Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

AgroChem : Predictive Training for AgroChemical Interaction Analysis using Deep Learning

[1] Sanjana S V, [2] Rashmi M, [3] Suhas J S, [4] Rohith P
Department of Computer Science and Engineering Malnad College of Engineering, Hassan, Karnataka, India

Kavyasri M N
Associate Professor
Department of Computer Science and Engineering Malnad College of Engineering, Hassan, Karnataka, India

***Abstract*—The increasing use of agrochemicals in modern agriculture has significantly improved crop productivity, but uncontrolled chemical interactions and excessive pesticide exposure continue to pose serious risks to environmental sustainability, soil fertility, water resources, and human health. Accurate prediction of agrochemical interactions and toxicity levels is therefore essential for ensuring safe agricultural practices and reducing ecological damage.**

In recent years, Machine Learning (ML) and Artificial Intelligence (AI) techniques have emerged as effective computational approaches for analysing complex agricultural and chemical datasets to identify harmful interaction patterns and predict chemical behaviour. Several ML algorithms including Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), Decision Tree (DT), Artificial Neural Networks (ANN), and Deep Learning methods have been explored for agrochemical safety prediction.

These models utilise features such as chemical composition, toxicity levels, environmental conditions, molecular descriptors, and crop-related parameters to estimate pesticide effectiveness and interaction risks. Data preprocessing techniques including normalization, feature selection, clustering, and dimensionality reduction improve prediction reliability and model performance.

This survey presents a comprehensive review of existing ML-based toxicity prediction systems, focusing on their methodologies, advantages, limitations, and research outcomes. The study highlights the importance of intelligent agricultural decision-support systems in promoting sustainable farming and improving agrochemical safety management.

***Index Terms*—Agrochemical Toxicity, Machine Learning, Deep Learning, Random Forest, Feature Selection, Clustering, Predictive Agriculture, Chemical Safety**

I. Introduction

Agrochemicals such as pesticides, herbicides, fungicides, and fertilizers play a major role in improving agricultural productivity and protecting crops from pests and diseases. However, excessive usage and harmful chemical interactions can negatively affect soil quality, biodiversity, groundwater resources, and human health. Predicting agrochemical toxicity and interaction behaviour is therefore essential for enabling safer agricultural practices and reducing environmental impact.

Traditional toxicity assessment methods primarily depend on laboratory testing and manual analysis, which are often expensive, time-consuming, and difficult to scale for large agricultural datasets. The rapid advancement of Artificial Intelligence (AI) and Machine Learning (ML) technologies has enabled the development of intelligent predictive systems capable of analysing complex toxicity-related datasets efficiently.

Machine Learning models such as Random Forest (RF), Artificial Neural Networks (ANN), Support Vector Machine (SVM), clustering algorithms, and Deep Learning architectures can identify hidden toxicity patterns and predict agrochemical behaviour using molecular descriptors and chemical properties. These systems improve prediction accuracy and reduce dependency on extensive laboratory testing.

Despite significant advancements, several challenges remain unresolved, including noisy and imbalanced datasets, limited feature selection methods, lack of explainable predictions, and limited deployment in real-world agricultural environments. This survey analyses existing ML-based toxicity prediction systems and identifies future research directions for intelligent agrochemical analysis.

II. Literature Survey

The application of Machine Learning techniques in toxicity prediction and agrochemical analysis has gained significant research attention due to the increasing demand for sustainable agriculture and intelligent chemical safety management.

Rawat et al. [1] proposed a Machine Learning-based toxicity prediction framework using algorithms such as Random Forest, Artificial Neural Networks (ANN), Bagging, K-Star, Decision Tree, and Linear Regression. Feature selection techniques and cross-validation were used to improve model performance and prediction accuracy. The study showed that Random Forest and Bagging algorithms achieved better performance for toxicity prediction.

The research presented in [2] focused on feature selection and clustering techniques for optimized toxicity prediction.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

Methods such as Principal Component Analysis (PCA), ReliefF, K-Means clustering, and Hierarchical Clustering were used to eliminate redundant features and improve prediction efficiency. The study highlighted that feature selection significantly improves dataset quality and model performance.

Sameh Zarif et al. [3] proposed toxicity prediction models using Machine Learning and Deep Learning algorithms such as ANN, CNN, LSTM, and SVM. The study demonstrated that Deep Learning architectures, particularly LSTM models, achieved high predictive accuracy by learning complex toxicity patterns from large datasets.

Rui Yao et al. [4] proposed a Random Forest-based feature selection approach for handling high-dimensional datasets. The method improved computational efficiency, prediction reliability, and interpretability by identifying important toxicity-related attributes and reducing unnecessary features.

Collectively, the reviewed studies demonstrate that Random Forest, feature selection techniques, clustering algorithms, and Deep Learning models provide strong predictive capabilities for toxicity analysis and intelligent agrochemical safety management. However, challenges such as preprocessing complexity, dataset imbalance, and computational requirements continue to motivate further research.

III. Comparative Analysis

Table I presents a structured comparison of the reviewed ML-based toxicity prediction systems, evaluating each against methodology, strengths, limitations, and research outcomes.

TABLE I: Comparative Analysis of Machine Learning-Based Toxicity Prediction Approaches

Title	Proposed Approach	Advantages	Limitations	Research Result
Performance Analysis of Drug Toxicity Prediction Using Machine Learning Approaches	Random Forest, ANN, Bagging, K-Star, Decision Tree, and Linear Regression for toxicity prediction and comparative evaluation.	High prediction accuracy and reduced overfitting.	Requires preprocessing and balanced datasets.	Random Forest achieved the best predictive performance.
Descriptive Analysis of Feature Selection and Clustering Algorithms for Optimized Drug Toxicity Prediction Model	PCA, ReliefF, K-Means, and Hierarchical Clustering for feature optimization and toxicity analysis.	Reduced redundant features and improved dataset quality.	High preprocessing complexity and computational cost.	Feature selection significantly improved prediction accuracy.
Advancing Toxicology Through Machine Learning and Deep Learning-Based Toxicity Prediction	ANN, CNN, LSTM, SVM, and KNN models for toxicity prediction and automatic feature extraction.	High accuracy and automatic feature learning from complex datasets.	Requires large datasets and high computational resources.	LSTM demonstrated the highest predictive accuracy.
Feature Selection Based on Random Forest for Partial Discharges Characteristic Set	Random Forest-based feature selection for identifying important toxicity-related attributes.	Improved computational efficiency and interpretability.	Requires parameter tuning and quality datasets.	Feature selection improved overall model performance.

IV. Proposed System

The proposed AgroChem system uses Machine Learning and Deep Learning techniques to predict agrochemical toxicity and pesticide effectiveness using molecular descriptors, chemical properties, and toxicity-related datasets.

The system applies preprocessing techniques such as data cleaning, normalization, feature selection, dimensionality reduction, and clustering to improve dataset quality and prediction efficiency. Random Forest-based feature selection is used to identify important toxicity-related attributes, while clustering algorithms help group chemicals based on toxicity behaviour.

The predictive models classify agrochemicals into Low, Moderate, and High toxicity levels. The system evaluates model performance using Accuracy, Precision, Recall, F1-score, Confusion Matrix, and visual analytics.

The dashboard provides decision-support insights through visual analytics for safer agrochemical usage and intelligent agricultural planning.

V. Preliminary Design Methodology

- **Data Collection:** Collect agrochemical datasets containing molecular descriptors, chemical properties, toxicity information, and pesticide-related attributes.
- **Data Preprocessing:** Apply data cleaning, normalization, feature selection, dimensionality reduction, and imbalance handling techniques.
- **Feature Selection and Clustering:** Identify important toxicity-related features and group chemicals using clustering algorithms.
- **Model Training:** Train Machine Learning and Deep Learning models such as Random Forest, ANN, SVM, CNN, and LSTM.
- **Prediction and Classification:** Predict agrochemical toxicity levels and pesticide effectiveness.
- **Decision Support:** Generate analytical insights and visual reports for safer agrochemical management.

VI. Conclusion

This survey paper examined the role of Machine Learning and Deep Learning techniques in predicting agrochemical toxicity and pesticide effectiveness for intelligent agricultural decision support. The reviewed systems demonstrate that algorithms such as Random Forest, ANN, CNN, SVM, and LSTM provide reliable predictive performance across toxicity-related datasets.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

Feature selection and clustering techniques improve dimensionality reduction, remove redundant features, and enhance prediction efficiency. However, challenges such as dataset imbalance, preprocessing complexity, computational requirements, and limited explainability remain important research concerns.

Future work should focus on developing lightweight, explainable, and real-time agrochemical prediction systems integrated with intelligent agricultural monitoring technologies. Greater emphasis on scalable datasets, automated analysis, and visual decision-support systems will help bridge the gap between research models and practical agricultural deployment.

References

[1] D. Rawat et al., "Performance Analysis of Drug Toxicity Prediction Using Machine Learning Approaches," 2023.

[2] D. Rawat et al., "Descriptive Analysis of Feature Selection and Clustering Algorithms for Optimized Drug Toxicity Prediction Model," 2023.

[3] Sameh Zarif et al., "Advancing Toxicology Through Machine Learning and Deep Learning-Based Toxicity Prediction," 2025.

[4] Rui Yao et al., "Feature Selection Based on Random Forest for Partial Discharges Characteristic Set," IEEE Access, 2020.