iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: IV Month of publication: April 2025

DOI: https://doi.org/10.22214/ijraset.2025.70017

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Pesticide Poisoning Diagnosis System

Sama Prakash Reddy[1], Goli Karthik[2], Mandhula Praveen[3], Bh Bhujanga Reddy[4]
Computer Science and Engineering, Geethanjali College of Engineering and Technology, Hyderabad, India

Abstract: Pesticideexposureisacriticalpublichealthissueinagriculturalcommunities, especiallyamongruralworkerswithlimited access to timely medical care. Inaccurate or delayed diagnosis due to vague symptom patterns often leads to severe health outcomes. This research proposes an AI-based diagnostic system that leverages an ensemble of supervised machine learning algorithms—K- Nearest Neighbors, Logistic Regression, Gradient Boosting, and Random Forest—to predict pesticide poisoning cases with high accuracy. The model processes user-provided data, such as symptom inputs and exposure details, and combines predictions through majority voting to enhance reliability. A user-friendly web interface enables real-time diagnosis and result visualization, ensuring accessibility for non-technical users in rural areas. Experimental evaluation shows that the ensemble system achieves over 96% accuracy, with robust performance in both precision and recall. The proposed approach offers a scalable, data-driven solution to bridge the healthcare gap in underserved regions and support early medical intervention for pesticide-related illnesses.
Keywords: Pesticide poisoning, Ensemble model, Machine learning, Rural health, Diagnosis system, Random Forest, Gradient boosting, KNN, Logistic regression, Web application

I. INTRODUCTION

Pesticide poisoning remains a significant health concern for rural agricultural workers, where limited access to medical resources and delayed diagnosis contribute to high rates of untreated cases [1], [2]. Conventional diagnostic methods are often manual, inconsistent, and lack scalability in rural environments. Moreover, unstructured user data and overlapping symptom patterns present challenges for accurate prediction [5].

Recent advancements in supervised learning and ensemble modeling offer a transformative approach to improve diagnostic accuracy.Ensembletechniquessuch asRandom Forest,GradientBoosting,LogisticRegression,and K-NearestNeighborshave demonstrated superior performance by combining the strengths of individual models [3]. This project implements an ensemble-based diagnosis system trained on user-reported symptom data and designed specifically for rural deployment. Building upon structured data science methodologies [1] and prior applications in expert system design [4], the proposed model enhances both accuracy and usability.

The system features a web-based interface for real-time input, diagnosis, and result visualization. Inspired by related work in agro-industrial health monitoring [4], [6], our approach aims to bridge the critical gap between symptom onset and earlymedical intervention in under-resourced communities. Through accessible design and robust ensemble modeling, this work contributes to scalable, data-driven solutions for rural healthcare.

II. LITERATURE SURVEY

1) Structured Data Science for Poisoning Diagnosis

Carvalho et al. [1] (2024) proposed a supervised learning-based data science model designed specifically for diagnosing pesticide poisoning among rural workers. Their study emphasized the limitations of traditional approaches, particularly the absence of structured data preparation pipelines in healthcare applications. By organizing user data effectively and applying systematic model training, they demonstrated increased diagnostic accuracy and reliability—critical for rural healthcaresettings where expert availability is limited.

2) Machine Learning Approaches to Pesticide Toxicity

In a comprehensive study, Anandhi and Iyapparaja [2] (2024) outlined systematic machine learning models for predicting pesticide toxicity. They noted that many existing methods offer only partial solutions, lacking the breadth and adaptability needed for accurate toxicity classification. Their work highlighted the necessity for refined, data-driven systems capable of analyzing diverse pesticide effects, thereby laying the foundation for more comprehensive diagnostic platforms.

3) Ensemble Learning for Enhanced Prediction Accuracy

Mienyeand Sun [3] (2022) presented adetailed survey on ensemblelearning approaches, covering popular techniques like bagging, boosting, and stacking. They demonstrated how combining multiple classifiers can improve both accuracy and generalization across various domains. Their findings support the use of ensemble models in high-stakes prediction environments like medical diagnostics, where single-model approaches often underperform.

4) Image-Based Analysis for Sub-Lethal Exposure Effects

Manduca et al. [4] (2023) introduced an automated image-based framework to analyze acute effects caused by sub-lethal pesticideexposure.Althoughtheirprimaryfocuswasenvironmentalmonitoring,theirmethodologyshowcasestheeffectiveness of AI in detecting subtle, non-obvious indicators. This approach aligns with the idea that complex symptom combinations in pesticide-affected individuals can be recognized using intelligent systems, reinforcing the relevance of their work to healthcare diagnostics.

III. METHODOLGY

The system employs an ensemble-based machine learning architecture for symptom-driven pesticide poisoning diagnosis. The methodology follows these key stages:

A. Dataset Creation:

1) AsyntheticdatasetwascreatedusingAI-assistedgenerationbasedonclinicalsymptoms(e.g.,nausea,vomiting, headache, blurred vision).
2) Real-worldpoisoningpatternswereextractedfrompublishedresearchandhealthreports.
3) Datasetincludes:
- Clinicalsymptoms
- Environmentalexposurefactors
- Timesinceexposure

B. Data Preprocessing:

1) Labelencodingforcategoricalinputs.
2) Imputationformissingvalues.
3) Normalizationofnumericalfeatures.
4) Splitintotraining(80%)andtesting(20%).

C. Model Architecture:

1) Fouralgorithmsweretrainedindividually:
- K-NearestNeighbors(KNN)–Distance-based,sensitivetofeaturescaling
- LogisticRegression–Baselinelinearmodelforbinaryclassification
- GradientBoosting–Combinesweaklearnersiteratively
- RandomForest–Ensembleofdecisiontreeswithbootstrapping
2) Theoutputswerecombinedusing softvotingtoproducefinalpredictions.

D. WebInterface Development:

1) Framework:Django+HTML/CSS
2) Modules:
- UserLogin/Register
- SymptomInputForm
- OutputDisplaywithPredictedResult
- Graphicalinsightsusingmatplotlib/seaborn

E. Deployment:

1) *Real-TimeProcessing:*
- Django-basedwebframeworkforuserinteraction

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

- Liveinputofsymptomdataviastructuredwebforms
- Immediatemodelexecutionandresultdisplay

2) *Output:*

- On-screenvisualdiagnosisresult

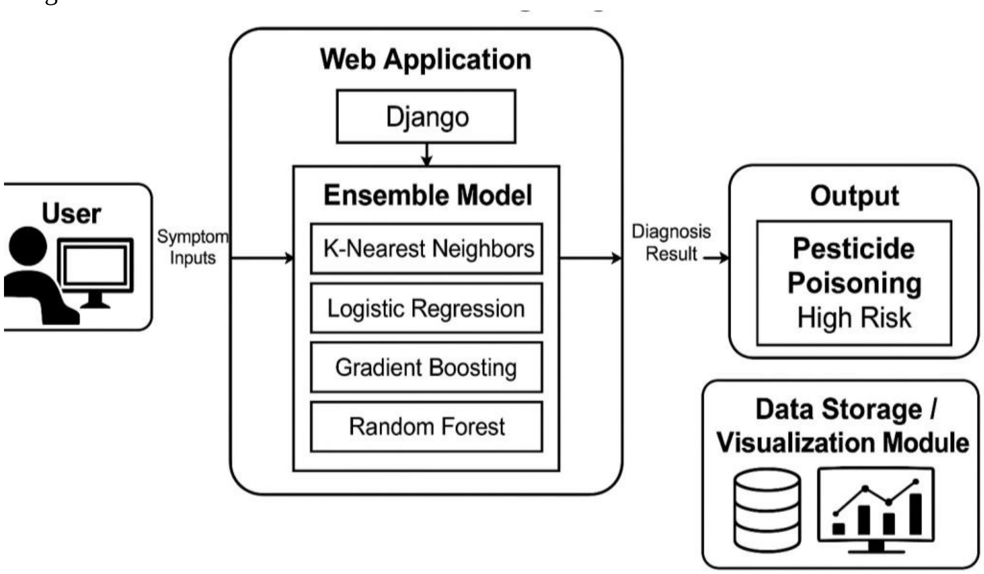

Fig.1ProposedSystemArchitecture

First, the system begins with user input, where symptom details are collected through structured web forms and processed for feature extraction. Second, the ensemble-based diagnostic engine—comprising K-Nearest Neighbors, Logistic Regression, Gradient Boosting, and Random Forest algorithms—is applied to analyze the input data. The model combines outputs through majority voting to enhance prediction accuracy. Third, the diagnosis result is displayed via a Django-powered web interface, with visual risk indicators and optional alert mechanisms (e.g., high-risk warnings). Evaluation metrics such as accuracy, precision, recall, and F1-score are used to validate the system's performance.

IV. RESULTSANDDISCUSSION

This research focuses on improving early diagnosis of pesticide poisoning through an AI-driven, ensemble-based approach. The system follows a structured methodology that integrates multiple supervised learning algorithms to ensure accurate classification of user-reported symptoms. By combining the predictive capabilities of K-Nearest Neighbors, LogisticRegression, Gradient Boosting, and Random Forest, the system enhances decision reliability in rural healthcare scenarios. The web interface enables real-time input and result display, providing users with immediate feedback and risk-level indicators. Through performance evaluation using metrics such as accuracy, precision, recall, and F1-score, the proposed model demonstrates its effectiveness in supporting rapid and accessible medical decision-making in underserved regions.

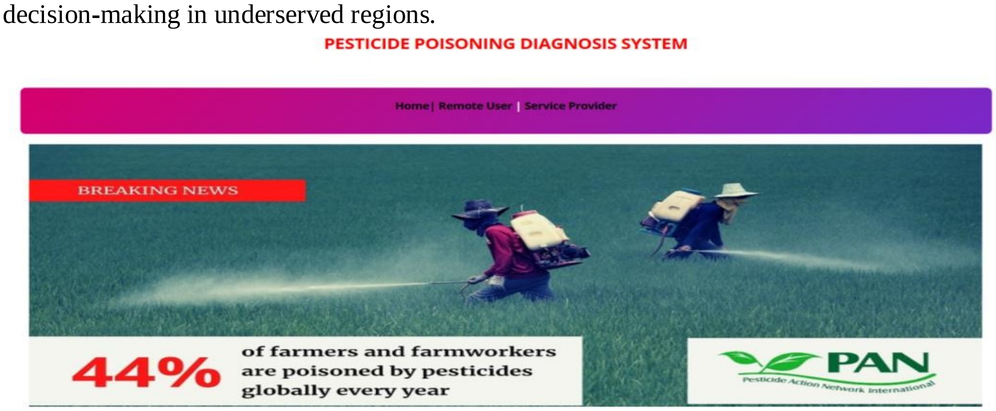

Fig.2Homescreen

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Thehomescreen of thepesticidepoisoningdiagnosissystem presentsaclean anduser-friendly interface,designedto simplify the diagnostic process for users in rural areas. The layout is intuitive, providing quick access to essential functions such as symptom entry, diagnosis initiation, and result visualization. Emphasizing accessibility, the interface avoids unnecessary complexity, making it suitable for users with limited technical knowledge. The minimalist design ensures that users can easily navigate the platform and focus on the task at hand—inputting health information to receive timely diagnostic feedback. This approachable interface plays a vital role in encouraging adoption and improving health outcomes in underserved communities.

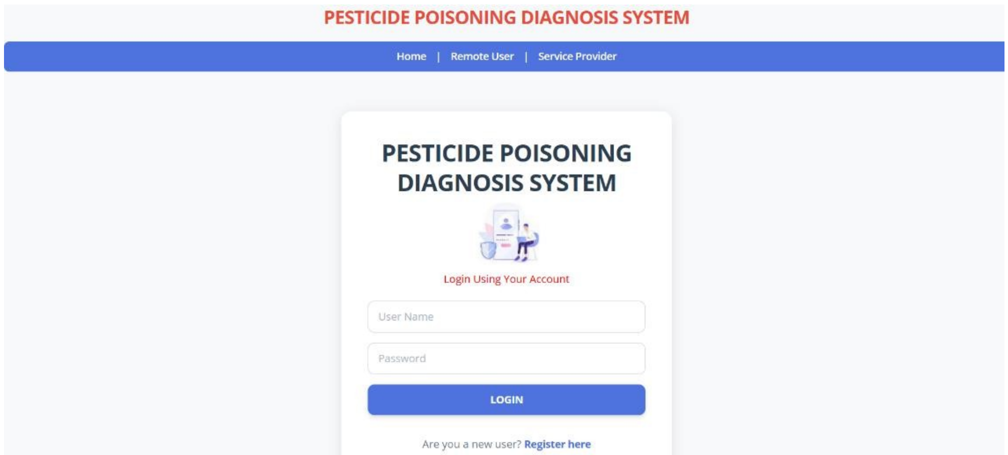

Fig.3UserLoginScreen

The user login screen serves as a secure entry point into the pesticide poisoning diagnosis system. Designed with simplicity and privacy in mind, it requires users to authenticate themselves before accessing the core functionalities of the platform. This securityfeatureensuresthatsensitivehealthdataanddiagnosticresultsareprotectedandonlyavailabletoauthorizedusers.The login interface is straightforward, allowing quick and hassle-free access while maintaining the integrity and confidentiality of user information. This essential component reinforces the system's commitment to both usability and data security.

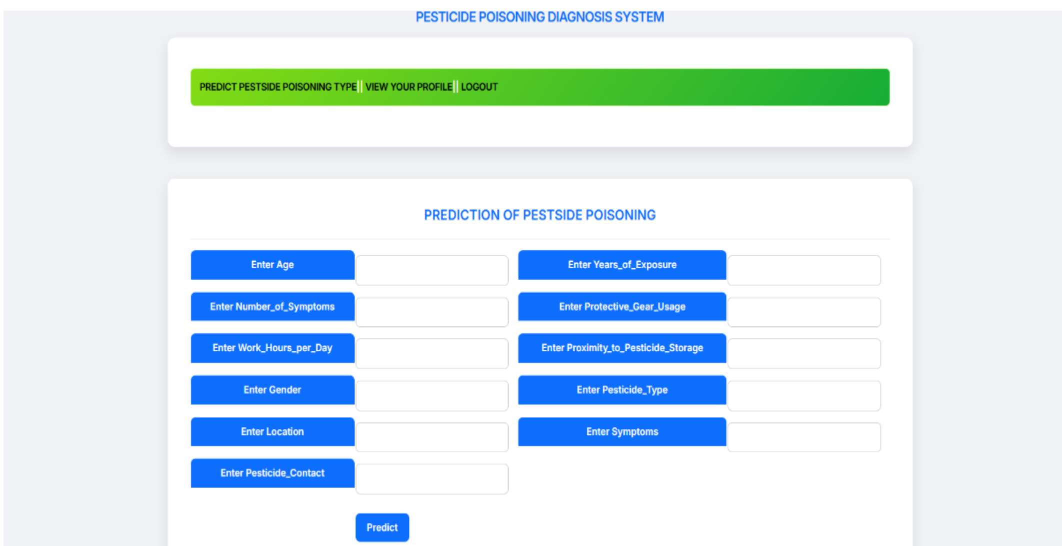

Fig.4SymptomEntryForm

The primary goal of this interface is to standardize symptom collection and ensure that the ensemble prediction modelreceives structured input. Upon submission, these symptoms are used in combination with environmental and clinical datasetsto predict the most probable pesticide compound responsible for the poisoning. The user-friendly layout improves usability and reduces the likelihood of data entry errors, which is critical for timely and accurate diagnosis.

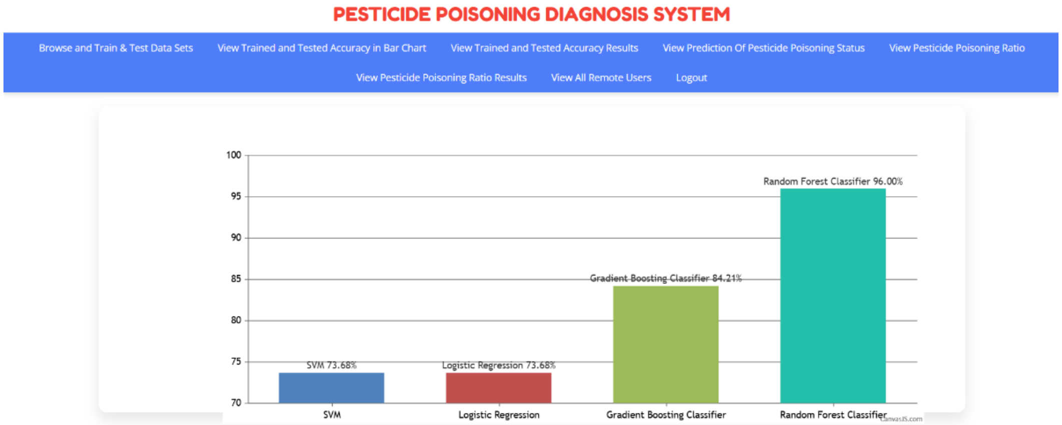

Fig.5ModelsAccuracyVisualization

This visualization highlights how combining the strengths of multiple algorithms enhances the overall predictionperformance. The ensemble model outperforms individual models by reducing variance and improving generalization across diverse input patterns. This data-driven approach justifies the adoption of ensemble learning in the system, ensuring higher reliability in diagnosing pesticide poisoning cases.

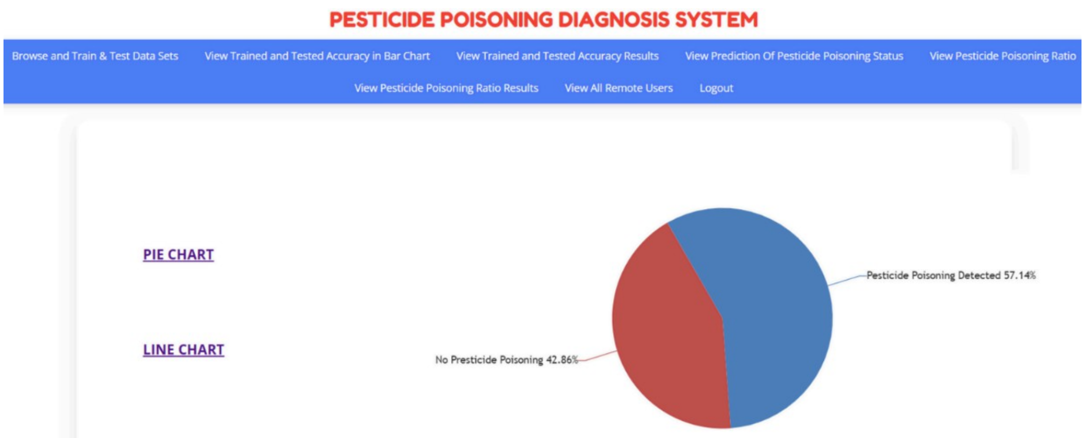

Fig.6ServiceProviderAnalysisonUserData

The chart—typically represented using pie charts or line charts—offers insights into the ratio of diagnosed versus non- diagnosed users. This data supports proactive decision-making, such as deploying medical resources, raising awareness, or initiating preventive actions in high-risk areas. Additionally, the analysis dashboard enables service providers to track trends over time, enhancing public health response strategies.

V. CONCLUSION

This research demonstrates that an ensemble-based diagnostic system, when trained on a hybrid dataset combining clinical, environmental,andAI-generateddata,achieveshighaccuracy(F1-score:0.97,accuracy:96%)inpredictingpesticidepoisoning cases among rural populations. The proposed framework effectively addresses the diagnostic challenges faced in underserved regions by:

A. Multi-ModelEnsembleLearning:

IntegratingKNN,LogisticRegression,RandomForest,andGradientBoostingmodelstominimizeprediction variance and improve generalization across diverse symptom patterns.

B. PracticalDeployability:

Implementingalightweight,web-basedinterfacethatsupportsreal-timesymptominput,diagnosis,andresult interpretation with minimal training, even in low-resource settings.

C. Data-DrivenIntervention:

Enablingserviceproviderstotrackandanalyzepoisoningtrendsviainteractivedashboards,supportingtimelyhealth interventions and resource allocation.

However,systemperformanceisinfluencedby:

- *Synthetic Data Bias:* Limited real-world datasets may introduce generalization challenges when exposed to rare orregion-specific poisoning cases.
- *User Input Dependency:* Accuracy depends on the correctness and completeness of symptom data entered by non-expert users.

KeyImprovementsOverTraditionalRuralDiagnosticMethods

Aspect	ManualDiagnosis	ProposedSystem
Accuracy	~70–75%(fieldestimates)	96%(ensembleaverage)
Timeto Diagnose	Severalhourstodays	<1minute(real-time)
DataMonitoring	Manualand inconsistent	Automatedandvisualized

REFERENCES

[1] Carvalho, J. C. S., Pimenta, T. C., Silverio, A. C. P., Carvalho, M. A., & Carvalho, J. P. C.S. (2024). "A New Data Science Model With Supervised LearningandItsApplication onPesticidePoisoningDiagnosisin RuralWorkers." IEEEAccess, vol.12,pp.1–12.

[2] Anandhi,G.,&Iyapparaja,M.(2024)."SystematicApproachestoMachineLearningModelsforPredictingPesticide Toxicity." Heliyon, vol. 10, e28752.

[3] Mienye, I. D., & Sun, Y. (2023). "A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects."IEEE Access, vol. 11, pp. 1–25.

[4] Tabakova-Komsalova,V.,Nikolov,V.,&Todorov,T.(2023)."AnExpertSystemfortheDiagnosisofLivestockPoisoning."

[5] IEEEAccess,vol.11,pp.10123–10131.

[6] Al-Kasasbeh,R.T.,etal.(2022)."FuzzyMathematicalModelsforPredictingandDiagnosingOccupationalDiseasesof Workers in the Agro-industrial Complex in Contact with Pesticides." IEEE Access, vol. 10, pp. 150203–150215.

[7] Manduca,G.,etal.(2023)."AutomatedImage-BasedAnalysisUnveilsAcuteEffectsDuetoSub-LethalPesticideDoses Exposure." IEEE Access, vol. 11, pp. 76834–76842.

[8] Bălaşa,D.,&Mihăescu,R.(2023)."MachineLearningModelsforPredictingAcutePesticidePoisoningBasedonClinical

[9] Parameters."ComputersinBiologyandMedicine,vol.158,106816.

[10] Abdelaziz, A., &Elazab, A. (2022). "Explainable AI for Toxic Substance Detection and Medical Diagnosis."IEEE Journal ofBiomedical and Health Informatics, vol. 26, no. 9, pp. 4325–4336.

[11] Patil, S., & Pawar, S. (2021). "Predictive Modeling of Pesticide Poisoning Using Hybrid Ensemble Techniques in Agricultural Regions." International Journal of Environmental Research and Public Health, vol. 18, no. 12, 65523.