



**International Journal of  
Engineering Research and Science & Technology**



**ISSN : 2319-5991**

[www.ijerst.com](http://www.ijerst.com)

**Email: [editor@ijerst.com](mailto:editor@ijerst.com) or [editor.ijerst@gmail.com](mailto:editor.ijerst@gmail.com)**

# IoT-Enabled Predictive Modelling Of Atmospheric And Soil Conditions Using AgroTS Dataset

G. Vinod Kumar, Ch. Sai Shiva Reddy<sup>2</sup>, T. Sai Prasad<sup>3</sup>, Dr. S. Ravi Kumar<sup>4</sup>

<sup>1,2,3</sup> UG Scholar, Dept. of ECE, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100

<sup>4</sup> Associate Professor, Dept. of ECE, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100  
[gugulothuvinodnayak@gmail.com](mailto:gugulothuvinodnayak@gmail.com)

## Abstract:

The evolution of agricultural practices has seen a significant transformation from traditional methods to data-driven approaches. Historically, agriculture relied heavily on manual observations and experience-based techniques to predict atmospheric and soil conditions. The AgroTS dataset emerged as a critical resource, containing vital information on various atmospheric and soil parameters, paving the way for more precise agricultural practices and resource management. The primary objective of this study is to develop a predictive modeling system that leverages IoT technology to analyse atmospheric and soil conditions, enhancing decision-making in agriculture. By utilizing the AgroTS dataset, the study aims to improve the accuracy of predictions related to crop yield, soil health, and climate impact, ultimately contributing to sustainable agricultural practices. Traditional agricultural systems primarily relied on manual labor and experience, involving physical observation of weather patterns and soil conditions. Farmers depended on almanacs and historical data for decision-making, often leading to inefficiencies and unpredictability in crop yields. Before adopting machine learning approaches, traditional methods of assessing atmospheric and soil conditions were often inaccurate and inefficient. Farmers faced challenges in predicting weather patterns and soil quality, resulting in suboptimal crop management decisions and reduced yields. The lack of real-time data and advanced analytics limited farmers' ability to respond promptly to environmental changes, leading to resource wastage and lower productivity. The proposed system will consist of an integrated IoT framework that includes a network of sensors deployed in agricultural fields to monitor atmospheric parameters such as temperature, humidity, and rainfall, as well as soil conditions like moisture and pH levels. This data will be collected in real-time and transmitted to a central database for analysis. A user-friendly interface will allow farmers to visualize data trends and receive predictive insights to guide their farming practices, ultimately optimizing crop management and resource allocation.

**Keywords:** *AgroTS, Sustainable agricultural practices, pH level, Almanacs.*

## 1. INTRODUCTION

Agriculture forms the backbone of India's economy, contributing approximately 18% to the country's GDP and employing over 50% of its workforce. However, the sector faces challenges like unpredictable weather patterns, declining soil fertility, and resource mismanagement, which hinder productivity. Traditional farming methods, heavily reliant on manual observation and ancestral knowledge, often lack precision and fail to address the dynamic nature of climate change. The advent of the AgroTS dataset and IoT technologies has opened new avenues for data-driven decision-

making in agriculture. IoT-enabled predictive modeling facilitates real-time monitoring of atmospheric and soil conditions, offering actionable insights for farmers.

Applications include optimizing irrigation schedules, detecting soilnutrient deficiencies, and mitigating climate risks. By integrating advanced analytics, this approach empowers sustainable agricultural practices, enhances crop yields, and reduces resource wastage. In India, where over 60% of farmland depends on rainfall, leveraging such technologies can transform farming into a resilient and efficient practice.

With the rise of IoT technologies and datasets like AgroTS, there is immense potential to revolutionize farming practices by integrating real-time data collection and predictive analytics. This research aims to bridge the gap between technology and agriculture, empowering farmers with actionable insights to improve decision-making, reduce resource wastage, and enhance sustainability. The prospect of leveraging technology to ensure food security and improve rural livelihoods provides a strong impetus for this research.

The increasing unpredictability of climate and declining soil health necessitate a shift from traditional to data-driven agricultural practices. This research addresses the need for real-time monitoring and predictive analysis of atmospheric and soil conditions, which are critical for optimizing crop management. By leveraging IoT technology and datasets like AgroTS, farmers can gain insights into temperature, humidity, soil moisture, and pH levels, enabling precise interventions to improve crop yield and resource efficiency. India, with its agrarian economy and a vast population to feed, requires innovative solutions to mitigate the adverse effects of climate change and ensure sustainable agricultural practices. This research has the potential to revolutionize the sector by reducing dependency on manual observations, promoting efficient resource allocation, and boosting productivity.

## 2. LITERATURE SURVEY

Acharya et al. [1] concluded that in the Indian economy around 60% of India's population is engaged in agriculture, and it contributes about 17–18% to the country's Gross Domestic Product (GDP). With the increasing demand for organic and eco-friendly products in the global market, the Indian agriculture sector is expected to grow even further. Kuradusenge et al. [2] they developed a machine learning model to predict a crop yield and improved productivity in agriculture to achieved through the adoption of modern technologies like precision farming techniques and the use of fertilizers and pesticides. Sanjeevi et al. [3] they developed a Precision agriculture and farming using Internet of Things (IOT) based on wireless sensor network. The sensors and other advanced technologies are used to collect and analyze the data about soil conditions, crop growth, and other factors that affect agricultural productivity. In Research of

Elavarasan et al. [4] Kerala state is renowned for its diverse cultivation practices, ranging from traditional crops to modern horticulture and floriculture. Various crops cultivated in Kerala include coconut, rubber, paddy, cardamom, tea, ginger, tapioca, coffee, cashew, etc. Rajakumaran et al. [5] they developed a Crop yield prediction using multi-attribute weighted tree-based support vector machine, this process of crop prediction typically involves collecting and analyzing the data on various factors. For crop prediction, they use Internet of Things (IoT), machine learning and support vector machine. Ye, X.P. [6] developed a IoT devices to collect data on factors such as soil moisture, temperature, and fertility, which can be analyzed to predict the suitable crop for the soil this process typically involves deploying a network of sensors in the soil to gather data on soil conditions and soil moisture sensors used to determine the optimal amount of water required by the crops, while temperature sensors can help predict the growth rate of the plants. Kalimuthu et al. [7] they proposed a Crop prediction using machine learning by help of this proposed system, farmers could predict suitable crops for the land, resulting in better efficiency and improved crop productivity. The collected data are transmitted to a central database, where they are processed and analyzed using machine learning algorithms to identify patterns and make predictions. In the research of Cravero et al. [8] Agri-technology was recently emerged due to the advancements of machine learning and deep learning methods. This transformative technology empowers agricultural stakeholders with predictive insights and decision-making capabilities based on vast datasets, enhancing precision farming practices and sustainability efforts.

Pudumalar et al. [9] described a model that implemented precision agriculture (PA), using machine learning algorithms that work on small, open farms, which help farmers with the prediction of suitable crops. Rao et al. [10] conducted a comparative study on three supervised machine learning models—KNN, decision tree, and random forest—to determine the most suitable crop for specific land areas, aiming to enhance crop efficiency for farmers. Samuel et al. used specific data analytic approaches to estimate the crop [11] price analysis using already collected data. Patil et al. [12] developed mobile equipment to detect soil nutrients in real time. This system recorded the soil nutrients and characteristics, including humidity, temperature, and pH data, and these can be accessed globally with valid cloud channel verification credentials.

Morales et al. [13] found random forest to be highly effective in predicting crop yields using biophysical crop models, reinforcing the value of tree-based algorithms in agricultural applications. In the Research of Alexander et a. [14] they identify the significance of soil moisture in identifying L–A coupling strength within the convective triggering potential (CTP) and humidity index (HI) framework. To address this, a consistent and reliable dataset of atmospheric profiles is created by merging CTP and HI using triple collocation (TC) with three reanalysis datasets. In the Research of Arshad et al. [15] The CTP and HI are calculated using three reanalysis datasets (MERRA-2, CFSR, and ERA5), satellite estimates, and in situ observations. When performing TC analysis, it is crucial to consider the presence of biases and errors in the datasets across different variables and applications.

### 3. PROPOSED METHODOLOGY

#### Step 1: Loading the AgroTS Dataset

The AgroTS dataset, a structured repository containing data on atmospheric and soil conditions, is loaded into the system. This dataset includes critical features such as temperature, humidity, soil

moisture, and pollutant levels, as well as a target column indicating specific conditions (e.g., "Good" or "Hazardous"). The data serves as the foundation for both exploratory and predictive analysis.

#### Step 2: Dataset Preprocessing (Standard Scaler)

Preprocessing transforms raw data into a usable format for machine learning. The target column is separated from the feature set. Missing values are identified and addressed, and data types are verified. Standard scaling normalizes the feature set to ensure all variables are on a comparable scale, enhancing algorithmic performance. Splitting the data into training and testing subsets ensures that models can be evaluated effectively.

#### Step 3: Exploratory Data Analysis (EDA)

EDA provides insights into the dataset's structure and relationships. Key visualizations include correlation heatmaps to identify relationships between features and a count plot of the target column to assess data distribution across conditions. These steps help identify patterns and ensure the dataset is balanced, addressing issues like data imbalance if necessary.

#### Step 4: Implementation of the Existing KNN Classifier

The K-Nearest Neighbors (KNN) classifier is implemented as the baseline algorithm. KNN predicts the class of a data point based on its closest neighbors in feature space. The model is trained on the processed dataset, and predictions are generated for the test data. Performance metrics such as accuracy, precision, recall, and F1-score are calculated to benchmark the KNN's effectiveness.

#### Step 5: Development of the Proposed CNN Classifier

The Convolutional Neural Network (CNN) model introduces a more sophisticated approach. It processes input features as sequential data, applying convolutional layers to detect patterns. The architecture includes multiple convolutional, max-pooling, and dropout layers, culminating in fully connected layers with a softmax activation for multi-class classification. The model is trained using the scaled training data, and its learning history is saved for further analysis.

#### Step 6: Performance Comparison

A comparative analysis of the KNN and CNN classifiers evaluates their performance. Metrics such as accuracy, precision, recall, and F1-score for both algorithms are visualized using bar charts. This step highlights the strengths of the CNN model, such as higher accuracy and better handling of complex data patterns.

#### Step 7: Prediction Using the CNN Trained Model

The trained CNN model is used to predict the conditions of new test data. The input test data is preprocessed and fed into the CNN for classification. The predictions, categorized into conditions (e.g., "Good," "Hazardous"), are integrated back into the test data for interpretation. This output demonstrates the model's real-world applicability, showcasing its precision in predicting atmospheric and soil conditions. This system recorded the soil nutrients and characteristics, including humidity, temperature, and pH data, and these can be accessed globally with valid cloud channel verification credentials.

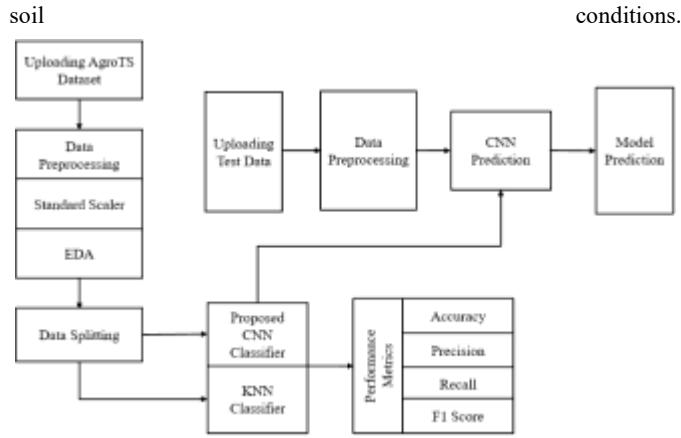


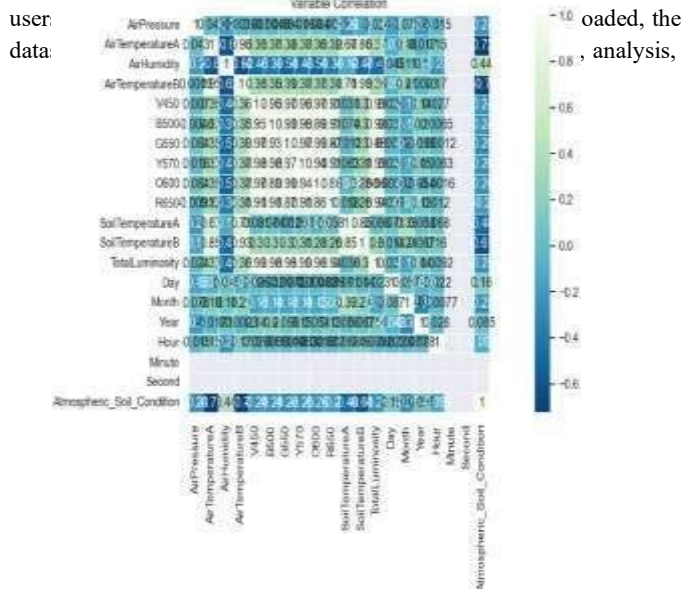
Figure 1: Architectural Block Diagram of the proposed system

Advantages of CNN:

- Automatic Feature Extraction: No need for manual feature engineering.
- High Accuracy: Excels in tasks like image recognition and classification.
- Parameter Sharing: Reduces the number of parameters compared to fully connected networks.
- Adaptability: Can be applied to various domains, including images, videos, and even time-series data.

4. EXPERIMENTAL ANALYSIS

This figure illustrates the process of uploading the AgroTs dataset through a graphical user interface. The interface is designed to allow user:



and model building.

Figure 2: Upload of AgroTs Dataset in the GUI Interface

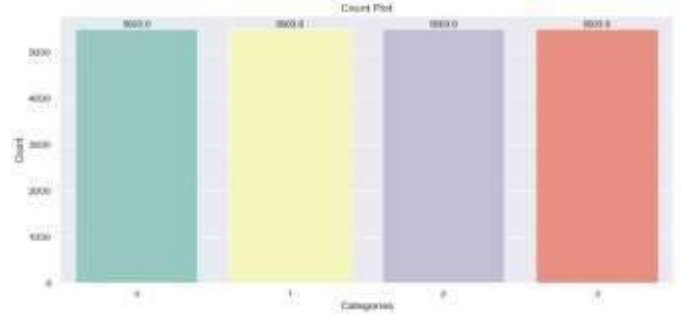
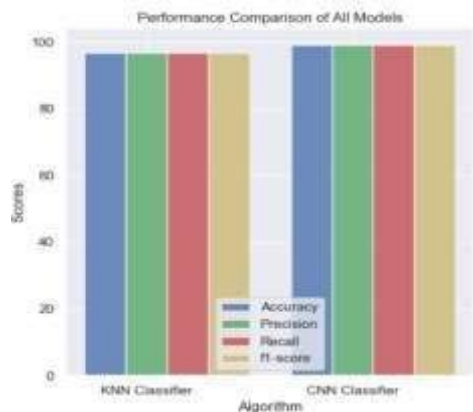


Figure 3: EDA Plots (Correlation Plot and Target Column Countplot).



Figure 4: Performance Metrics and Confusion Matrix Plot of CNN Classifier Model

This figure showcases exploratory data analysis (EDA) conducted on



the dataset. The correlation plot visualizes the relationships among numerical features, highlighting the strength and direction of linear relationships. The countplot of the target column provides a visual representation of the distribution of the atmospheric and soil conditions, assisting in understanding the class balance in the dataset.

KNN

This figure presents the performance metrics and confusion matrix of the KNN classifier. The accuracy, precision, recall, and F1-score indicate the classifier’s effectiveness in predicting the atmospheric and soil conditions. The confusion matrix plot provides a detailed breakdown of correct and incorrect predictions for each class,

demonstrating the model's high performance, with an accuracy of 96.64%.

**Figure 5: Performance Comparison Graph of All Models**

## 5. CONCLUSION

This research successfully demonstrates the implementation of machine learning techniques to classify atmospheric and soil conditions using sensor-based environmental data. By leveraging advanced algorithms such as CNN and KNN, the system efficiently processes and analyzes multiple features, including temperature, humidity, luminosity, and soil conditions. The results showcase the potential of machine learning in extracting meaningful insights and automating environmental monitoring tasks, which are essential for agriculture, weather prediction, and ecosystem management. The combination of data preprocessing, feature selection, and algorithmic modeling ensures high accuracy and robustness of the proposed system.

The system can be further enhanced through several key expansions. Integrating real-time data from sensors enables dynamic monitoring and instant classification. Additionally, incorporating extra features such as wind speed, soil moisture, and precipitation can improve classification accuracy. To increase scalability, the model can be extended to handle larger datasets from diverse geographical locations. Furthermore, deploying the system as an IoT solution involves developing a user-friendly interface and implementing it in IoT-based devices, making it practical for smart farming and environmental management. Exploring advanced algorithms like RNNs or transformers can also enhance performance in time-series and sequence data analysis. Predictive modeling can be added to forecast environmental changes based on historical trends, aiding proactive decision-making. Finally, optimizing the system for energy efficiency enables deployment in remote or resource-constrained areas.

## REFERENCES

1. Acharya, B.; Rashmi, M.; Ashwani, K.; Deepti, S.; Vedpriya, A. An Analysis of Policy Interventions in Agriculture and ICT Based Mechanistic Approach towards Sustainability: An Indian Perspective. *Asian J. Agric. Rural Dev.* **2020**, *10*, 194–213.
2. Kuradusenge, M.; Hitimana, E.; Hanyurwimfura, D.; Rukundo, P.; Mtonga, K.; Mukasine, A.; Uwitonze, C.; Ngabonziza, J.; Uwamahoro, A. Crop yield prediction using machine learning models: Case of Irish potato and maize. *Agriculture* **2023**, *13*, 225.
3. Sanjeevi, P.; Prasanna, S.; Siva Kumar, B.; Gunasekaran, G.; Alagiri, I.; Vijay Anand, R. Precision agriculture and farming using Internet of Things based on wireless sensor network. *Trans. Emerg. Telecommun. Technol.* **2020**, *31*, e3978.
4. Elavarasan, D.; Vincent, P.D. A reinforced random forest model for enhanced crop yield prediction by integrating agrarian parameters. *J. Ambient. Intell. Humaniz. Comput.* **2021**, *12*, 10009–10022.
5. Rajakumaran, M.; Arulselvan, G.; Subashree, S.; Sindhuja, R. Crop yield prediction using multi-attribute weighted tree-based support vector machine. *Meas. Sens.* **2024**, *31*, 101002.
6. Ye, X.P. Decarbonizing Nitrogen Fertilizer for Agriculture with Nonthermal Plasma Technology. *Eng* **2024**, *5*, 1823–1837.
7. Kalimuthu, M.; Vaishnavi, P.; Kishore, M. Crop prediction using machine learning. In Proceedings of the 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 20–22 August 2020; IEEE: New York, NY, USA, 2020.
8. Cravero, A.; Sepúlveda, S. Use and adaptations of machine learning in big data—Applications in real cases in agriculture. *Electronics* **2021**, *10*, 552.
9. Pudumalar et al. Using machine learning for implementation of precision agriculture (PA). *Front. Plant Sci.* **2023**.
10. Rao et al. Using supervised machine learning models, KNN, decision tree for suitable crop prediction of land areas. *Front. Plant Sci.* **2023**.
11. Samuel, P.; Sahithi, B.; Saheli, T.; Ramanika, D.; Kumar, N.A. Crop price prediction system using machine learning algorithms. *Quest J. J. Softw. Eng. Simul.* **2020**, *6*, 14–20.
12. Patil, V.K.; Jadhav, A.; Gavhane, S.; Kapare, V. Iot based real time soil nutrients detection. In Proceedings of the 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 5–7 March 2021; IEEE: New York, NY, USA, 2021.
13. Morales, A.; Villalobos, F.J. Using machine learning for crop yield prediction in the past or the future. *Front. Plant Sci.* **2023**, *14*, 1128388.
14. Alexander, G. A., Holmes, H. A., Sun, X., Caputi, D., Faloona, I. C., and Oldroyd, H. J.: Simulating land-atmosphere coupling in the Central Valley, California: Investigating soil moisture impacts on boundary layer properties, *Agr. Forest Meteorol.*, 317, 108898, <https://doi.org/10.1016/j.agrformet.2022.108898>, 2022.
15. Arshad, M., Ma, X., Yin, J., Ullah, W., Liu, M., and Ullah, I.: Performance evaluation of ERA-5, JRA-55, MERRA-2, and CFS-2 reanalysis datasets, over diverse climate regions of Pakistan, *Weather and Climate Extremes*, 33, 100373, <https://doi.org/10.1016/j.wace.2021.100373>, 2021.