Journal of Artificial Intelligence and Capsule Networks (ISSN: 2582-2012)
www.irojournals.com/aicn/

AI-Powered Drug Detection System Utilizing Bioactivity Prediction and Drug Release Tracking

Hari Krishnan Andi

Centre for Postgraduate Studies, Asia Metropolitan University, Malaysia

E-mail: hari.andi@amu.edu.my

Abstract

In recent years, Artificial Intelligence (AI) and Machine Learning technologies have played an emerging trend aiding in the creation of new medicines. Simply said, deep learning algorithms and artificial neural networks have brought a new level of sophistication to this field. In recent years, Artificial Intelligence through Machine Learning have been used in this area, and its use is supported by historical data. Additionally, freshly created modelling algorithms relied heavily on unique data mining, duration, and management strategies, which were compared to gauge overall efficiency. This paper suggests the AI powered Drug Detection System using Bioactivity Prediction and Drug Release Tracking. The experimental findings show that the suggested systems effectively recognize the illegal drug advertisements. Datasets with millions of posts gathered using the Google+ API have been used to meticulously verify both the methods. The experimental evidence shows that both approaches can be used to accurately identify medicines.

Keywords: Machine learning, drug discovery, biomolecular, bioactivity, pharmaceutical industry

1. Introduction

Over the last two decades, chemists and biologists have struggled mightily to incorporate perfect methods for the safe and effective delivery of medicinal chemicals to specific areas of the body [1]. However, limitations like accuracy and efficiency are imposed by these methods [2]. As a result, there has been a rise in the use of unique self-fulfilling techniques, which aim to overcome problems plaguing conventional computational approaches [3]. This is especially true with Deep Learning (DL) and Machine Learning (ML)

DOI: https://doi.org/10.36548/jaicn.2022.4.003
Received: 28.09.2022, received in revised form: 13.10.2022, accepted: 30.10.2022, published: 09.11.2022

algorithms. The process of finding a successful treatment for an illness adds an enormous burden. Therefore, pharmaceutical firms' primary concern is to balance cost and speed [5]. The use of Artificial Intelligence (AI) has provided straightforward, empirical answers to all these issues, cutting down the time and money required for this procedure. Pharmaceutical businesses and chemists do vital research in the field of drug creation and development. However, medication design and discovery face obstacles such as limited effectiveness, off-target delivery, lengthy processes, and increased costs. Genomics, protoplasmic, microarray, and clinical trial data are all complex and data-heavy, which slows down the drug development process.

The creation of novel pharmaceuticals is still a major obstacle to advancing modern biomedicine. Bioinformatics and cheminformatics research have been using computational tools to better understand molecular pathways and provide potential therapy options for a variety of ailments for almost three decades. New possibilities in biomedicine and other areas of data science have arisen due to the recent advancements in computing power (such as massively parallel computing and computing on Graphical Processing Units, or GPUs) and data analysis and inference techniques [5-8]. The study and development of creating intelligent devices and superior computer programmes is the field of study known as artificial intelligence. It's somewhat unlike utilizing computers to deduce how humans think, however AI isn't limited to employing techniques found in biology [9].

Due to the widespread use of AI in numerous fields, including but not limited to robotics, voice translation, image analysis, etc., AI approaches are now accustomed. In an effort to create intelligent machines, researchers in the field of computer science have devised a number of novel strategies, some of which have been discovered to be relevant in the fields of chemistry, biology, and pharmaceuticals. Artificial intelligence is being employed in the creation of new organic synthesis schemes, the comprehension of complicated biological systems, the creation of new Application Programming Interfaces (APIs), and the creation of new analytical or diagnostic equipment and procedures. As with practically every other area of the pharmaceutical sciences, AI approaches have applications in drug research and development, drug repurposing, drug metabolism prediction, drug toxicity analysis, increasing pharmaceutical production, and even clinical trials [10].

To put it simply, drug discovery is one of the first processes in the drug development pipeline, and it involves determining the functions of bioactive molecules in the creation of novel pharmaceuticals. Conventional drug development begins with the definition of a

treatment's biomolecular targets, followed by high-throughput screening trials to discover compounds with the desired bioactivity against those targets. Discovering promising drug candidates is the goal of high-throughput screening. The development of high-throughput screening technology has made it feasible to run tests to screen hundreds of chemicals for bioactivity levels on specific target proteins [11]. However, high-throughput screening experiment design is labor- and time-intensive, and it requires access to sophisticated chemical and biological libraries.

In addition, it would be impractical to do high-throughput screening tests on all known chemicals and on all expressed proteins in the human genome [8]. The high failure rates of high-throughput screening are a further issue that prevents the discovery of new medications [9]. The situation becomes much more dire when, how new medicines are created, is considered. Drug development encompasses the whole process of bringing a new medication to market, from the first stages of research and testing through the last stages of human clinical trials.

The length of time required, and the total cost of manufacturing are the two main issues with medication design and development. Drug delivery and development face additional obstacles such as inefficiency, improper target delivery, and incorrect dose. Computer-aided drug designing using AI algorithms has the potential to do away with the problems and restrictions of conventional medication research and design. Supervised learning, unsupervised learning, and reinforcement learning all fall under the umbrella of machine learning, which in turn is a subset of artificial intelligence. Further, the machine learning subfield is known, and it has seen widespread use in the pharmaceutical industry [12].

2. Literature Survey

For the best results while treating HIV with antiretroviral therapy, Shen et al. [13], created an artificial intelligence procedure called AI-PRS to figure out which medication combinations and dosing schedules work best. It is based on method that uses a parabolic response curve to connect medication combinations and dosages to effectiveness. A total of 10 HIV-positive patients were given tenofovir, efavirenz, and lamivudine in their trial, and after some time, the researchers discovered that the dosage of tenofovir could be lowered by 33 percent without triggering a viral relapse utilising the PRS technique.

It takes a lot of time and effort to create and keep track of how drug-like something works. This is why there are a variety of internet resources available to help with medication release analysis and verifying the legitimacy of certain bioactive chemicals used as carriers. Following this, the computational analysis is validated using benchmark datasets. A pharmacophore based on a chemical characteristic is most appropriate for this kind of assessment. Specifically, Mackey et al. [14], fabricated an ad, promising people they could get medications without a doctor's prescription. The advertisement was published on Twitter, Instagram, and Facebook. The experiment's bogus drug adverts were simple to identify, with the exception of one account that was later removed due to questionable behaviour. One fifth of all internet postings promote illegal and/or fake goods, according to the research by Stroppa et al. [15]. In addition, their findings demonstrated the need for the development and deployment of cutting-edge and individualized screening and detection systems in the quest to locate unlawful cyber merchants and online strategies.

Several research projects have focused on the computational aspects of identifying harmful and/or unwanted social media marketing. One such methodology, for identifying microblogging spammers, was developed by Hu et al. [16]. In addition, Zheng and coworkers [17] created an SVM-based machine learning model to identify Sina Weibo spammers. An unsupervised approach that may be used to locate social media postings relevant to current events analysis [18], was developed by Agrawal et al. To combat the difficulties of text mining on microblogged platforms, Jain et al. [19], combined hybrid neural networks to identify spam. In contrast to the prior research, this work concentrates on the presence of drug advertising on social media sites. The purpose of this study is to use epidemiological techniques for the investigation of online infrastructures that facilitate opiate addiction.

3. Bioactivity Prediction and Drug Release Tracking

3.1 Impact of Artificial Intelligence

Datasets that are too large and complex to be processed by traditional methods of data analysis are referred to as "big data". Big data is defined by its volume, velocity, and variety. Volume refers to the massive quantity and weight of information created, while velocity refers to the pace at which this information is replicated. The variety characterizes the heterogeneity of the datasets [20]. Recent advances in technologies such as microarrays, have resulted in an explosion of biological data, ushering modern drug development into the age of big data.

3.2 Large Datasets on the Pharmaceutical Industry

The development of AI has simplified big data analytics by providing several ML algorithms that may aid in the extraction [21]. Gene expression, for instance, is often utilised in target identification in order to learn about disease processes and locate causative genes.

3.3 Computational Drug Design

The whole procedure of drug design has been revolutionized by the widespread use of computer tools in recent years. Traditional computational approaches have been around for a while, but they still have their fair share of problems, such as lengthy run times, high costs, and a lack of confidence in the results [22]. Artificial intelligence has the potential to eliminate these roadblocks in computational drug design and expand the usefulness of computational approaches to drug development. The novel medications are developed based on how a protein interacts with its ligands in three dimensions, which is a crucial information for the drug development process.

Furthermore, quantum mechanics calculates protein-ligand interactions during medication development and defines the characteristics of molecules at the subatomic level. Quantum physics, however, may be computationally expensive and demanding when using traditional computing approaches, which can impact its accuracy [23]. However, artificial intelligence has the potential to make quantum physics more accessible and practical. Traditional drug discovery may also benefit from the use of several text mining-based methods that have recently been created. The goal of text mining is to extract useful information from large amounts of unstructured text stored in a variety of sources by using Natural Language Processing (NLP) to convert the text into structured data. This AI-based algorithm processes and interprets human languages like voice and written text. In order to make use of these AI-driven methods, several text mining tools have been created [24].

3.4 Evaluation of Drug Dosage and Method of Administration

A patient may have unwanted and sometimes fatal side effects if they are given the wrong dosage of a medication. That's why it's so important to find the optimal therapeutic dosage for a medicine. Identifying the optimal dose of a medicine that produces the intended benefit while minimizing hazardous side effects has proven difficult throughout the years [25]. Since the advent of AI, many scientists have turned to machine learning and deep learning algorithms to aid them in dosing medicines correctly.

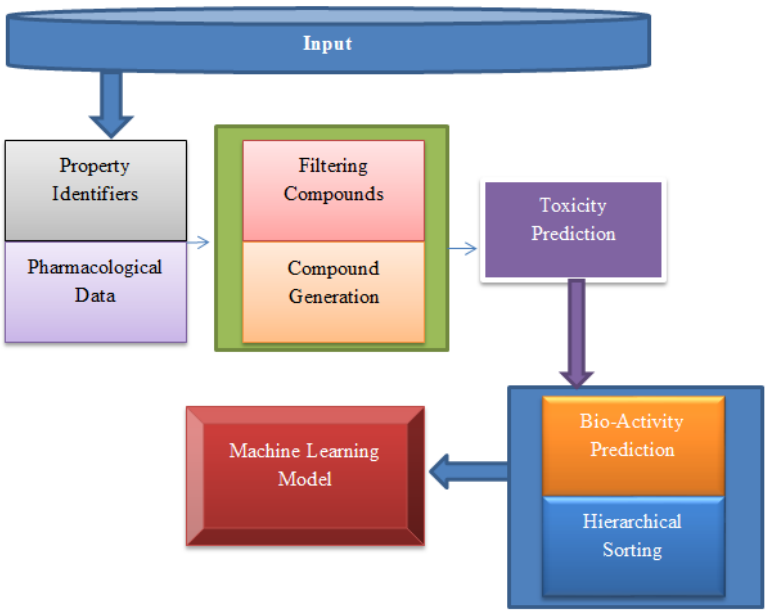

Figure 1. Bioactivity Prediction Model

3.5 Predictive Modelling of Drug Action based on Structure and Ligand

The process involves locating a specific chemical that binds to a medication target. In order to find effective medicinal molecules, Virtual Screening can quickly sort through a large pool of potential candidates [26]. As a result, it's become a crucial resource for high-throughput screening; a process plagued by the twin issues of high cost and poor accuracy. This method determines which two or more highly active molecules has the missing functional ligand by comparing their chemical and physical properties. Because it is independent of the target protein's 3-D structure, this approach may be used even when that structure is unavailable or when just a rough outline is known [27]. However, Structure Based Virtual Screening has been used when the structure is three-dimensional. The goal of this technique is to determine which amino acid residues in a therapeutic target are responsible for binding the active ligand.

3.6 An SVM-based Classification

The preprocessed Term Frequency-Inverse Document Frequency features are used to train a Support Vector Machine (SVM) model, which is then utilized to predict the labels of unseen postings.

3.7 Networks of Neural Convolutions

Due to its single neural network layer, TextCNN is both extremely scalable and sensitive to text categorization. The overarching structure of TextCNN [28] is developed. The

word vector is considered to have a size ‘d’. Next, filters are applied to convolutions of the sentence matrix to produce feature maps. Here, region sizes of 2, 3, and 4 with two filters per area are used. A max-pooling procedure is then performed on the feature map to extract the largest possible set of features. A feature vector consisting of six items from six feature maps may therefore be used as input to a softmax layer.

4. Results and Discussion

The accuracy is determined by how many unlawful advertisements are anticipated and located among all the typical real-world items. Increasing prices and decreased efficiency are major obstacles for the pharmaceutical business when creating a new medicine. However, there is promising potential for ML methods and recent advances in DL to cut down on these expenses, boost productivity, and shorten the duration of the drug development process [29]. This simple access to massive amounts of data suggests that this is the cusp of a third generation of AI. Many pharmaceutical corporations have formed partnerships with AI firms as a consequence of the widespread interest in the role AI is playing in the drug development process. Additionally, by June of 2020 [30], there will have been 230 new companies in this space. Table 1 shows obtained results computations.

Table 1. Overall efficiency calculation tabulation

Proposing Model	Prediction	Accuracy	Efficiency	Size of Dataset	Detection Rate (%)
Existing Work(Manuel)	54.2%	32%	41%	Large	42.12%
Existing Work(Computerized)	60.27%	58%	58%	Large	59.2%
AI + Machine learning Model(SVM based)	80.27%	87.2%	85.3%	Large	87.23%
Proposed Hybrid SVM + TextCNN	94.7%	98%	97.8%	Large	**98.30%**

Moreover, DL techniques overcome the limitations of AI and ML by integrating data at several levels using nonlinear models. The DL and ML combination is through the dynamic interaction between the output and the molecular docking. However, the DL algorithm's advantages come into play due to the integration of data at several levels, resulting in high accuracy and precision. It is now simple to implement AI applications in picture processing, which have surpassed human capability.

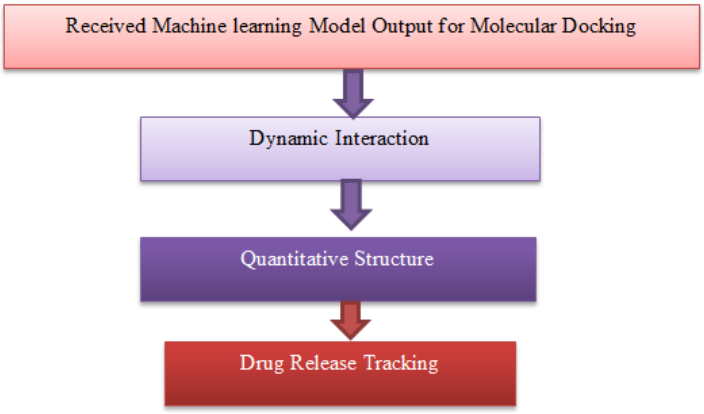

Figure 2. Model Evaluation through Drug Release Tracking

Even more importantly, the proliferation of ML models, each of which claims to be the most recent, has left laypeople in need of guidance in selecting the most appropriate model with which to deal. Users and developers should thus agree on a universally accepted objective assessment before checking the model's efficacy. It is also important to remember that most nations do not provide patent protection for innovations generated solely by AI systems. Figure 3 graph shows the overall performance evaluation.

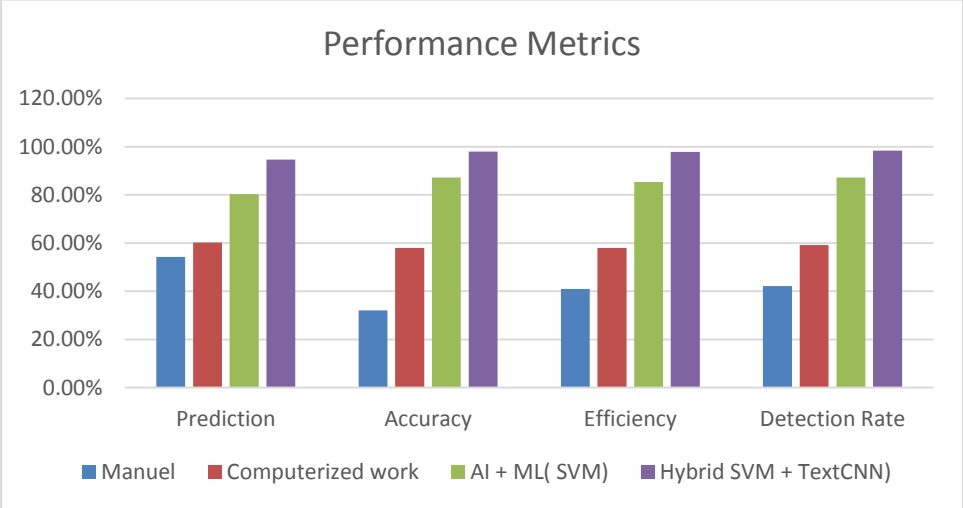

Figure 3. Overall Performance Evaluation

Drug discovery firms that use AI must also go through a rigorous copyrighting procedure in order to protect their intellectual property. Concerns about privacy and safety are warranted because of the sensitive nature of the data needed to provide the tailored care that drives AI. Finally, quicker calculation will be needed when working with massive volumes of data. Current-generation supercomputers are predicted to be obsoleted in the not-

too-distant future by quantum computers or similar technologies that can do the same task in seconds or minutes.

5. Conclusion

The present opioid crisis may be in large part due to the ease with which illegal drugs may be bought and sold on social networking sites. For further epidemiology research and to create intervention applications, it is necessary to have tools for monitoring and evaluating online drug marketplaces. In this study, the social network is exploited to show that machine learning -based approaches can successfully spot ads of illegal drugs in social media messages. The data relating to the opiate misuse crisis might be extracted and analyzed using the suggested methods, which would be useful to the researchers, law enforcement, etc. Moreover, these findings have important implications for informing the development of guidelines and strategies for implementing public health interventions in virtual communities.

References

[1] Xing J, Wenchao L, Liu R, Wang Y, Xie Y, Zhang H, Shi Z, Jiang H, Liu Y-C, Chen K et al (2017) Machinelearning-assisted approach for discovering novel inhibitors targeting bromodomain-containing protein 4. J Chem Inf Model 57(7):1677–1690.

[2] Xue LC, Dobbs D, Bonvin AMJJ, Honavar V (2015) Computational prediction of protein interfaces: A review of data driven methods. FEBS Lett 589(23):3516–3526

[3] Yamanishi Y, Araki M, Gutteridge A, Honda W, Kanehisa M (2008) Prediction of drug–target interaction networks from the integration of chemical and genomic spaces. Bioinformatics 24(13):i232–i240

[4] Yavuz BÇ, Yurtay N, Ozkan O (2018) Prediction of protein secondary structure with clonal selection algorithm and multilayer perceptron. IEEE Access 6:45256–45261

[5] Youjun X, Pei J, Lai L (2017) Deep learning based regression and multiclass models for acute oral toxicity prediction with automatic chemical feature extraction. J Chem Inf Model 57(11):2672–2685

[6] Wu Z, Cheng F, Li J et al (2017) SDTNBI: an integrated network and chemoinformatics tool for systematic prediction of drug-target interactions and drug repositioning. Brief Bioinform. https:// doi.org/10.1093/bib/bbw012

[7] Zeng X, Zhu S, Liu X et al (2019) DeepDR: a network-based deep learning approach to in silico drug repositioning. Bioinformatics.

[8] Chen H, Cheng F, Li J (2020) IDrug: Integration of drug repositioning and drug-target prediction via cross-network embedding. PLoS Comput Biol. https://doi.org/10.1371/journal.pcbi.1008040

[9] Li B, Dai C, Wang L et al (2020) A novel drug repurposing approach for non-small cell lung cancer using deep learning. PLoS ONE. https://doi.org/10.1371/journal.pone.0233112

[10] Schneider P, Walters WP, Plowright AT et al (2020) Rethinking drug design in the artifcial intelligence era. Nat Rev Drug Discov 19:353–364. https://doi.org/10.1038/s41573-019-0050-3

[11] Chen H, Engkvist O, Wang Y et al (2018) The rise of deep learning in drug discovery. Drug Discov Today 23(6):1241–1250. https://doi.org/10.1016/j.drudis.2018.01.039

[12] Zhang W, Xu H, Li X et al (2020) DRIMC: an improved drug repositioning approach using Bayesian inductive matrix completion. Bioinformatics. https://doi.org/10.1093/bioinformatics/ btaa062

[13] Shen Y, Liu T, Chen J et al (2020) Harnessing artifcial intelligence to optimize long-term maintenance dosing for antiretroviral-naive adults with HIV-1 Infection. Adv Ther 3:1900114. https://doi.org/10.1002/adtp.201900114

[14] Mackey, T.K., Liang, B.A.: Global reach of direct-to-consumer advertising using social media for illicit online drug sales. Journal of medical Internet research 15(5) (2013)

[15] Stroppa, A., di Stefano, D., Parrella, B.: Social media and luxury goods counter feit: a growing concern for government, industry and consumers worldwide. The Washington Post (2016)

[16] Hu, X., Tang, J., Zhang, Y., Liu, H.: Social spammer detection in microblogging. In: Twenty-Third International Joint Conference on Artifificial Intelligence (2013)

[17] Zheng, X., Zeng, Z., Chen, Z., Yu, Y., Rong, C.: Detecting spammers on social networks. Neurocomputing 159, 27–34 (2015)

[18] Agrawal, M., Velusamy, R.L.: R-salsa: A spam filtering technique for social networking sites. In: 2016 IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS), pp. 1–7 (2016). IEEE

[19] Jain, G., Sharma, M., Agarwal, B.: Spam detection in social media using convolutional and long short term memory neural network. Annals of Mathematics and Artificial Intelligence 85(1), 21–44 (2019)

[20] Miotto R, Wang F, Wang S, et al. Deep learning for healthcare: review, opportunities and challenges. Brief Bioinform 2017; 1–11; doi: 10.1093/bib/bbx044.

[21] Ching T, Himmelstein DS, Beaulieu-Jones BK, et al. Opportunities And Obstacles For Deep Learning In Biology And Medicine. bioRxiv 2017;142760:

[22] Goh GB, Hodas NO, Vishnu A. Deep learning for computational chemistry. J Comput Chem 2017;38(16):1291–307.

[23] Boughorbel S, Jarray F, El-Anbari M. Optimal classifier for imbalanced data using Matthews correlation coefficient metric. PLoS One 2017;12:1–17.

[24] Mervin LH, Afzal AM, Drakakis G, et al. Target prediction utilising negative bioactivity data covering large chemical space. J Cheminform 2015;7:51.

[25] Segler MHS, Kogej T, Tyrchan C, et al. Generating focused molecule libraries for drug discovery with recurrent neural networks. ACS Cent Sci 2018;4(1):120–31.

[26] M. Vettoretti, G. Cappon, A. Facchinetti, and G. Sparacino, "Advanced diabetes management using artifificial intelligence and continuous glucose monitoring sensors," Sensors, vol. 20, no. 14, p. 3870, 2020.

[27] A. Beneyto, B. W. Bequette, and J. Vehi, "Fault tolerant strategies for automated insulin delivery considering the human component: current and future perspectives," Journal of diabetes science and technology, vol. 15, no. 6, pp. 1224– 1231, 2021.

[28] X. Sun, M. Rashid, N. Hobbs, R. Brandt, M. R. Askari, and A. Cinar, "Incorporating prior information in adaptive model predictive control for multivariable artificial pancreas systems," Journal of Diabetes Science and Technology, vol. 16, no. 1, pp. 19–28, 2022.

[29] A. Esteva, B. Kuprel, R. A. Novoa et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, no. 7639, pp. 115–118, 2017.

[30] N. P. Long, K. H. Jung, N. H. Anh et al., "An integrative data mining and omics-based translational model for the identification and validation of oncogenic biomarkers of pancreatic cancer," Cancers, vol. 11, no. 2, p. 155, 2019.

Author's biography

Hari Krishnan Andi is presently working as the director, Centre for Postgraduate Studies, Asia Metropolitan University, Malaysia. His major area of research includes emotional intelligence, data mining, soft skills, business management, psychological development, and mentoring techniques.