Toxins prediction in Chemicals using Machine Learning Techniques

[1]Rama Prabha K.P,[2]Asha N, [3]V. Sujatha, [4]M B Benjula anbu malar, [5]P. Vijayalakshmi, [6]Jayanthi.R

[1]Assistant Professor Senior, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai 600127, Tamil Nadu, India
ramaprabha.kp@vit.ac.in

[2]Assistant Professor Senior, School of Information Technology and Engineering, VIT, Vellore 632014, Tamil Nadu, India
nasha@vit.ac.in, ORCID:Orcid: 0000-0003-4303-3884

[3]Assistant Professor, School of Advanced Sciences, Department of Mathematics, Vellore Institute of Technology, Vellore 632014, Tamil Nadu, India.
sujathavishnu03@gmail.com

[4]Assistant Professor, School of Information Technology and Engineering, VIT, Vellore 632014, Tamil Nadu, India
benjula2929@gmail.com

[5]Associate Professor, Department of Computer science and Engineering
Panimalar Engineering College
pasramviji@gmail.com

[6]Associate Professor, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai 600127, Tamil Nadu, India
jayanthi.r@vit.ac.in

Abstract:

There are thousands of chemicals, both manmade and naturally occurring, to which humans are exposed. It is now common knowledge that a drug's potency is only one factor in determining its success. Absorption, distribution, metabolism, excretion, and toxicity are only some of the features that need to be taken into account. Due to the extensive resources needed to evaluate a chemical, little is learned about its potential toxicity. There are a wide range of toxicities or adverse medication effects that need to be assessed during both the preclinical as well as clinical trial phases to ensure patient safety. To determine the safety of certain chemicals, scientists have typically used in in vitro as well as in vivo trials. Yet, not only are such research costly and time-consuming, but animal testing experiments specifically are increasingly being criticized for being unethical. The field has seen some success employing conventional machine learning (ML) techniques. When combined with Big Data and AI, machine learning's successes in fields like NLP, speech recognition, image identification, combinatorial chemistry, and genomics suggest it may be useful for toxicity prediction in the modern era. In this piece, we will apply state-of-the-

art machine learning techniques to the problem of toxicity prediction. These techniques include deep learning, regression trees, k-nearest neighbors, and support vector machines, among others. We also talk about how changing the machine learning algorithm's input parameters, such as moving from a focus on chemical structure description alone to additionally including study of human transcriptome data, can significantly improve its prediction accuracy.

Keywords- *Toxicity prediction, machine learning, chemical structure, binary toxicity, multiclass toxicity, categorical toxicity.*

1. Introduction:

Toxicity assessment is difficult because in vivo systems are notoriously complicated. This makes concerns about medication safety a major factor in the discontinuation of drugs during either the experimental or therapeutic stages of development. Protection agencies have grown increasingly frustrated over the years due to the high rate of harmful test failures. Studies have shown that 90% of drugs that make it to the clinical trial stage will fail during either the phase I, II, or III testing required for drug approval. Clinical trial evaluations performed between 2010 and 2017 revealed that uncontrollable toxicity was the cause of 30% of medication development failures. Several pharmaceuticals are pulled from the market despite having been approved, because of concerns about their safety. Patients, medical professionals, investors, and government agencies all lose faith in the sector as a result of this. [1-6] Thus, computational toxicity forecasting is most useful at the outset of drug discovery, when it may be used to screen out compounds that are highly unlikely to succeed in human clinical trials.

It is crucial to test drugs for toxicity before approving them for use. It's common knowledge that medications need to pass clinical trials before they can be sold legally [1, 2]. There is always some danger involved with participating in a research trial. Studies conducted late in the development of new medications on people have reportedly showed that over half of these drugs are either dangerous or ineffective.[3]. Specifically and irreversibly hepatotoxicity in men [4, 5] led to the immediate withdrawal of the medication Sitaxentan (Figure 1). The need for preclinical analyses to screen out potentially harmful medications is highlighted by the fact that unsafe clinical trials draw attention to this area of research. The animal trial is frequently used for preclinical research, however it has limited utility. The trial, on only one hand, is time-consuming and costly. However, due to heterogeneity between species and illness models [6,7] the findings provide little direction for human toxicological reactions. For instance, while hepatotoxicity with Sitaxentan was notably present in people no such harm was observed in animal trials [8]. Thus, there is no risk reduction to be gained from using animals to test new medications [9].

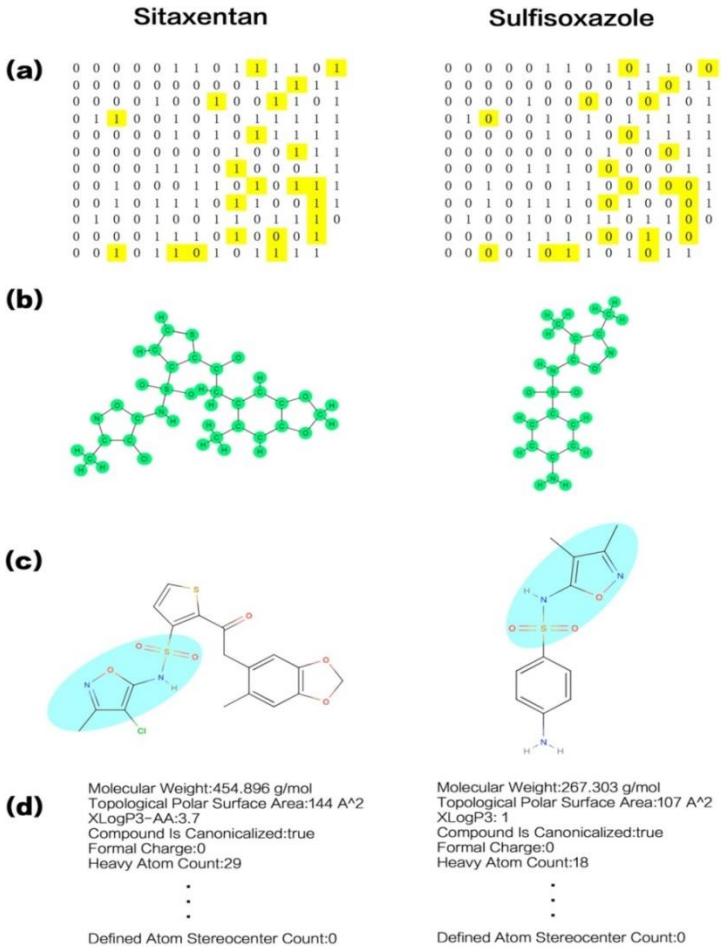

Figure 1. Sitaxentan with Sulfisoxazole: Chemical Structure Description

High-throughput computer toxicity forecasts are essential for lowering the costs and reducing the inherent uncertainties of animal research. Quantitative structure-activity relations (QSAR) using chemical structural factors are a well-established and popular approach to toxicity prediction [10]. This technique makes use of statistical methods to determine a quantifiable connection between a medicinal compound's structural or physicochemical features and its physiological activity [11]. Toxicological profiles and other aspects of the substance can be extrapolated from the relationship. The Hansch approach, developed in 1962 [12] and still extensively used today, implies independent of the factors regulating the biological activities of the compounds. The QSAR model is obtained by free-energy-related statistical techniques like linear regression. In 1964, Free and Wilson established a method for analyzing the relationship between chemical structure and physiological activity using regression analysis [13]. Using QSAR regression analysis [14–16] to anticipate drug toxicity was first implemented in the 1980s. As we entered the new millennium, scientists performed toxicity prediction using either a single or many physicochemical mechanisms [17]. Most research into this process made use of regression analyses, multivariate, or neural network models. Knowledge based systems, in addition to statistical methods, were frequently utilized to analyze the interplay of the various mechanisms. Since the volume of data continues to grow

at an exponential rate, it becomes increasingly challenging for knowledge-based systems to perform fully automated work with a large number of data [18]. Yet, it is challenging to forecast with high accuracy using statistical methods like regression analysis, multivariate, even early shallow models of neural networks.

The Rest of the Paper is organized as Follows. Section 2 Mainly discusses about the Previous survey of works regarding toxins prediction using Various ML Techniques.Section3 Discusses about the proposed method and its architecture for prediction of toxins using novel ML Techniques. Section 4 Evaluates the model by comparing with existing algorithms AND FINALLY Section5 Concludes the work

2. Literature Survey:

Machine learning is a subfield of AI that employs complex mechanisms to endow systems with the capability to "learn" & "predict" [19]. Artificial neural networks (ANN), decision trees, neural networks with support vector machines (SVM), & Bayesian classifiers are the primary algorithms of machine learning [20]. These algorithms originated from the research of clustering analysis & pattern identification. These has many connections to data mining [21], including cluster analysis & pattern identification. In order to forecast toxicity, a growing number of researchers are turning to machine learning [22]. This is due to the many benefits of machine learning, including its speed, low cost, and high accuracy. Multiple algorithms, including the evolutionary algorithms (GA) [23,24], the decision tree model (RF) [25–27], the neural network algorithm (ANN) [28–30], as well as other systems [31–33], have been used to improve upon the accuracy of conventional QSAR models for predicting the toxicities and biological activities of drugs. The results of various machine learning techniques vary. The effectiveness can be greatly influenced by data sets and computational representations.

Rosenblatt proposed a perceptron model, which mimics neuron structure and may serve as a classifier [34], in 1957. In the beginning, there was Window and Hoff [35] who employed. Use of delta rules in perceptron training provides a solid foundation for a linear classifier. The k nearest algorithm [36] was proposed by Cover and Hart in 1967 and allows computers to categorize sample points based on spatial attributes. Quilan first presented the algorithm for decision trees [37] in 1986. The basic concept behind SVM, developed by Cortes et al. in 1995, is to identify the boundary between two classes that separates them by the greatest amount of data. Linear high-dimensional categorization has been demonstrated to be possible with SVM [38], in addition to linear classification. Breiman introduced the first version of the RF method [39], a multi-tree classifier, in 2001. Each tree gives its forecast in a separate category, and these categories all vote on the final classification [40]. It finds extensive application in the resolution of multiclass issues. Both SVMs and RFs benefit from more density and structure in their data sets, where their statistical foundations shine.

For nonlinear mapping, Back in 1986, Hinton et al. [41] created the back-propagating method (BP) of multi-layer perception (MLP) using a sigmoid activated function, proving the effectiveness of ANN in this context. In the subsequent year, 1991 [42], it was observed that BP with a function of sigmoid activation has the vanishing grade difficulty, making it

challenging to replicate more complex training. So-called "shallow learning" ANN designs. The vanishing gradients problem of the sigmoid function was addressed in 2011 [43] with the introduction of the ReLU (Rectified Logistic Unit) creation function. With this development, deep learning entered the world. ReLU-based algorithms have achieved impressive results in the field of picture recognition

Deep learning, an offshoot of ANN, has become a thriving area in machine learning. Pattern recognition speech recognition processing of natural languages picture and video recognition and life science are just a few of the areas where it has been crucial in driving innovation. For optimal performance, deep learning requires big, diverse, and sparse training data. The neural network models recurrent neural network models (RNN) and convolutional neural network model (CNN) have both seen widespread application in deep learning in recent years. Natural language analysis and series data prediction are two examples of applications where the former shines. In contrast, the latter excels at identifying features with a spatial layout, like the shapes seen in photographs and graphics . Increases in computing power, the widespread adoption of distributed clusters and graphics processing units (GPUs) and the proliferation of optimization algorithms have drastically shortened the time required to train deep neural networks, making them applicable in fields as diverse as bioinformatics and chemistry

3. Methodology:

Although while it typically necessitates processing power and enormous quantities of information to learn from, one of ML's key advantages is that it enables the modeling or prediction of difficult issues. Machine learning methods can be classified into two broad types: supervised & unsupervised. The first type of method uses annotated data to automatically map inputs to outputs, while the latter type of method can discover underlying relationships straight from a dataset. Supervised learning is frequently employed in the field of medication safety evaluation because of its capacity to examine compound input features in relation to desired outputs such as biological activities or harmful outcomes. Models such as k-nearest neighbors, vector machines, random forests, and algorithms based on neural networks have all seen extensive usage in the field of toxicity assessment (DL, deep learning). One of the most popular ML techniques, DL-based models are built on top of several layers of neural nets. There is a huge breadth of DL designs and algorithms. Multilayer perceptron (MLP) network, recurrent neural networks, & convolution networks are some of the most common. The decision between the two depends on the nature of the problem at hand and the molecular representation being used. Data gathering and preparation, molecular representation identification, model creation via testing on new dataset after training and validating on a portion of the data set are the standard steps in model design. ToXicity prediction using ML is depicted in general in Figure 2. Information about the building blocks and properties of ML models is abundant in the literature; as a result, that is not the subject of this article.

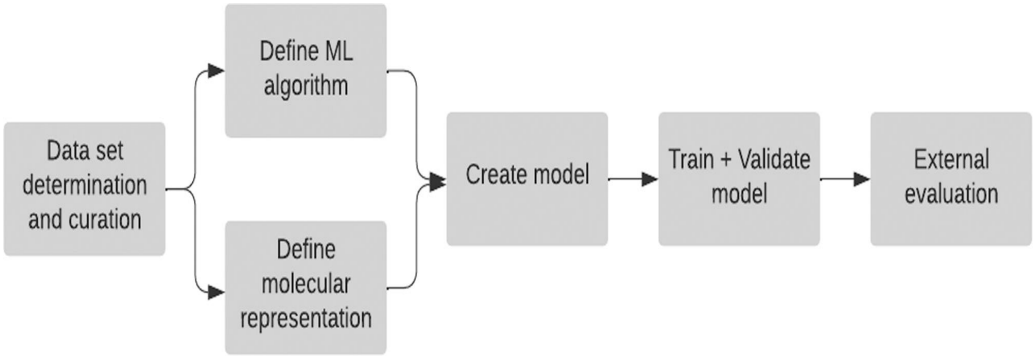

Figure 2.Proposed ML Method for Prediction of Toxins

One of the most important factors in computerized toxicity prediction is the molecular representation chosen. A molecular structure can be represented in a number of different ways, including as a labeled graph where nodes represent atoms and edges represent atom-to-atom bonds as numerical character traits or both qualitative and quantitative data determined from physicochemical properties; as short ASCII chords recognized as SMILES strings; and as molecular fingerprint images, which consist on bits chosen to represent the either the presenilin-1 (psi) or specialists. Toxicological fingerprinting is commonly done using Klekotha-Roth and Extended Connectivity Fingerprinting (ECFP), as well as the Molecular Access System (MAS) key and Pub Chem (PC) substructure fingerprints.

The current state of the art in molecule representation and Machine learning algorithms for descriptive research and causal is discussed, along with examples of the top performing methods. Predictions of toxicity at multiple terminal stages have been made in computer simulations. Tox21 Data Challenge end points, sharp oral toxic, hepatotoxicity, cardiotoxicity, endocrine disruption, & other similar measures are frequently studied. The use of a classifier model is appropriate when toxicity is considered as a binary choice, as in the case of Tox21 terminals, whereas the use of a regression model is appropriate when a quantitative forecast, such as LD50 forecasting, is sought. The many data sources for ML building on each toxic end point are listed in Table 1. Because of differences in complexity, class distributions, and the amount of chemical space covered, it is difficult to compare the relative efficacy of different algorithms across a wide range of harmful endpoints when working with ML methods that produce inconsistent results across data sets. Different evaluation criteria complicate the comparison even more, as their determination is highly dependent on both the ML technique and the database in question. There has been a lot of research into the importance of selecting an appropriate metric to use while constructing ML models.

And getting a satisfiability modulo or comprehension of the identified hazardous responses is one of the biggest hurdles in computational toxicology today. Even while ML techniques, and specifically DL models, are increasingly used in the real world, they are still typically viewed as mysterious black boxes that can solve difficult problems effectively but provide little in the way of explanation when asked to explain their predictions. In recent models, many methods have been used in an effort to improve interpretability. A demonstration that

neural networks may be trained to identify patterns that are identical to well-known toxins. To help scientists determine which characteristics are most strongly linked to toxicity, Wenzel et al. produced response maps. New approaches are appearing that may prove useful for in silico toxicity.

4. Results and Discussion:

It was built using the Java AWT (Abstract Windows Toolkit). The.arff file contains information about 4589 patients. There are two data sets: one for practice and one for evaluation, are then crafted from the collected information. Inside the UI, you can go into the test results and pick a prediction algorithm using a drop-down menu. The findings are summarized and displayed in greater detail. Predictions are made regarding the algorithms at each iteration of the study, with the quality parameters in mind.

In this software, we employ Machine Learning techniques to design a GUI for predicting the presence of toxins during the pharmaceutical production process. The component includes the computer and the user. Each model requires data plus algorithm runs before the system can generate results. Instead, users can see only their own data and the model's projections. But first they have to register and sign in. Once the user has registered, they will have access to all of the models and their predicted outcomes.

1. **Correctly classified instances:** This will reveal the proportion of successful test cases.
2. **Relative Absolute Error:** To determine how much of a discrepancy there is between the observed and expected values, one can employ this technique.
3. **Root Relative Squared Error:** This would have been an improvement above just a simple prediction. An estimate generated from a simple prediction by halving the total squared error is the relative mean - squared error.

Table 1.Analysis Results of Various Quality Parameters in comparison with algorithms

Quality Parameters	**Performance of Various Prediction Algorithms**			
	Linear Regression	**Logistic Regression**	**Ada Boost**	**Proposed Method**
Correctly Classified Instances	12.25	23.25	24.56	34.58
Relative Absolute Error	13.25	14.25	15.25	7.85
Root Relative Error	18.25	16.36	13.25	10.25
R-Squared Error	22.25	18.59	18.25	8.25

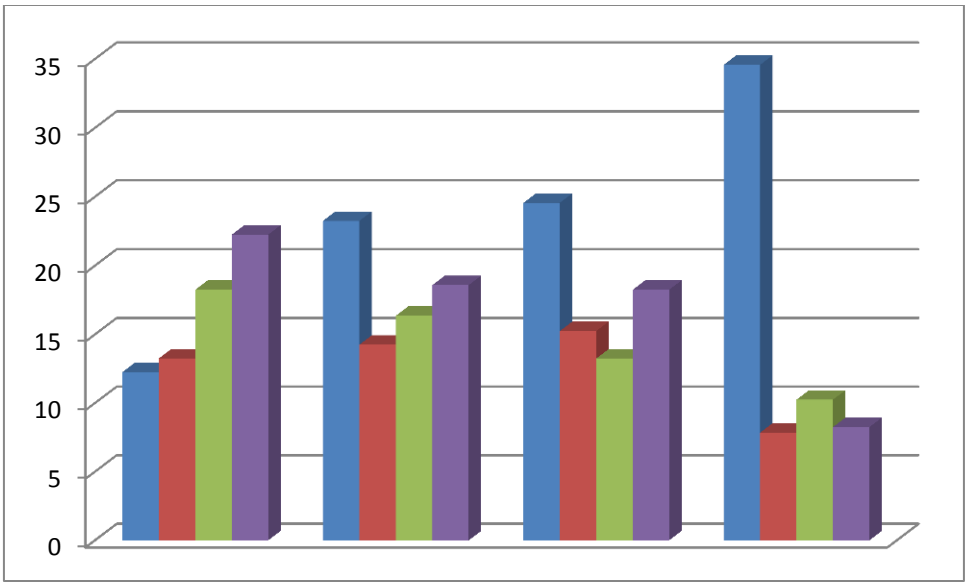

Figure 4.Evaluation Results of Various Parameters

The evaluation results of the different prediction algorithms are shown in the above graph. It is clear from the graph that the proposed method outperforms the other prediction algorithms.

5. Conclusion:

Estimating drug candidates' toxicity is crucial to avoiding unexpectedly high costs, late-stage failures, and significant withdrawals during the drug discovery process. Evidence suggests that Ml algorithms, despite their persisting flaws that were resolved in this research, may be a potential strategy to serve as early filter of dangerous chemicals inside the process of drug discovery. A higher quantity of responsible for considerable and a more generalizable set of methods are needed which will enhance their potential for incorporation into the organic drug development pipeline. To acquire the quantity and quality of information required by ML for developing accurate predictive toxicology models, pharmaceutical companies' cooperation will become increasingly important, despite the growing availability of public data. This paper analyzes and summarizes the modern machine learning (ML) techniques currently being established for each toxic end point, shedding light on areas for future improvement.

References:

1. Ting, N. (Ed.) Introduction and New Drug Development Process. In Dose Finding in Drug Development; Springer: New York, NY, USA, 2006; pp. 1–17.
2. Janodia, M.D.; Sreedhar, D.; Virendra, L.; Ajay, P.; Udupa, N. Drug Development Process: A review. Pharm. Rev. 2007, 5, 2214–2221.
3. Davanam, G., Kumar, T. P., & Kumar, M. S. (2021). Efficient energy management for reducing cross layer attacks in cognitive radio networks. Journal of Green Engineering, 11, 1412-1426.
4. Balaji, K., P. Sai Kiran, and M. Sunil Kumar. "Power aware virtual machine placement in IaaS cloud using discrete firefly algorithm." Applied Nanoscience (2022): 1-9.
5. Kumar, M. Sunil, et al. "Deep Convolution Neural Network Based solution for Detecting Plant Diseases." Journal of Pharmaceutical Negative Results (2022): 464-471.

6. Davanam, G., Pavan Kumar, T., & Sunil Kumar, M. (2021). Novel Defense Framework for Cross-layer Attacks in Cognitive Radio Networks. In International Conference on Intelligent and Smart Computing in Data Analytics (pp. 23-33). Springer, Singapore.
7. Sangamithra, B., Neelima, P., & Kumar, M. S. (2017, April). A memetic algorithm for multi objective vehicle routing problem with time windows. In 2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE) (pp. 1-8). IEEE.
8. Balaji, K., P. Sai Kiran, and M. Sunil Kumar. "Resource aware virtual machine placement in IaaS cloud using bio-inspired firefly algorithm." Journal of Green Engineering 10 (2020): 9315-9327.
9. Ganesh, Davanam, Thummala Pavan Kumar, and Malchi Sunil Kumar. "Optimised Levenshtein centroid cross-layer defence for multi-hop cognitive radio networks." IET Communications 15.2 (2021): 245-256.
10. Kumar, M. S., & Harshitha, D. (2019). Process innovation methods on business process reengineering. Int. J. Innov. Technol. Explor. Eng.
11. Natarajan, V. Anantha, M. Sunil Kumar, Rizwan Patan, Suresh Kallam, and Mohamed Yasin Noor Mohamed. "Segmentation of nuclei in histopathology images using fully convolutional deep neural architecture." In 2020 International Conference on computing and information technology (ICCIT-1441), pp. 1-7. IEEE, 2020.
12. Sreedhar, B., BE, M. S., & Kumar, M. S. (2020, October). A comparative study of melanoma skin cancer detection in traditional and current image processing techniques. In 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) (pp. 654-658). IEEE.
13. Malchi, S. K., Kallam, S., Al-Turjman, F., & Patan, R. (2021). A trust-based fuzzy neural network for smart data fusion in internet of things. Computers & Electrical Engineering, 89, 106901.Quinn, F.R.; Neiman, Z.; Beisler, J.A. Toxicity and quantitative structure-activity relationships of colchicines.J. Med. Chem. 1981, 24, 636–639.
14. Denny, W.A.; Cain, B.F.; Atwell, G.J.; Hansch, C.; Panthananickal, A.; Leo, A. Potential antitumor agents. 36. Quantitative relationships between experimental antitumor activity, toxicity, and structure for the general class of 9-anilinoacridine antitumor agents. J. Med. Chem. 1982, 25, 276–315.
15. Denny, W.A.; Atwell, G.J.; Cain, B.F. Potential antitumor agents. 32. Role of agent base strength in the quantitative structure-antitumor relationships for 4′-(9-acridinylamino) methane sulfonanilide analogs. J. Med. Chem. 1979, 22, 1453–1460.
16. Barratt, M.D. Prediction of toxicity from chemical structure. Cell Biol. Toxicol. 2000, 16, 1–13.
17. Compton, P.; Preston, P.; Edwards, G.; Kang, B. Knowledge Based Systems That Have Some Idea of Their Limits. CIO 2000, 15, 57–63.
18. Mitchell, T.M. Machine Learning; McGraw Hill: Ridge, IL, USA, 1997; Volume 45, pp. 870–877.
19. Sangamithra, B., Manjunath Swamy, B.E., Sunil Kumar, M. (2022). Personalized Ranking Mechanism Using Yandex Dataset on Machine Learning Approaches. In: Kumar, A., Ghinea, G., Merugu, S., Hashimoto, T. (eds) Proceedings of the International Conference on Cognitive and Intelligent Computing. Cognitive Science and Technology. Springer, Singapore. https://doi.org/10.1007/978-981-19-2350-0_61
20. Kumar, M. S., Siddardha, B., Reddy, A. H., Reddy, C. V. S., Shaik, A. B., & Ganesh, D. (2022). Applying The Modular Encryption Standard To Mobile Cloud Computing To Improve The Safety Of Health Data. Journal of Pharmaceutical Negative Results, 1911-1917.

21. Burada, S., Swamy, B. E., & Kumar, M. S. (2022). Computer-Aided Diagnosis Mechanism for Melanoma Skin Cancer Detection Using Radial Basis Function Network. In Proceedings of the International Conference on Cognitive and Intelligent Computing (pp. 619-628). Springer, Singapore.
22. Prasad, T. G., Turukmane, A. V., Kumar, M. S., Madhavi, N. B., Sushama, C., & Neelima, P. (2022). Cnn Based Pathway Control To Prevent Covid Spread Using Face Mask And Body Temperature Detection. Journal of Pharmaceutical Negative Results, 1374-1381.
23. Kumar, M. S., Chandra, K. G., & Gupta, K. K. (2022). Stocks Analysis and Prediction of Indian Oil Trading Using Big Data Analytics. International Journal of Mechanical Engineering, 7(1), 6734-6738.
24. AnanthaNatarajan, V., M. Sunil Kumar, and V. Tamizhazhagan. "Forecasting of Wind Power using LSTM Recurrent Neural Network." Journal of Green Engineering 10 (2020).
25. Natarajan, V. A., Kumar, M. S., Tamizhazhagan, V., & Chevdumoi, R. M. (2022). Prediction Of Soil Ph From Remote Sensing Data Using Gradient Boosted Regression Analysis. Journal of Pharmaceutical Negative Results, 29-36.
26. Sunil Kumar, M., & Rama Mohan Reddy, A. (2016). An Efficient Approach for Evolution of Functional Requirements to Improve the Quality of Software Architecture. In Artificial Intelligence and Evolutionary Computations in Engineering Systems (pp. 775-792). Springer, New Delhi.Agrafiotis, D.K.; Cedeño, W.; Lobanov, V.S. On the use of neural network ensembles in QSAR and QSPR.J. Chem. Inform. Comput. Sci. 2002, 42, 903–911.
27. Wikel, J.H.; Dow, E.R. The use of neural networks for variable selection in QSAR. Bioorgan. Med. Chem. Lett.1993, 3, 645–651.
28. Lu, X.; Ball, J.W.; Dixon, S.L.; Jurs, P.C. Quantitative structure-activity relationships for toxicity of phenols using regression analysis and computational neural networks. Environ. Toxicol. Chem. 1998, 13, 841–851.
29. Lu, J.; Peng, J.; Wang, J.; Shen, Q.; Bi, Y.; Gong, L.; Zheng, M.; Luo, X.; Zhu, W.; Jiang, H. Estimation of acute oral toxicity in rat using local lazy learning. J. Cheminform. 2014, 6, 26.
30. Mazzatorta, P.; Cronin, M.T.D.; Benfenati, E. A QSAR Study of Avian Oral Toxicity using Support Vector Machines and Genetic Algorithms. Qsar Comb. Sci. 2010, 25, 616–628.
31. Srinivasan, A.; King, R.D. Using Inductive Logic Programming to construct Structure-Activity Relationshipsp;AAAI: Menlo Park, CA, USA, 1999; pp. 64–73.
32. Rosenblatt, F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain; MIT Press: Cambridge, MA, USA, 1988; pp. 386–408.
33. Widrow, B.; Hoff, M.E. Adaptive Switching Circuits. In Neurocomputing: Foundations of Research; Ire Wescon Conv. Rec; MIT Press: Cambridge, MA, USA, 1966.
34. Cover, T.; Hart, P. Nearest neighbor pattern classification. IEEE Trans. Inf. Theory 2002, 13, 21–27.
35. Quinlan, J.R. Induction of decision trees. Mach. Learn. 1986, 1, 81–106.
36. Cortes, C.; Vapnik, V.; Cortes, C.; Vapnik, V.; Llorens, C.; Vapnik, V.N.; Cortes, C.; Côrtes, M.V.C.B. Support-vector networks. Mach. Learn. 1995, 20, 27–297.
37. Breiman, L. Random Forests. Mach. Learn. 2001, 45, 5–32.
38. Tin Kam, H. Random Decision Forests. In Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, Canada, 14–16 August 1995; Volume 1, pp.

278–282.

39. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors.Read. Cognit. Sci. 1986, 323, 399–421.
40. Hochreiter, S. The Vanishing Gradient Problem during Learning Recurrent Neural Nets and Problem Solutions. Int. J. Uncertain. Fuzz. Knowl.-Based Syst. 1998, 6, 107–116.
41. Glorot, X.; Bordes, A.; Bengio, Y. Deep Sparse Rectifier Neural Networks. In Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, Fort Lauderdale, FL, USA, 11–13 April 2011;
42. Geoffrey, G., David, D., Miroslav, D., Eds.; PMLR Proceedings of Machine Learning Research, 2011; Volume 15, pp. 315–323.
43. Zahangir Alom, M.; Taha, T.M.; Yakopcic, C.; Westberg, S.; Hasan, M.; Van Esesn, B.C.; Awwal, A.A.S.; Asari, V.K. The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches. arXiv 2018, arXiv:1803.01164.