Received February 5, 2025, March 14, 2025, date of publication May 17, 2025.

*Digital Object Identifier 10.21608/ijaici.2025.330535.1000*

# Chemical Toxicity Prediction Based on Artificial Intelligence: A Review

**Eman Shehab[1], Hamada Nayel[2], Mohamed Taha[2]**

[1] Department of Computer Science, Faculty of Computers and Artificial Intelligence, University of Sadat City, Sadat, Egypt

[2] Department of Computer Science, Faculty of Computers and Artificial Intelligence, Benha University, Benha, Egypt.

Corresponding author: Eman Shehab (e-mail: eman.shehab@fcai.usc.edu.eg).

**ABSTRACT** The increasing number of chemicals has aroused public concern due to their negative influence on the environment and human health. To protect the environment and human health, the toxicity of these compounds must be assessed. Traditional in vitro and in vivo toxicity testing are time-consuming, expensive, and complex, and they may pose ethical considerations as well. Due to these restrictions, alternative methods for assessing the toxicity of a chemical are required. Numerous toxicity prediction models have been developed recently using a variety of machine learning and deep learning algorithms such as support vector machines, random forests, k-nearest neighbors, ensemble learning, and deep neural networks by integrating classical ML techniques or Deep Learning (DL) with molecular representations such as fingerprints or 2D graphs. This paper presents an overview of chemical toxicity and the drug Discovery Process. It summarizes current ML and DL models for predictive toxicology with a brief objective and the limitations and challenges AI faces in toxicity prediction.

**INDEX TERMS** Chemical Toxicity, Drug Discovery Process, Molecular Representations, Machine Learning, Deep Learning.

## I. INTRODUCTION

Chemical toxicity is any harmful effect that can happen when you are exposed to chemicals. It can be measured in different ways, such as long-term toxicity or effects specific to a particular organ, like genotoxicity and carcinogenicity. This can then be translated into quantitative or qualitative parameters like LD50, or low, moderate, or high toxicity. Toxicity studies are designed to precisely discover these adverse effects on humans, animals, plants, or the environment, whether through acute exposure (in a single dose) or cumulative exposures (in repeated doses over time). Many factors influence the toxicity of chemicals: the route of exposure (oral, dermal, or inhaled), the dose, frequency, and duration of exposure, specific properties related to Absorption, Distribution, Metabolism, and Excretion/Elimination (ADME), interactions between exogenous or endogenous substances, subject characteristics (age, sex, or body mass), and specific physicochemical properties (lipophilicity, solubility, boiling point, among others) [1,2].

To ensure public safety by minimizing exposure to hazardous chemicals, regulatory decision-making bodies such as the European Medicines Agency (EMA), U.S. Food and Drug Administration (FDA), Environmental Protection Agency (EPA), and European Environment Agency (EPA) have employed toxicity assessments [3]. Animal tests are the foundation of current conventional toxicity evaluation procedures. These tests are limited, nevertheless, by financial, time, and ethical considerations. Furthermore, testing such many substances through animal experiments is not feasible for regulatory, toxicological, or medicinal development objectives.

To overcome these obstacles, it is critical to create quick and affordable substitutes for performing animal toxicity testing, such as in vitro and silico techniques. Numerous computer techniques, including read-across, structural warnings, and the Quantitative Structure-Activity Relationship (QSAR), have been applied in recent decades to forecast the toxicological consequences of compounds [4]. QSAR establishes a quantitative connection between a chemical's physicochemical or structural properties and its harmful effects. It's been a popular technique for creating toxicity prediction models.

Recently, QSAR based on ML and DL has become increasingly common in predictive toxicology [5]. This is

because of the ongoing advancements in processing power, the rise of big data, and the quick development of ML and DL methodologies. ML and DL are highly appealing computational algorithms for predicting toxicity for a wide range of substances because of their capacity to learn from data and create predictions automatically. While ML and DL-based models have made significant progress in predicting toxicology, there is a rising interest in generating more accurate models. A comprehensive assessment of ML and DL models in predictive toxicology can increase their reliability and provide insight for future improvement.

This review summarizes papers on the in-silico concept and the prediction of chemical toxicity. The search used the drug's key terms in silico, prediction, and chemical toxicity. Among the items searched, the literature was aimed to use the most recent studies (from 2021 to 2024). The remainder of this paper is arranged as follows: section 2 exposes a brief overview of drug discovery process. Section 3 shows a molecular representation. Section 4 displays the models for ML and DL. Section 5 discusses challenges of AI in toxicity prediction. Section 6 presents the paper's conclusion.

## II. The Drug Discovery Process: A Brief Overview

Drug research and development is a multidimensional, complex endeavor. The approach involves four key phases: identifying & validating targets & compound screening and lead optimization & preclinical investigations, and clinical trials. Figure 1 presents the process of drug discovery and development [6]. The initial phase in this approach is identifying pathophysiological variables and biological targets. Bioinformatics, genomics, and proteomic research are required to determine cellular and genetic targets. The first molecule, or hit with activity against the given target, is initially discovered. This can be accomplished by creating chemical libraries or isolating natural compounds from plants, bacteria, and fungi. The next phase is to identify the lead chemical with the most promising potential for medication development. Lead optimization involves changing a selected lead to boost specificity and effectiveness at lower doses. Therapeutic candidates undergo an iterative process that includes cellular tests and structure-activity connections to improve their functional qualities. Animal models are utilized for in-vivo studies, including pharmacokinetic and toxicity assessments. Following preclinical research, the medication candidate is tested on patients in clinical trials [7,8]. Clinical trials are crucial for assessing drug efficacy and patient safety. The method is time-consuming and inefficient. Pharmaceutical businesses therefore look for ways to cut costs and expedite their projects. Artificial Intelligence (AI) refers to a machine's capacity to mimic human cognitive processes, such as learning and problem-solving. AI systems that are based on technology can mimic human intellect via the use of a variety of advanced tools and networks. AI-based technologies are increasingly being deployed at various drug discovery phases to save time and increase profitability. These encompass a range of activities such as computational organic synthesis, compound production, quantum mechanics (QM)-based compound attribute calculation, real-time cell sorting, cell classification, and more [9].

## III. Molecular representations

An essential aspect of AI-based medication discovery and analysis is the conversion of molecules into a computer-readable format while preserving their inherent physicochemical characteristics, given the rapid expansion of natural products [10]. A range of descriptors have been suggested to describe medications; these descriptors can be categorized into four groups based on their dimensionality. Several open-source toolkits, like OpenBabel [11] and ChemmineR [12], have been proposed to speed up drug development by calculating molecular descriptors and structures.

The simplest molecular representation is the zero-dimensional (0D) descriptor, derived from medication chemical formulas [13]. The 0D descriptor often comprises molecular weight, atom number, atom type count, and other basic characteristics, such as the quantity of heavy atoms. The 0D descriptor is simple and only extracts shallow information.

Drugs are encoded using the one-dimensional (1D) descriptor based on their substructures, including the number of rings, functional groups, substituent atoms, and fragments centered on atoms. Typically, the 1D descriptor's elements are binary—for example, 1/0 denotes the presence or absence of a substituent atom—or the frequency at which certain substructures occur. A simplified molecular-input line-entry system (SMILES) is another kind of 1D descriptor besides the property-based 1D descriptor [14]. SMILES represents medications using a string of characters. A medication may have multiple SMILES representations based on its atom order.

The two-dimensional (2D) descriptor takes into account adjacency, connectedness, and other topological properties of the atoms to provide more information than the one-dimensional (1D) descriptor. Consequently, medication is usually represented as a graph with nodes denoting atoms and edges denoting bonds to extract 2D descriptors. Graph invariants, connectivity bonds, graph-based substructures, and topological descriptors are examples of property-based 2D descriptors. The molecular fingerprint (FP), which encodes molecules in binary form, was proposed to extract more information [15]. FP, denoted by 1/0, indicates whether a given substructure is present or absent in a string of a given length. The fingerprints from the molecular access system [16], the daylight-like fingerprint [11], and the extended-

connectivity fingerprint are the most widely utilized 2D FPs [17].

The three-dimensional (3D) descriptor represents a molecule in 3D space [18], with each atom identified by its x, y, and z coordinates. The 3D descriptor provides detailed information about spatial and geometric configurations. 3D

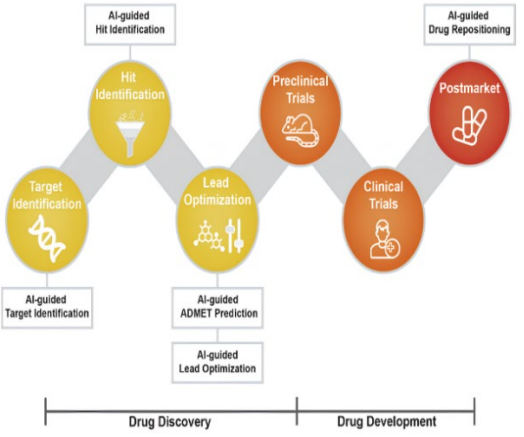


**FIGURE 1.** **The process of drug discovery and development.**

descriptors provide information on surface area, volume, and steric properties. Geometrical fingerprint [19] and pharmacophore fingerprint [20] are examples of non-property-based 3D descriptors. They are commonly employed in drug development and virtual screening due to their ability to reflect complicated physicochemical features accurately.

Figure 2 depicts schematic diagrams of compound representations employing 0D-3D descriptors [21]. Recently, graph-based approaches for encoding molecules have been developed, in addition to existing schemes. Examples of graph-based systems include convolutional networks for spectral and spatial graphs. A recent review provides more information concerning graph-based molecular representation approaches. The molecular graph representation is based on mapping atoms and bonds into sets of nodes and edges.

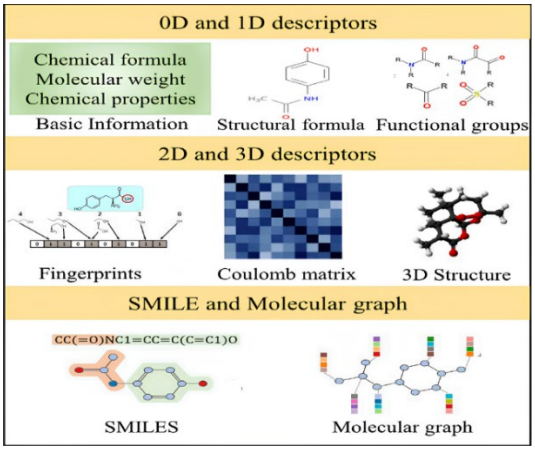


**FIGURE 2.** **Compound representations employing 0D-3D descriptors.**

It makes intuitive sense to regard the bonds in a molecule as edges and the atoms as nodes, but there is no reason why one might not think of alternate mappings. Generally, nodes are represented by circles or spheres, and edges by lines in graph representations. Rather, the nodes in molecular graphs are typically depicted by points where the bonds connect (for carbon atoms) or letters that indicate the type of atom (as in the periodic table).

A molecular graph representation is a 2D object that may represent 3D data, such as bond angles, chirality, and atomic coordinates. It is necessary to express any geographic links between the nodes as node and/or edge attributes because nodes in a graph, which is a mathematical object, only have pairwise relationships rather than formal spatial positions [10]. Numerous software programs, such as ChemDraw [22], Mercury [23], Avogadro [24], VESTA [25], and VMD [26], may readily visualize 2D and 3D graph representations.

## IV. ML and DL models

Animal models can be used in experiments to test a chemical's toxicity, but these investigations are expensive and time-consuming. As a result, ML and DL have emerged as desirable methods for assessing chemical toxicity.

Figure 3 illustrates the basic processes of the ML and DL modeling that consist of (a) defining questions, (b) collecting data, (c) preprocessing data, (d) splitting data, (e) training models, (f) optimizing models, (g) evaluating models, (h) interpreting models, (i) deploying models [27].

Regression and classification models are the two different categories of machine learning models. Classification models are based on categorical toxicity values, whereas regression models are based on quantitative toxicity values like LD50 and LC50. Classification models are more often used in the field of predictive toxicology. Toxicology prediction models have been created using various ML and DL methods, including SVM, RF, KNN, and neural networks (NN) [28]. The ML and DL algorithms employed in the published toxicity prediction models are included in Table 1.

Support Vector Machines (SVM): The SVM computes the best linear decision boundary for class separation based on a sample that defines class borders. The SVM technique separates data by generating a hypersurface through linear classification. The SVM model successfully identified the optimal hypersurface for distinguishing two classes by a significant margin. SVM is a prominent machine learning technique that can be used for classification and regression. SVM is commonly used in classification problems, recognizing hyperplanes to enhance the margin between classes [29].

Decision Tree (DT): the decision tree structure is similar to a tree's. Trees consist of roots, nodes, branches, and leaves. The decision tree shares structure with decision nodes, leaf nodes, and branches. The dataset is divided into decreasing entropy levels by the leaf and terminal nodes, which display

the class label of a tree for final prediction. The tree began with a route node and ended by splitting nodes [30].

KNN- is one of the simplest machine learning algorithms. It predicts a chemical's activity by selecting the k chemicals with the shortest distances from it in a chemical space represented by a collection of descriptors. Based on majority voting, compounds are often assigned to the class with the highest number of k-nearest chemicals in classification. The KNN technique is simple to grasp and produces highly interpretable prediction models [28].

A subclass of machine learning methods known as "deep learning algorithms" uses multi-layered neural networks—thus the name "deep"—to simulate intricate patterns in huge datasets. Because these algorithms can automatically extract features and build hierarchical representations of data, they have attracted much interest. This makes them especially useful for tasks like natural language processing, picture and audio recognition, and more. Figure 4 presents a neural network with possible inputs. The order of the input, from up to down, concord with the bias introduced by the user [31].

Important architectures for deep learning are: CNNs, or convolutional neural networks:

CNNs are deep learning types that function similarly to feed-forward neural networks. CNN uses rotational and transitional approaches for data analysis. The input data for the convolutional operation is applied to this neural layer. It additionally filters the input data. This network automatically trained itself based on features and patterns [32]. This algorithm specifically handles and examines visual data, such as pictures and videos.

RNNs (Recurrent Neural Networks): Because RNNs are built for sequential data, they are perfect for time series prediction and language modeling applications. Preserve a memory of past inputs through hidden states, enabling them to record temporal dependencies [32].

Generative Adversarial Networks (GANs): These networks are composed of a discriminator and a generator that are trained concurrently. The discriminator assesses the data samples that the generator produces, creating highly realistic [33].

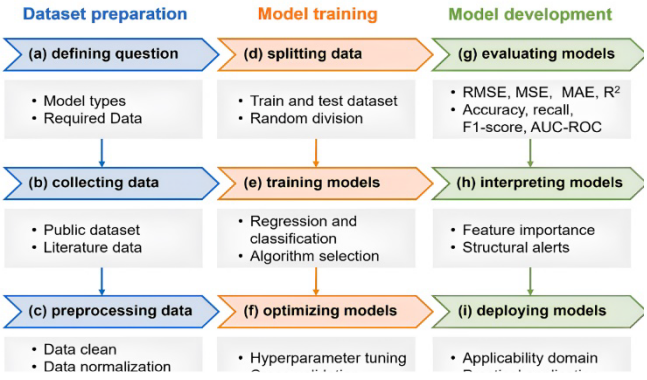


**FIGURE 3.** **Basic processes of the ML and DL modeling.**

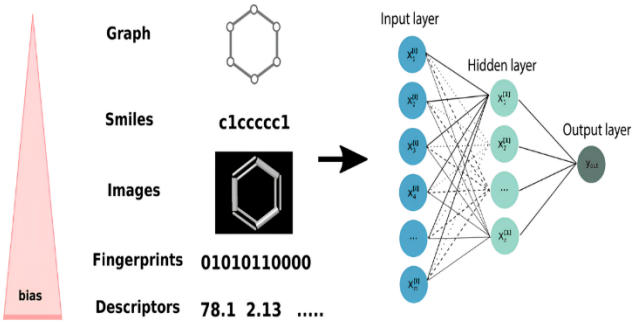


**FIGURE 4.** **A neural network with possible molecular inputs.**

## V. Challenges of AI in Toxicity Prediction

Even with artificial intelligence's potential, there are still a lot of obstacles standing in the way of its ultimate realization. These difficulties can be methodically divided into five main areas: data, interpretability, model creation, generalizability, and tool-related obstacles.

### *A. Data-related barriers*

Developing the best prediction models is hampered by the limited availability and inconsistent quality of toxicity data. It is possible to explain this difference by looking at the different experimental strategies used to address different toxicity data points, including liver toxicity, toxicity resulting from pharmacokinetic features, idiosyncratic reactions, and cardiogenic toxicity. Toxicological datasets such as Tox21 and ToxCast frequently suffer from data imbalances. The overrepresentation of benign chemicals in these databases is a characteristic. The unbalanced outcome has led to a bias in the models identifying compounds as harmless, so ignoring potentially hazardous chemicals and causing a serious risk to human health. Also, this might influence regulatory choices leading to extended public exposure to harmful substances and long-term health effects.

Predictions may also be based on an over-reliance on molecular descriptors, such as structural traits and physicochemical characteristics, to represent toxicity data. Structural features determine toxicophore qualities, while physicochemical aspects explain ADME (Absorption, Distribution, Metabolism, and Excretion) characteristics. They might, however, be unable to adequately convey the biological dynamics and interactions of in vivo settings [66]. Prediction errors may arise especially when toxicity results from interactions with the biological system or biotransformation produces toxic metabolites. High feature counts can also result in issues including data sparsity, overfitting, and higher processing demands. The "curse of dimensionality" refers to this phenomenon, in which excessive data dimensionality has a negative effect on model performance.

### *B. Models-related barriers*

Traditional methods, such as decision trees and support vector machines, are different from sophisticated deep learning (DL) models in terms of data handling and representations regarding toxicity prediction. Conventional models cannot capture complex toxicological patterns and interactions since they rely on manually chosen weighted features. On the other hand, DL models analyze data hierarchically, enabling them to recognize minute toxicological patterns. Additionally, task-specific characteristics that indicate toxicophoric and are essential for toxicity predictions can be produced using DL. Traditional models are simple to use, but not accurate enough for complex toxicological scenarios. While more accurate, DL models have drawbacks because of their great computational complexity.

New developments in toxicity prediction include learning strategies based on graphs. The DeepTox pipeline [67] is a prime example of these techniques, effectively improving toxicity data representations and exhibiting considerable gains in predictive efficiency. Attaining model robustness is still difficult, though. Scenarios with "toxicity cliffs" make this clear. A study on substituted phenols illustrates these situations well since small molecular changes such as fluoro- or bromo-substitution resulted in significantly varied toxicity [68, 69]. Certain models may not be able to identify these subtle structural alterations, leading to notable differences in their toxicity predictions.

### *C. Generalizability-related barriers*

Because in vivo biological systems are complicated and have a wide range of toxicity endpoints, generalizability in machine learning models for toxicity prediction is difficult. By highlighting the various interactions that occur within biological systems, these endpoints evaluate the toxicity profiles of the substances. Models are usually designed for particular endpoints, and each one needs a different set of data parameters. Thus, a model that works well for one kind of toxicity may not work well for another [70].

### *D. Interpretability-related barriers:*

While ML models' accuracy is important, their prediction transparency is just as important. Functioning frequently as "black boxes," particularly in DL models with intricate structures. The intricate topologies of these models, with numerous layers of interconnected nodes that independently choose features for predictions, make it difficult to understand the logic underlying their conclusions. Even the model developers cannot fully understand these models' internal workings. Decisions made in predictive toxicology, where choices have broad effects on drug development and public health, require a grasp of the ratio of underlying toxicity projections. Without unambiguous interpretability, scientists may find it difficult to respond to important queries such as "How toxic is this compound?" or "Does the found toxicity justify excluding the compound, or does it still meet the criteria for potential hit?" For toxicologists and chemists, this knowledge is essential, especially when the substance under investigation has great promise for future drug discovery. It may be possible to modify the substance to reduce toxicity while maintaining its therapeutic qualities by having a thorough knowledge of the rationale for the modality [70]. Furthermore, in regulatory contexts, choosing too cautiously and maybe missing out on useful substances can arise from an inability to understand toxicity model predictions.

### *E. Tool-related barriers:*

When predicting toxicity, the decision between commercial and open-access techniques is crucial. Although open-access technologies are widely available, their usability, dependability, and predictability are typically compromised by their lack of advanced predictive capabilities and user-friendly services. On the other hand, commercial tools are more expensive and may not be entirely transparent, but they are also more effective and user focused. The effectiveness and safety of drug development are greatly impacted by this decision [71].

To maximize the drug discovery process and lower the chance of late-stage drug withdrawals, the main problem is striking a balance between the accuracy of commercial tools and the accessibility of open-access technologies [72].

## VI. CONCLUSION

Estimating the toxicity of drug candidates is crucial in drug discovery, as it can lead to high costs, failures in later phases, and withdrawals. Based on the existing data, ML and DL models could be viable methods for serving as early filters of dangerous chemicals during the drug discovery process, even with their ongoing limitations discussed in this review. This review highlights a brief Overview of chemical toxicity and drug processes. We focus on the recent progress and outstanding challenges in the area, describing the state-of-the-art models implemented for chemical toxicity prediction. The type of molecular representation, ML and DL algorithms are also explained.

TABLE I
SUMMARIZES CURRENT ML AND DL MODELS FOR PREDICTIVE TOXICOLOGY

| Ref# | Objectives | Methods Used | Limitations | year |
|---|---|---|---|---|
| [34] | Predict toxicity of chemicals using molecular similarity and machine-learning models. Provide a freely available computational platform for toxicity prediction. | Molecular similarity and machine-learning models. In silico methods for toxicity prediction. | No limitations mentioned. | 2024 |
| [35] | Develop a new approach method for chemical-mixture toxicity assessment. Integrate AI with pathophysiology to predict toxicity mechanisms comprehensively. | AI-HNN and CPTM for toxicity prediction. Integration of AI-HNN and CPTM into AI-CPTM framework. | Absence of comprehensive methods for toxicity assessment of mixtures. Standalone AI models have limitations in identifying toxic chemicals. | 2024 |
| [36] | Identify chemical substructures responsible for hidden neuron activation. Explain model predictions using associated substructures for individual compounds. | SHAP and integrated gradients for feature attribution. Novel technique identifies chemical substructures activating hidden neurons. | Understanding predictions made by complex neural network models is difficult. Current techniques do not explain how compounds are transformed in layers. | 2024 |
| [37] | Develop 3MTox model for toxicity identification. Achieve state-of-the-art performance on toxicity benchmark datasets. | 3MTox model with BERT backbone and motif graph input. Classical ML and DL methods for toxicity prediction | Over-reliance on artificial features. Easy overfitting with classical ML and DL methods | 2024 |
| [38] | Develop deep learning models for predicting compound toxicity. Integrate models for virtual screening of low-toxicity drug candidates. | Graph convolutional network (GCN) regression model for acute toxicity prediction. Multiple GCN binary classification models for different toxicity types | Addressed data size, label type, and distribution variations. No specific limitations mentioned in the abstract section | 2024 |
| [39] | Review deep learning for predictive toxicology assessment. Highlight early detection of adverse drug reactions. | Deep learning techniques. Traditional methods like animal testing. | Limited data availability and quality for training deep learning models. Imbalanced toxicity datasets can lead to biased model performance. | 2024 |
| **[40]** | Develop hybrid quantum-classical neural network for drug toxicity prediction. Transfer learnable weights from quantum to classical devices | Quantum-classical neural network for drug toxicity prediction. Hadamard test for efficient inner product estimation in quantum computing | Noisy intermediate-scale quantum devices face decoherence and gate errors. Quantum-classical neural network aims to address computational complexity challenges. | 2024 |
| [41] | Develop hybrid model to predict chemical hepatotoxicity. Improve risk assessments for environmental and health concerns | Hybrid model combining in vitro assay and chemical structures. Machine learning for quantitative structure-activity relationship (QSAR) modeling | Predictivity of hepatotoxicity model initially at 0.59. Improved to 0.8 with inclusion of 37 structural alerts | 2024 |
| [42] | Develop models for predicting ocular toxicity of chemicals. Enhance model interpretability using SHAP and attention weights analysis. | Machine learning and deep learning algorithms. SHAP method and attention weights analysis | Reliance on data quality and quantity for model performance. Importance of balancing data quality and model interpretability. | 2024 |
| [43] | Predict acute dermal toxicity using machine learning. Identify important features and structural fragments associated with toxicity. | Machine learning and deep learning algorithms. SARpy, Shapley additive explanation, attentive FP heatmap | No direct identification of acute dermal toxicity through animal experiments. Difficulty in assessing acute dermal toxicity of potential compounds. | 2024 |
| [44] | Assess in silico chemical toxicity prediction for occupational cancer prevention. Develop predictive models for different toxicities in recent years. | Machine learning techniques for quantitative regression and qualitative classification studies. Development of predictive models for different toxicities in recent years. | Obstacles and shortcomings in drug safety assessment. Enhancements needed for future drug safety assessment. | 2024 |
| [45] | Develop a problem formulation framework for in silico toxicology. Identify gaps and inconsistencies in in silico toxicology problem formulations. | Developed a problem formulation framework. Modified and applied a PF framework from risk assessment literature | PFs for in silico toxicology lack consistency in components. PFs need to address higher-level conceptual questions. | 2024 |

| Ref# | Objectives | Methods Used | Limitations | year |
|---|---|---|---|---|
| [46] | Develop small molecule toxicity prediction model. Improve efficiency of research and development in drug design. | Graph attention network model proposed. Attention mechanism used to mine connection relationships between atoms | Traditional ML methods can't use molecules directly as inputs. Difficulty in accurately extracting molecular features. | 2023 |
| [28] | Summarize machine learning and deep learning toxicity prediction models. Highlight importance of dataset quality for model performance. | Machine learning algorithms: SVM, random forest, k-NN, ensemble. Deep learning algorithms: Neural network | Traditional toxicity assays are complicated, costly, and time-consuming. Different datasets may impact model performance in toxicity prediction. | 2023 |
| [47] | Investigate EGNNs for toxicity prediction using 3D molecular structures. Enhance ML models for toxicity prediction with 3D geometry information. | Equivariant Graph Neural Networks (EGNNs). Equivariant transformer (ET) model in TorchMD-NET | Physicochemical property total energy not related to toxicity prediction. No direct relationship between 3D molecule representations and toxicity. | 2023 |
| [48] | Predict toxicity accurately using multi-task deep learning model. Provide contrastive molecular explanations for toxicity predictions. | Multi-task deep learning model for toxicity prediction. Contrastive explanation method for model predictions | Highly skewed ClinTox test set with few 'toxic' molecules. Minimal recovery of known toxicophores for clinical endpoints. | 2022 |
| [49] | Develop predictive models for drug toxicity. Integrate informatics and biology to build foundation for predictions. | Integration of informatics and biology. Development of DTox for predicting drug toxicity | Lack of comprehensive data on drug toxicity. Challenges in integrating diverse data sources | 2022 |
| [50] | Analyze EPA CompTox Chemical Dashboard data and tools. Evaluate utility for next generation risk assessment and toxicity prediction | NAMs: in chemico, in silico, in vitro approaches. Data curation, predictive tools, case studies discussed | Limited availability of experimental data for some substances. Challenges in integrating diverse data sources for predictive modeling. | 2023 |
| [51] | Predict broad toxicities using induced pluripotent stem cells. Achieve high accuracy in toxicity predictions for various categories | ES cell gene networks combined with developmental toxicity testing. Transfer learning from ES cell data to predict toxicities | Animal testing limited applicability to humans. Desire for effective alternatives due to animal protection concerns. | 2021 |
| [52] | Link mass spectra of chemicals to toxicity endpoints. Utilize machine learning and experimentation for analysis. | Machine learning. Experimentation. | Limited number of chemicals studied. Machine learning model performance variability | 2022 |
| [53] | Improve compound toxicity prediction using GCN and SSL. Investigate if GCN is superior to other ML methods. | Semi-supervised learning (SSL) algorithms. Graph Convolution Neural Network (GCN) with Mean Teacher (MT) SSL | Limited availability of annotated toxicity data. Time-consuming and costly traditional toxicity testing methods | 2021 |
| [54] | Develop models to predict respiratory toxicity of chemicals. Identify significant molecular descriptors for accurate predictions | Eight machine learning models utilized for prediction. Methods include SVM, MLP, XGB, RF, LR, ABDT, KNN, NB | Lack of data availability statement. No conflict of interest declaration provided | 2022 |
| [55] | Develop ML models for toxicology predictions. Implement consensus approach to improve predictive performance | VenomPred platform employs in-house Machine Learning models. Consensus approach combining results of different ML models used. | In vitro and in vivo methods limited by resources. Ethics, time, budget constraints in toxicity prediction methods. | 2022 |
| [56] | Compare traditional and deep learning approaches for toxicity prediction. Evaluate performance of different models on toxicity datasets. | Traditional physico-chemical descriptor and machine learning-based approaches. Descriptor-free, SMILES-based, deep learning BERT architectures | No limitations mentioned in the paper. | 2022 |
| [57] | Predict pharmacokinetics and toxicity for diverse targets. Support molecular optimization and interpretation for input molecules. | Graph neural networks and graph-based signatures. Deep learning-based pharmacokinetic and toxicity prediction platform | Current methods limited in diversity, accuracy, interpretability, and extensibility. Challenges in providing pharmacokinetics and toxicity for diverse targets. | 2024 |
| [58] | Develop machine learning model for toxic dispersion casualty prediction. | Machine learning based quantitative property- | - | 2022 |

| Ref# | Objectives | Methods Used | Limitations | year |
|---|---|---|---|---|
| | Estimate affected area caused by toxic chemical release accurately. | consequence relationship model. Toxic dispersion casualty database construction for prediction model development | | |
| [59] | Develop predictive toxicity models for nonionic alcohol ethoxylate surfactants. Estimate chronic HC5s using acute-to-chronic ratio and regressions | ICE models for toxicity prediction. Comparison of acute and chronic HC5s | - | 2021 |
| [60] | Predict chemical respiratory toxicity using machine learning. Develop in silico models for toxicity assessment. | Machine learning. In silico prediction. | Limited by the availability of high-quality data. Performance may vary based on different chemical classes. | 2021 |
| [61] | Develop predictive computational model for developmental toxicants. Create alternative chemical developmental toxicity evaluations. | Computational modeling. Data extraction from PubChem and the ToxCast program | Animal testing guidelines are costly, time-consuming, and require many animals. Computational modeling aims to provide a cost-effective alternative. | 2022 |
| [62] | Predict reproductive toxicity of chemicals using ensemble learning methods. Utilize molecular fingerprints for predictive modeling of reproductive toxicity. | Ensemble learning methods. Molecular fingerprints for predicting reproductive toxicity of chemicals | Limited by the quality of available data sources. Challenges in interpreting complex interactions between chemicals and toxicity. | 2021 |
| [63] | Propose FS method for class-imbalance datasets in toxicity prediction. Boosting with fast clustering-based FS and fast correlation-based filter. | Boosting-based feature selection ensemble. Fast clustering-based FS and fast correlation-based filter | Class-imbalance damages feature selection performance for QSAR models. Standard methods less efficient compared to proposed FS method. | 2021 |
| [64] | Develop hybrid deep learning model for toxicity prediction. Prioritize chemicals for experimental testing accurately | Hybrid neural network (HNN) deep learning model. Ensemble methods: Random Forest, Bagging, Adaboost | Resource-intensive to assess chemicals in-vivo. Limited evaluation of chemicals in commercial use. | 2021 |
| [65] | Predict toxicity risk after chemotherapy sessions for individual patients. Classify chemotherapy induced complications based on predefined toxicity levels. | Machine learning models. Predicting toxicity levels in chemotherapy patients. | Imbalance of chemical balance may lead to severe complications. Side effects depend on various factors like drug type and dose. | 2022 |

## REFERENCES


[1] E. P´erez Sant´ın, R. Rodr´ıguez Solana, M. Gonz´alez Garc´ıa, M. D. M. Garc´ıa Su´arez, G. D. Blanco D´ıaz, M. D. Cima Cabal, J. M. Moreno Rojas, and J. I. L´opez S´anchez, "Toxicity prediction based on artificial intelligence: A multidisciplinary overview," Wiley Interdisciplinary Reviews: Computational Molecular Science, vol. 11, no. 5, p.e1516, 2021.

[2] A. B. Raies and V. B. Bajic, "In silico toxicology: computational methods for the prediction of chemical toxicity," Wiley Interdisciplinary Reviews: Computational Molecular Science, vol. 6, no. 2, pp.147–172, 2016.

[3] H. Hong, S. Thakkar, M. Chen, and W. Tong, "Development of decision forest models for prediction of drug-induced liver injury in humans using a large set of fda-approved drugs," Scientific reports, vol. 7, no. 1, p. 17311, 2017.

[4] Y. Huang, X. Li, S. Xu, H. Zheng, L. Zhang, J. Chen, H. Hong, R. Kusko, and R. Li, "Quantitative structure–activity relationship models for predicting inflammatory potential of metal oxide nanoparticles," Environmental health perspectives, vol. 128, no. 6, p. 067010, 2020.

[5] H. Hong, J. Zhu, M. Chen, P. Gong, C. Zhang, and W. Tong, "Quantitative structure–activity relationship models for predicting risk of drug-induced liver injury in humans," Drug-induced liver toxicity, pp. 77–100, 2018.

[6] H. Kim, E. Kim, I. Lee, B. Bae, M. Park, and H. Nam, "Artificial intelligence in drug discovery: a comprehensive review of data-driven and machine learning approaches," Biotechnology and Bioprocess Engineering, vol. 25, pp. 895–930, 2020.

[7] R. Dhudum, A. Ganeshpurkar, and A. Pawar, "Revolutionizing drug discovery: A comprehensive review of ai applications," Drugs and Drug Candidates, vol. 3, no. 1, pp. 148–171, 2024.

[8] H. S. Chan, H. Shan, T. Dahoun, H. Vogel, and S. Yuan, "Advancing drug discovery via artificial intelligence," Trends in pharmacological sciences, vol. 40, no. 8, pp. 592–604, 2019.

[9] M. J. Lamberti, M. Wilkinson, B. A. Donzanti, G. E. Wohlhieter, S. Parikh, R. G. Wilkins, and K. Getz, "A study on the application and use of artificial intelligence to support drug development," Clinical therapeutics, vol. 41, no. 8, pp. 1414–1426, 2019.

[10] L. David, A. Thakkar, R. Mercado, and O. Engkvist, "Molecular representations in ai-driven drug discovery: a review and practical guide," Journal of Cheminformatics, vol. 12, no. 1, p. 56, 2020.

[11] N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, and G. R. Hutchison, "Open babel: An open chemical toolbox," Journal of cheminformatics, vol. 3, pp. 1–14, 2011.

[12] Y. Cao, A. Charisi, L.-C. Cheng, T. Jiang, and T. Girke, "Chemminer: a compound mining framework for r," Bioinformatics, vol. 24, no. 15, pp. 1733–1734, 2008.

[13] F. Grisoni, D. Ballabio, R. Todeschini, and V. Consonni, "Molecular descriptors for structure–activity applications: a hands-on approach," Computational toxicology: methods and protocols, pp. 3–53, 2018.

[14] D. Weininger, "Smiles, a chemical language and information system. 1. introduction to methodology and encoding rules," Journal of chemical information and computer sciences, vol. 28, no. 1, pp. 31–36, 1988.

[15] A. Capecchi, D. Probst, and J.-L. Reymond, "One molecular fingerprint to rule them all: drugs, biomolecules, and the metabolome," Journal of cheminformatics, vol. 12, pp. 1–15, 2020.
[16] M. Seo, H. K. Shin, Y. Myung, S. Hwang, and K. T. No, "Development of natural compound molecular fingerprint (NC-MFP) with the dictionary of natural products (DNP) for natural product-based drug development," Journal of Cheminformatics, vol. 12, no. 1, p. 6, 2020.
[17] D. Rogers and M. Hahn, "Extended-connectivity fingerprints," Journal of chemical information and modeling, vol. 50, no. 5, pp. 742–754, 2010.
[18] H. Matter and T. P¨otter, "Comparing 3d pharmacophore triplets and 2d fingerprints for selecting diverse compound subsets," Journal of chemical information and computer sciences, vol. 39, no. 6, pp. 1211–1225, 1999.
[19] S. Yin, E. A. Proctor, A. A. Lugovskoy, and N. V. Dokholyan, "Fast screening of protein surfaces using geometric invariant fingerprints," Proceedings of the National Academy of Sciences, vol. 106, no. 39, pp. 16 622–16 626, 2009.
[20] D. J. Wood, J. d. Vlieg, M. Wagener, and T. Ritschel, "Pharmacophore fingerprint-based approach to binding site subpocket similarity and its application to bioisostere replacement," Journal of chemical information and modeling, vol. 52, no. 8, pp. 2031–2043, 2012.
[21] W. Chen, X. Liu, S. Zhang, and S. Chen, "Artificial intelligence for drug discovery: Resources, methods, and applications," Molecular Therapy-Nucleic Acids, vol. 31, pp. 691–702, 2023.
[22] ChemDraw. PerkinElmer Informatics.
[23] C. F. Macrae, I. Sovago, S. J. Cottrell, P. T. Galek, P. McCabe, E. Pidcock, M. Platings, G. P. Shields, J. S. Stevens, M. Towler et al., "Mercury 4.0: From visualization to analysis, design and prediction," Applied Crystallography, vol. 53, no. 1, pp. 226–235, 2020.
[24] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, and G. R. Hutchison, "Avogadro: an advanced semantic chemical editor, visualization, and analysis platform," Journal of cheminformatics, vol. 4, pp. 1–17, 2012.
[25] K. Momma and F. Izumi, "Vesta 3 for three-dimensional visualization of crystal, volumetric and morphology data," Journal of applied crystallography, vol. 44, no. 6, pp. 1272–1276, 2011.
[26] W. Humphrey, A. Dalke, and K. Schulten, "Vmd: visual molecular dynamics," Journal of molecular graphics, vol. 14, no. 1, pp. 33–38, 1996.
[27] S. Cui, Y. Gao, Y. Huang, L. Shen, Q. Zhao, Y. Pan, and S. Zhuang, "Advances and applications of machine learning and deep learning in environmental ecology and health," Environmental Pollution, p. 122358, 2023.
[28] W. Guo, J. Liu, F. Dong, M. Song, Z. Li, M. K. H. Khan, T. A. Patterson, and H. Hong, "Review of machine learning and deep learning models for toxicity prediction," Experimental Biology and Medicine, vol. 248, no. 21, pp. 1952–1973, 2023.
[29] S. M. A. Mahnaz Ahmadi and F. Ghorbani-Bidkorpeh, "Toxicity prediction of nanoparticles using machine learning approaches," Toxicology, vol. 501, 2024.
[30] M. Maniruzzaman, M. J. Rahman, B. Ahammed, and M. M. Abedin, "Classification and prediction of diabetes disease using machine learning paradigm," Health information science and systems, vol. 8, pp. 1–14, 2020.
[31] J. Hemmerich and G. F. Ecker, "In silico toxicology: From structure–activity relationships towards deep learning and adverse outcome pathways," Wiley Interdisciplinary Reviews: Computational Molecular Science, vol. 10, no. 4, p. e1475, 2020.
[32] Jessi J and White, "New methods for predicting drug molecule activity using deep learning," Bioscience Methods, vol. 15, no. 1, pp. 28–36, 2024.
[33] F. Mostafa and M. Chen, "Computational models for predicting liver toxicity in the deep learning era," Frontiers in Toxicology, vol. 5, p.1340860, 2024.
[34] P. Banerjee, E. Kemmler, M. Dunkel, and R. Preissner, "Protox 3.0: a webserver for the prediction of toxicity of chemicals," Nucleic Acids Research, p. gkae303, 2024.
[35] S. Limbu, E. Glasgow, T. Block, and S. Dakshanamurthy, "A machine learning-driven pathophysiology-based new approach method for the dose-dependent assessment of hazardous chemical mixtures and experimental validations," Toxics, vol. 12, no. 7, p. 481, 20241.
[36] M. Walter, S. J. Webb, and V. J. Gillet, "Interpreting neural network models for toxicity prediction by extracting learned chemical features," Journal of Chemical Information and Modeling, vol. 64, no. 9, pp. 3670–3688, 2024.
[37] Y. Zhu, Y. Zhang, X. Li, and L. Wang, "3mtox: A motif-level graph based multi-view chemical language model for toxicity identification with deep interpretation," Journal of Hazardous Materials, vol. 476, p. 135114, 2024.
[38] K. M. Saravanan, J.-F. Wan, L. Dai, J. Zhang, J. Z. Zhang, and H. Zhang, "A deep learning based multi-model approach for predicting drug-like chemical compound's toxicity," Methods, vol. 226, pp. 164–175, 2024.
[39] B. I. Priyadarshini, "Deep learning for predictive toxicology assessment early detection of adverse drug reactions." Power System Technology, vol. 48, no. 1, pp. 680–697, 2024.
[40] A. M. Smaldone and V. S. Batista, "Quantum-to-classical neural network transfer learning applied to drug toxicity prediction," Journal of Chemical Theory and Computation, 2024.
[41] E. Chung, X. Wen, X. Jia, H. L. Ciallella, L. M. Aleksunes, and H. Zhu, "Hybrid non-animal modeling: A mechanistic approach to predict chemical hepatotoxicity," Journal of Hazardous Materials, vol.471, p. 134297, 2024.
[42] Y. Zhou, Z. Wang, Z. Huang, W. Li, Y. Chen, X. Yu, Y. Tang, and G. Liu, "In silico prediction of ocular toxicity of compounds using explainable machine learning and deep learning approaches," Journal of Applied Toxicology, 2024.
[43] S. Lou, Z. Yu, Z. Huang, H. Wang, F. Pan, W. Li, G. Liu, and Y. Tang, "In silico prediction of chemical acute dermal toxicity using explainable machine learning methods," Chemical Research in Toxicology, vol. 37, no. 3, pp. 513–524, 2024.
[44] P. M. S and D. V. N, "In-silico models in molecular property and toxicity assessments," Novel Aspects on Chemistry and Biochemistry, vol. 9, 2024.
[45] J. Achar, M. T. Cronin, J. W. Firman, and G. ¨Oberg, "A problem formulation framework for the application of in silico toxicology methods in chemical risk assessment," Archives of Toxicology, vol. 98, no. 6, pp. 1727–1740, 2024.
[46] Y. Tong, Q. Guo, J. Cui, and X. Peng, "Toxicity prediction study of small molecules based on graph attention networks," in 2023 2nd International Conference on Artificial Intelligence, Human-Computer Interaction and Robotics (AIHCIR), pp. 603–607, 2023.
[47] J. Cremer, L. Medrano Sandonas, A. Tkatchenko, D.-A. Clevert, and G. De Fabritiis, "Equivariant graph neural networks for toxicity prediction," Chemical Research in Toxicology, vol. 36, no. 10, pp.1561–1573, 2023.
[48] B. Sharma, V. Chenthamarakshan, A. Dhurandhar, S. Pereira, J. A. Hendler, J. S. Dordick, and P. Das, "Accurate clinical toxicity prediction using multi-task deep neural nets and contrastive molecular explanations," Scientific Reports, vol. 13, no. 1, p. 4908, 2023.
[49] M. J. Sniatynski and B. S. Kristal, "Predicting drug toxicity at the intersection of informatics and biology: Dtox builds a foundation," Patterns, vol. 3, no. 9, 2022.
[50] D. Kim and J. Choi, "Utility of EPA comptox chemical dashboard for next generation risk assessment and toxicity prediction of chemicals," J. Korean Soc. Environ. Eng, vol. 45, no. 5, pp. 244–255, 2023.
[51] J. Yamane, T. Wada, H. Otsuki, K. Inomata, M. Suzuki, T. Hisaki, S. Sekine, H. Kouzuki, K. Kobayashi, H. Sone et al., "Prediction of broad chemical toxicities using induced pluripotent stem cells and gene networks by transfer learning from embryonic stem cell data," bioRxiv, pp. 2021–11, 2021.
[52] S. Hu, G. Liu, J. Zhang, J. Yan, H. Zhou, and X. Yan, "Linking electron ionization mass spectra of organic chemicals to toxicity endpoints through machine learning and experimentation," Journal of Hazardous Materials, vol. 431, p. 128558, 2022.
[53] J. Chen, Y.-W. Si, C.-W. Un, and S. W. Siu, "Chemical toxicity prediction based on semi-supervised learning and graph convolutional neural network," Journal of cheminformatics, vol. 13, pp. 1–16, 2021.
[54] K. Jaganathan, H. Tayara, and K. T. Chong, "An explainable supervised machine learning model for predicting respiratory toxicity

of chemicals using optimal molecular descriptors," Pharmaceutics, vol. 14, no. 4, p. 832, 2022.
[55] S. Galati, M. Di Stefano, E. Martinelli, M. Macchia, A. Martinelli, G. Poli, and T. Tuccinardi, "Venompred: A machine learning based platform for molecular toxicity predictions," International Journal of Molecular Sciences, vol. 23, no. 4, p. 2105, 2022.
[56] U. Norinder, "Traditional machine and deep learning for predicting toxicity endpoints," Molecules, vol. 28, no. 1, p. 217, 2022.
[57] Y. Myung, A. G. de S´a, and D. B. Ascher, "Deep-pk: deep learning for small molecule pharmacokinetic and toxicity prediction," Nucleic Acids Research, p. gkae254, 2024.
[58] Z. Jiao, Z. Zhang, S. Jung, and Q. Wang, "Machine learning based quantitative consequence prediction models for toxic dispersion casualty," Journal of Loss Prevention in the Process Industries, vol. 81, p. 104952, 2023.
[59] A. C. Bejarano and J. R. Wheeler, "Predictive toxicity models for chemically related substances: a case study with nonionic alcohol ethoxylate surfactant," Environmental Toxicology and Chemistry, vol. 40, no. 7, pp. 2071–2080, 2021.
[60] Z. Wang, P. Zhao, X. Zhang, X. Xu, W. Li, G. Liu, and Y. Tang, "In silico prediction of chemical respiratory toxicity via machine learning," Computational Toxicology, vol. 18, p. 100155, 2021.
[61] [61] H. L. Ciallella, D. P. Russo, S. Sharma, Y. Li, E. Sloter, L. Sweet, H. Huang, and H. Zhu, "Predicting prenatal developmental toxicity based on the combination of chemical structures and biological data," Environmental science & technology, vol. 56, no. 9, pp. 5984–5998, 2022.
[62] H. Feng, L. Zhang, S. Li, L. Liu, T. Yang, P. Yang, J. Zhao, I. T. Arkin, and H. Liu, "Predicting the reproductive toxicity of chemicals using ensemble learning methods and molecular fingerprints," Toxicology Letters, vol. 340, pp. 4–14, 2021.
[63] A. Antelo-Collado, R. Carrasco-Velar, N. Garc´ıa-Pedrajas, and G. Cerruela-Garc´ıa, "Effective feature selection method for class-imbalance datasets applied to chemical toxicity prediction," Journal of Chemical Information and Modeling, vol. 61, no. 1, pp. 76–94, 2020.
[64] S. Limbu, C. Zakka, and S. Dakshanamurthy, "Predicting environmental chemical toxicity using a new hybrid deep machine learning method," 2021.
[65] I. Boudali and I. B. Messaoud, "Machine learning models for toxicity prediction in chemotherapy," in International Conference on Intelligent Systems Design and Applications, 2022, pp. 350–364.
[66] A. Lysenko, A. Sharma, K. A. Boroevich, and T. Tsunoda, "An integrative machine learning approach for prediction of toxicity related drug safety," Life science alliance, vol. 1, no. 6, 2018.
[67] A. Mayr, G. Klambauer, T. Unterthiner, and S. Hochreiter, "DeepTox: Toxicity Prediction using Deep Learning," Frontiers in Environmental Science, vol. 3, p. 80, 2016.
[68] Y. Y. Pang, W. K. Yeo, K. Y. Loh, M. L. Go, and H. K. Ho, "Structure–toxicity relationship and structure–activity relationship study of 2-phenylaminophenylacetic acid derived compounds," Food and chemical toxicology, vol. 71, pp. 207–216, 2014.
[69] M. Wen, Z. Zhang, S. Niu, H. Sha, R. Yang, Y. Yun, and H. Lu, "Deep-learning-based drug–target interaction prediction," Journal of proteome research, vol. 16, no. 4, pp. 1401–1409, 2017.
[70] A. Abou Hajal and A. Z. Al Meslamani, "Overcoming barriers to machine learning applications in toxicity prediction," Expert Opinion on Drug Metabolism & Toxicology, vol. 20, no. 7, pp. 549–553, 2024.
[71] T. T. V. Tran, A. Surya Wibowo, H. Tayara, and K. T. Chong, "Artificial intelligence in drug toxicity prediction: recent advances, challenges, and future perspectives," Journal of chemical information and modeling, vol. 63, no. 9, pp. 2628–2643, 2023.
[72] S. Yang and S. Kar, "Application of artificial intelligence and machine learning in early detection of adverse drug reactions (ADRs) and drug-induced toxicity," Artificial Intelligence Chemistry, p. 100011, 2023.

**Eman Shehab** is currently a Lecturer Assistant in the Department of Computer Science, Faculty of Computers and Artificial Intelligence, University of Sadat City, Egypt, since 2020. She achieved her master's degree from Menoufia University. She has worked on several research topics; her research interests are neural network and machine learning. She can be contacted at email: eman.shehab@fcai.usc.edu.eg.

**Hamada Nayel** is an Assistant Professor at the Department of Computer Science, Faculty of Computers and Artificial Intelligence, Benha University, Benha, Egypt. In 2019, he received his Ph.D. from Mangalore University, India. His research interests include Arabic NLP, biomedical NLP, and social media analysis. He can be contacted at email: hamada.ali@fci.bu.edu.eg

**Mohamed Taha** is an Associate Professor at Benha University, Faculty of Computers and Artificial intelligence, Computer Science Department, Egypt. He received his M.Sc. degree and his Ph.D. degree in computer science at Ain Shams University, Egypt, in February 2009 and July 2015. He is the founder and coordinator of" Networking and Mobile Technologies" program, Faculty of Computers and Artificial Intelligence, Benha University. His research interest's concern: Computer Vision (Object Tracking-Video Surveillance Systems), Digital Forensics (Image Forgery Detection – Document Forgery Detection - Fake Currency Detection), Image Processing (OCR), Computer Network (Routing Protocols - Security), Augmented Reality, Cloud Computing, and Data Mining (Association Rules Mining-Knowledge Discovery). Taha has contributed more than 30+ technical papers in international journals and conferences. He can be contacted at email: mohamed.taha@fci.bu.edu.eg.