

Exploring the Consequences of Implementing Machine Learning in the Pharmaceutical Field

A. Ramesh¹ and Dr. Pawan Kumar²

Research Scholar, Department of Computer Science & Engineering¹

Research Guide, Department of Computer Science & Engineering²

NIILM University, Kaithal, India

Abstract: *The pharmaceutical and consumer health industries are greatly impacted by artificial intelligence and machine learning. These technologies are crucial for patient identification because of their improved intelligence applications, which include disease detection and diagnostics for clinical testing, pharmaceutical manufacturing, and predictive forecasting. Recent developments in a wide range of analytical tools and machine learning algorithms have opened up new possibilities for novel machine learning applications in many pharmaceutical science disciplines. This research examines the past, present, and future effects of machine learning on several sectors, including medicine development and design. Artificial neural networks are employed in pharmaceutical machine learning because they are capable of simulating the nonlinear interactions that are often seen in pharmaceutical research. Research is being done on artificial intelligence (AI) and learning machines for use in everyday medicine, industry, and regulation.*

Keywords: Machine Learning, Pharmaceutical Industry, Drug Discovery

I. INTRODUCTION

AI might change the pharmaceutical industry's business environment. Big data and AI-driven analytics have changed pharmaceutical innovation [1].

Machine learning may improve value chain efficiency and results and spur innovation (Nagy et al., 2019). Pharmaceutical firms may boost their value by encouraging innovation and new business methods [2]. Drug managers are looking for ways to apply AI and machine learning in biotechnology and health [3]. Many organizations are embracing current applications, paving the road for the industry's digital future in technology, studies say. AI suppliers interact with top pharmaceutical firms on medication discovery, development, and research [4].

Reports show that 72% of healthcare firms believe AI is essential to their future business strategies and 62% wish to invest in it soon [4]. Pharma News Intelligence examines the best AI and machine learning applications and their future to understand AI's future in the sector.

Data collection, particularly in the pharmaceutical business, has increased daily. Many disciplines of research are using "big data" increasingly often. Data-driven enterprises are also demonstrating how big data may benefit many industries. There are numerous definitions of "big data." "4 Vs." is a popular definition. Douglas Laney's [5] definition incorporates the "3 Vs"—volume, velocity, and variety. IBM later added the fourth "V." to this definition [6]. However, there is no quantitative, universally acknowledged definition of "big data."

It was called "new oil" because of its data value [7]. Social media, user content, genetics, textbooks and publishing, electronic health, and sensor networks affect "big data" riches and diversity. Data volume has skyrocketed due to technological and storage advances [8]. Nearly 2.5 million scientific publications are published annually [9]. The "pharmaceutical industry" domain registered almost 15,000 articles in 2019. (MedSource).

Thus, in the pharmaceutical industry, "big data" may be good and bad. Machine learning allows computers to "learn" and perform activities, increasing big data application in the pharmaceutical industry. Machine learning is the most frequently employed AI technique in the pharmaceutical industry, therefore its breadth is unique. Natural language processing (NLP), expert systems, and robotics are being used more in illness detection, patient monitoring, and robotic surgery [10]. Compared to machine learning, these technologies have received less pharmaceutical industry interest.

This article discusses several pharmaceutical industry-wide machine learning approaches.

Machine Learning and Pharmaceutical Industry

The notion of mainstream AI is still unknown. The basic goal of AI is to computerize human intelligence. Warren McCulloch and Walter Pitts created an artificial neuron computer model in 1943. Like human neurons, artificial neurons activate or not [11]. John McCarthy invented "A.I." at the 1956 Dartmouth Conference [12]. AI has fluctuated since then [13]. AI is being used in health care, engineering, and transportation [14–16]. Big data in healthcare and rapid analytical tool development drove this focus on AI applications [10].

Supervised and unsupervised machine learning employ generalisations from supervised learning. an input-output dataset. Goals are reliable output data. Finally, the machine learning model predicts victory [17].

Table 1. Parametric model vs non-parametric model

Parametric Model	Non-parametric Model
<p>Definition A learning model summarizing data with a set of fixed size parameters is called a parametric model (depending on the number of training examples). No matter how often data we throw on a parametric model, how many variables it needs does not change its mind [22].</p>	<p>Definition Nonparametric techniques are suitable if we have a large amount of data and no prior information, and we do not want to be concerned with finding the exact right functionality [22].</p>
<p>Popular Algorithms Logistic Regression- It is possible to predict a data value using logistic regression, which is a statistical analytic approach based on previous observations of a data collection. To predict a dependent data variable, a logistic regression model examines the relationship between one or more already-existing independent variables [23]. Linear Discriminant Analysis- Dimensionality reduction is frequently used for supervised classification problems, and it is also known as linear discriminant analysis (LDA), normal discriminant analysis (NDA), or discriminant function analysis (DFA). In this context, it refers to differences between groups, such as those that exist between two or more classes. It's a technique for bringing features from a higher-dimensional realm down to a more mundane (geeksforgeeks.org). Naive Bayes- The Naive Bayes method, unlike the others on this list, is based on the Bayes theorem, which takes a probabilistic approach. The algorithm does not go immediately into the data since prior probabilities have been specified for each of your target's classes. The programme creates the posterior probability by updating these prior probabilities depending on the new information we have provided (edureka.co ,2021).</p>	<p>Popular Algorithms k-Nearest Neighbors- KNN is a non- parametric learning algorithm since it makes no assumptions about the underlying data distribution. The "training" phase of lazy learning is brief. Its purpose is to use a large number of data points classified into several groups to predict the categorization of a new sample point. Decision Trees- On the Decision Tree, each node represents an attribute test, and each branch reflects the outcome of that test. This is how the Decision Tree can be summarised. The leaf nodes include the predicted labels. The comparison of attribute values continues until we reach a leaf node, at which point we return to the tree's root (edureka.co ,2021). Support Vector Machines- An SVM is the only one of its kind since it seeks to sort the data with the widest possible margins between two groups. Max margin separation is what it's called. A final thing to bear in mind is that SVMs only employ the support vectors to plot the hyper plane, as opposed to linear regression, which makes use of the entire dataset (edureka.co ,2021).</p>

<p>Benefits Simpler: it is easier to understand and interpret these methods. Speed: Data can be used to learn parametric models reasonably quickly. Less data: We don't require as much learning and we can work well even if we don't fit the data precisely.</p>	<p>Benefits Flexibility: fit many useful forms. Power: no following function assumptions (or weak assumptions). Performance: Can lead to higher performance prediction models.</p>
<p>Limitations Constrained: these procedures are extremely limited to the stated form when selecting a functional form. Complexity limited: Methods are better adapted for simpler issues. Low fit: the approaches will probably not match the fundamental transformation matrix in practice.</p>	<p>Limitations More data: Requires a lot more training data to estimate the mapping function. Slower: Far slower to train, because they often have far more parameters for the train. Overfitting: It is tough to describe more risks of overfitting the training data and why certain predictions are made.</p>

Regression, SVMs, random forests, and ANNs are unsupervised machine learning methods. MCA (Multiple correspondence analysis) extracts functions without examples in unattended learning [17]. (PCA). ANNs and SVMs support supervised models (Lo et al., 2018). PCA is an unsupervised dimensional reduction algorithm that finds smaller-dimensional axes in unlabeled data [12].

Pharmaceutical researchers investigate machine learning using fluffy logical algorithms. Logical expressions represent fuzzy set membership [12]. It requires less system skill, data noise consideration, and predictability [18]. Fuzzy gene expression prediction models [18]. GA optimises populations. Pharma researchers employ GA to choose QSAR [19,20]. Innovating pharmaceutical machine learning methods like light gradient boosting. This machine learning method has several benefits. Modern AI uses transfer learning. A model is retrained to accomplish a new objective [21]. A large original model dataset improves transfer learning. Machine learning models are parametric or non-parametric. Table 1 compares parametric and non-parametric models.

ANN and Deep Learning

ANNs are biologically inspired computer models that learn by doing. Each brain has thousands of neurons, or processing units. Synaptic connections provide total neuron communication [24]. Information is transmitted between cells via axons. Neurons are biological cells [24].

Weights or coefficients link artificial and human neurons in an ANN [24]. ANNs typically have three structural parts. visible and concealed input layers Artificial neuron dendrites mimic actual ones. Hidden layer between input and output. Hidden layer (weights) links two layers. Every layer has thousands of nodes, or neurons. ANNs' supervised layer neuron count is frequently determined via test-and-ander. [25]. Over-storage or over-composition of training data owing to hidden layer neuron imbalance may reduce ANN generalization.

Deep Learning (DL) is a representation learning machine [26]. Recently developed neural networks are employed in DL. DL has three or more hidden layers and nodes per layer. Thus, DL uses several representation layers to learn complex functions. Deep learning's applicability is limited by big training sets. Other sources have examined fully connected feed forward networks (FCFF), RNNs, and CNNs [27]. DL was recognized for medication discoveries [29] and medicine formulations [28] and was interested in pharmaceutical research [29]. Their prediction and overall performance are superior than SVMs and RF master learning [28].

Applications of AI and ML in Pharmaceutical Industry

In the pharmaceutical sector, artificial intelligence (AI) has several uses, from medication research to marketing. Pharma businesses may run more smoothly, economically, and effectively by incorporating AI technology into essential processes. The best part is that artificial intelligence (AI) systems are an effective instrument for pharmaceutical research and development since they are made to improve outcomes as they gain knowledge from fresh information and experience.

The following are a few notable uses of AI in the pharmaceutical industry:

R & D

Pharma corporations deploy powerful AI and ML systems to speed up drug development. These artificial intelligence tools are intended to find complex patterns in large amounts of data in order to address biological network problems.

This is a great way to look at patterns of sickness and find out which disorders react well to certain treatments. Nowadays, pharmaceutical companies could focus on developing medications that have the greatest potential to treat a certain disease or medical condition.

Drug development

AI might improve science and innovation. AI is able to validate and identify certain medicines in addition to discovering novel compounds.

Only 13.8% of drugs make it past clinical trials, according to data from MIT. A pharmaceutical company must approve a medicine when a clinical research is completed, which may cost anywhere from \$161 million to \$2 billion. Pharmaceutical companies are using AI more and more to save operating costs, increase the success rate of new pharmaceuticals, and provide more reasonably priced ad treatments.

Diagnosis

They are able to collect, process, and analyze massive amounts of patient data. Global healthcare providers use machine learning (ML) technology to secure cloud patient data or a central storage system. These are EMRs, or electronic medical records.

These details may help medical professionals comprehend how a medication or a genetic trait impacts a patient's health. ML algorithms have the potential to foresee diagnosis and treatment in real time using EMR data. Machine learning's ability to quickly collect and analyze enormous amounts of data has the potential to save millions of lives.

Disease prevention

Pharmaceutical businesses may use AI to cure common and unusual diseases like Alzheimer's and Parkinson's. Therapy for rare diseases has a poor return on investment (ROI) when compared to the time and resources needed to develop drugs for uncommon ailments.

For over 95% of uncommon illnesses, there are no FDA-approved therapies. But the creative powers of AI and ML are rapidly changing the game.

Epidemic forecasting

Healthcare providers and organizations utilize AI and ML extensively to track worldwide diseases. These technologies use data from the internet to examine the relationships between biological, environmental, and geological elements and public health in various geographic places. Less developed countries do not have the infrastructure for healthcare and the financial resources needed to contain epidemics. An example of an AI application is the ML-based outbreak prediction model, which helps health care providers recognize and contain possible malaria epidemics.

Remote Monitoring

Good news for healthcare and pharma. Companies have created AI devices to remotely monitor critical patients. Tencent Holdings and Medopad's AI technology decreases engine function evaluation time to 30 minutes and provides remote Parkinson's monitoring. AI and smartphone apps can remotely open and shut patients' hands.

Smartphone cameras capture Parkinson's disease symptoms and hand motions. Doctors may remotely modify medications by estimating severity using a patient's frequency and amplitude of movements. The doctor will be alerted if things worsen. Remote technology may reduce patient travel and doctor's office wait times.

Manufacturing

AI may help pharmaceutical companies produce drugs more quickly, efficiently, and productively. Any part of the industrial process may be improved and controlled using AI.

- Quality control
- Predictive upkeep
- Waste reduction
- Design optimization
- Process automation

Pharmaceutical businesses may use AI to market their products more quickly and cheaply than they can with conventional methods. AI also lowers human error by limiting human involvement in the industrial process, all the while increasing return on investment.

Marketing

AI may be used in pharmaceutical sales and marketing. Pharmaceutical firms may enhance revenue and brand awareness using AI-based marketing.

Marketing technology (leading transformation) converts website visitors into consumers. AI can help organize clients. This lets pharmaceutical businesses focus on revenue- and conversion-boosting marketing.

Past outcomes may help AI determine the best cost-effective marketing strategies. Creating current marketing tactics saves time and money. in addition to marketing campaign predictions.

AI is fast altering the pharmaceutical sector, but there are difficulties. Most pharmaceutical companies' IT infrastructure is outdated and AI-unready. Industrial capability for AI adoption and integration is lacking. This may help the pharmaceutical industry use AI: Academic AI R&D Collaboration to Support AI Pharmaceutical Businesses AI-powered discovery firms provide expert advice, powerful tools, and industry expertise. Inform manufacturing and R&D teams on AI tools and technology.

The Current Impact of Pharmaceutical Industry

The vast variety of applications in this study will help the pharmaceutical industry. From idea to commercialization, AI may be employed in drug development.

AI may spend heavily to speed up and optimize this process, with a 10-year average of \$2.6 billion [30]. Pharmaceutical companies and startups have used AI at a rapid rate in the last decade. Pfizer and Novartis teamed or bought IBM Watson AI [31]. Mak & Pichika [32] covered AI and pharmaceutical businesses, including medicine discovery and reuse. Pharmaceutical companies are studying automation, robotics, and AI marketing [31]. High, complex data availability and advanced algorithms support industrial efficiency and speedy data digitalization. AI can help us make better judgments and provide better cures. Insufficient experience, switching to alternative scientific methods, and lack of funding were Henstock's problems [31]. Author suggests internal data management and AI skills to solve these difficulties [31].

Mary and her colleagues surveyed pharma and biotech organizations in 2019 to investigate AI's effect and adoption. In 217 organizations, AI is used for patient selection and recruitment and medication data collection. Lack of qualified workers, safety, regulation, and financial limits prevent utilization [4].

Challenges

The pharmaceutical industry has many obstacles when it comes to machine learning. Here are few instances:

- To construct onerous constraints on medicine improvement, a clear technique is needed. Machine deduction requires the application of causal reasoning.
- Data scientists are essential to the pharmaceutical industry. Establishing a strong talent pipeline is essential.
- Data governance is a hot subject at the moment. Besides, this is hardly a legitimate way to get access to medical data, which are confidential. This suggests that people take their privacy rights for granted.
- The pharmaceutical industry has generally been hesitant to fund or modify research, despite the fact that doing so directly benefits the company financially [33].

II. CONCLUSION AND FUTURE ASPECTS

Research on artificial neural networks (ANNs) in pharmaceuticals seems promising. This happens in gene delivery and medicine. Peptides (CPPs) carry drugs into cells. Recently, ANNs were used to study CPP efficacy. This model provided 13 noteworthy predictions and evaluations [34]. Another advantage of these technology is drug reuse [35].

The COVID-19 epidemic continues, but the first AI drug is not yet available. IA is needed to create anti-COVID-19 drugs. Benevolent AI [36] discovered COVID-19 medicines in clinical trials using engine learning

Recent studies show that most multinational healthcare firms will adopt AI by 2025. Cancer and chronic sickness will get greater money. AI is being used to improve chronic disease care and cut costs, yet most Americans die from them. Management of diabetes, renal disease, cancer, and IPF is planned. So AI will affect medical progress. AI helps find the best test candidates and analyze patients rapidly [37–40].

Reducing clinical research barriers eliminates the need for a large test group. AI aids patient screening and diagnosis. Data-driven AI can extract MRI pictures and mammograms. AI and machine learning will help design drugs. The pharmaceutical and industrial industries are also using AI.

Digitalizing pharmaceutical research may be the finest application of AI and machine learning. Machine learning is being rapidly used in pharmaceutical firms, proving its adaptability. In fact, dataset type and size may affect learning. This makes it task-specific. With enough data, high-value AI applications may become mainstream. Cost-effective digital pharmaceutical research is expected [41–44].

REFERENCES

- [1]. Gil Y, Selman B. A 20-year community roadmap for artificial intelligence research in the US; 2019. arXiv preprint arXiv:1908.02624
- [2]. Reymen I, Berends H, Oudehand R, Stultiëns R. Decision making for business model development: a process study of effectuation and causation in new technology-based ventures. *R&D Management*. 2017;47(4):595-606.
- [3]. Shah P, Kendall F, Khozin S, Goosen R, Hu J, Laramie J, Schork N. Artificial intelligence and machine learning in clinical development: a translational perspective. *NPJ digital medicine*. 2019;2(1):1-5.
- [4]. Lamberti MJ, Wilkinson M, Donzanti BA, Wohlhieter GE, Parikh S, Wilkins RG, Getz K. A study on the application and use of artificial intelligence to support drug development. *Clinical therapeutics*. 2019;41(8):1414-1426.
- [5]. Chen M, Mao S, Liu Y. Big data: A survey. *Mobile Netw Appl*. 2014;19(2):171–209.
- [6]. Heidari S, Alborzi M, Radfar R, Afsharkazemi MA, Rajabzadeh GA. Big data clustering with varied density based on MapReduce. *J Big Data*. 2019;6(1):77.
- [7]. Javornik M, Nadoh N, Lange D. Data is the new oil. *Towards user-centric transport in Europe*: Springer. 2019;295–308.
- [8]. Hashem IAT, Yaqoob I, Anuar NB, Mokhtar S, Gani A, Khan SU. The rise of “big data” on cloud computing: review and open research issues. *Inf Syst*. 2015; 47:98–115.
- [9]. Ware M, Mabe M. *The STM report: An overview of scientific and scholarly journal publishing*; 2015.
- [10]. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol*. 2017;2(4):230–43.
- [11]. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophysics*. 1943;5(4):115–33.
- [12]. Russell SJ, Norvig P. *Artificial intelligence: A modern approach*. 3rd ed: Pearson Education Limited; 2016.
- [13]. Fast E, Horvitz E. Long-term trends in the public perception of artificial intelligence. *Thirty-First AAAI Conference on Artificial Intelligence*; 2017.
- [14]. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. *NatMed*. 2019;25(1):24–9.
- [15]. Frank MR, Wang D, Cebrian M, Rahwan I. The evolution of citation graphs in artificial intelligence research. *Nat Mach Intell*. 2019;1(2):79–85.
- [16]. Zhao J, Liang B, Chen Q. The key technology toward the self-driving car. *Int J Intell Unmanned Syst*. 2018;6(1):2–20.
- [17]. Marsland S *Machine learning: An algorithmic perspective*. 2nd ed: CRC Press; 2015.
- [18]. Woolf PJ, Wang Y. A fuzzy logic approach to analyzing gene expression data. *Physiol Genomics*. 2000;3(1):9–15.
- [19]. Serra A, Önlü S, Festa P, Fortino V, Greco D. MaNGA: A novel multi-niche multi-objective genetic algorithm for QSAR modelling. *Bioinformatics*. 2020;36(1):145– 53.

- [20]. De P, Bhattacharyya D, Roy K. Exploration of nitroimidazoles as radiosensitizers: application of multilayered feature selection approach in QSAR modeling. *Struct Chem.* 2020;1–13.
- [21]. Li X, Fourches D. Inductive transfer learning for molecular activity prediction: Next-gen QSAR models with MolPMoFit. *J Cheminformatics.* 2020;12:1–15.
- [22]. Díaz RF, Sanchez-Robles B. Non- parametric analysis of efficiency: An application to the pharmaceutical industry. *Mathematics.* 2020;8(9):1522.
- [23]. Rosencrance L, Burns E. What is logistic regression?; 2019. definition from whatis.com. Available: <https://searchbusinessanalytics.t echtarget.com/definition/logistic-regression>.
- [24]. Agatonovic-Kustrin S, Beresford R. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *J Pharm Biomed Anal.* 2000;22(5):717–27.
- [25]. Shin-ike K. A two phase method for determining the number of neurons in the hidden layer of a 3-layer neural network. *ProcSICE Ann Conf.* 2010; 2010:238– 42.
- [26]. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521(7553):436–44.
- [27]. Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T. The rise of deep learning in drug discovery. *Drug Discov Today.* 2018;23(6):1241–50.
- [28]. Yang Y, Ye Z, Su Y, Zhao Q, Li X, Ouyang D. Deep learning for In vitro prediction of pharmaceutical formulations. *ActaPharm Sin B.* 2019;9(1):177–85.
- [29]. Ma J, Sheridan RP, Liaw A, Dahl GE, Svetnik V. Deep neuralnets as a method for quantitative structure–activity relationships. *J Chem Inf Model.* 2015;55(2):263–74.
- [30]. Mohs RC, Greig NH. Drug discovery and development: Role of basic biological research. *Alzheimer's & Dementia: Translational Research & Clinical Interventions.* 2017;3(4):651–7.
- [31]. Henstock PV. Artificial intelligence for pharma: time for internal investment. *Trends Pharmacol Sci.* 2019;40(8):543–6.
- [32]. Mak K-K, Pichika MR. Artificial intelligence in drug development: present status and future prospects. *Drug Discovery Today.* 2019;24(3):773–80.
- [33]. Ramsundar B, Liu B, Wu Z, Verras A, Tudor M, Sheridan RP, Pande V. Is multitask deep learning practical for pharma?. *Journal of chemical information and modeling.* 2017;57(8):2068-2076.
- [34]. Damiani SA, Alaofi AL, Dhar P, Alhakamy NA. Novel machine learning application for prediction of membrane insertion potential of cell-penetrating peptides. *Int J Pharm.* 2019;567:118453.
- [35]. Paranjpe MD, Taubes A, Sirota M. Insights into computational drug repurposing for neurodegenerative disease. *Trends Pharmacol Sci.* 2019;40(8):565–576.
- [36]. Richardson PJ, Corbellino M, Stebbing J. Baricitinib for COVID-19: A suitable treatment?—Authors’ reply. *Lancet Infect Dis;* 2020.
- [37]. Nagy J, Oláh J, Erdei E, Máté D, Popp J. The role and impact of Industry 4.0 and the internet of things on the business strategy of the value chain—the case of Hungary. *Sustainability.* 2018;10(10):3491-3516.
- [38]. PubMed. Database resources of the national center for biotechnology information. Bethesda (MD): U.S. National Library of Medicine. Available: <https://www.ncbi.nlm.nih.gov/pubmed/?term0pharmaceutical+sciences>. Data accessed August 17, 2019.
- [39]. Wakabayashi G, Sasaki A, Nishizuka S, Furukawa T, Kitajima M. Our initial experience with robotic hepatobiliary- pancreatic surgery. *J Hepato-Biliary- Pancreat Sci.* 2011;18(4):481–7.
- [40]. Lo Y-C, Rensi SE, Torng W, Altman RB. Machine learning in chemoinformatics and drug discovery. *Drug Discov Today.* 2018;23(8):1538–46.
- [41]. Kachrimanis K, Rontogianni M, Malamataris S. Simultaneous quantitative analysis of mebendazole polymorphs A–C in powder mixtures by DRIFTS spectroscopy and ANN modeling. *J Pharm Biomed Anal.* 2010;51(3):512– 20.
- [42]. Lamberti MJ, Wilkinson M, Donzanti BA, Wohlhieter GE, Parikh S, Wilkins RG, et al. A study on the application and use of artificial intelligence to support drug development. *Clin Ther.* 2019;41(8):1414– 26.