, 1414–1426
- Beneke, F. and Mackenrodt, M.-O. (2019) Artificial intelligence and collusion. IIC Int. Rev. Intellectual Property Competition Law 50, 109– 134