



Ferti Forecast: Identification Of Fertilizers For Soil Based On Their NPK Levels Using Machine Learning Algorithms

¹B. Likitha, ²N. Mahendra Swamy, ³E. Charan Teja, ⁴N. Davidson, ⁵Mr. P. Darwin

¹Student(22551A0466), ²Student(22551A0499), ³Student(22551A0473), ⁴Student(22551A0497), ⁵Associate Professor

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING,
GODAVARI INSTITUTE OF ENGINEERING AND TECHNOLOGY(AUTONOMOUS),
RAJAHMAHENDRAVARAM

Abstract: Ferti Forecast is an IoT-enabled, machine learning-based system designed to provide intelligent fertilizer recommendations based on real-time soil nutrient analysis and environmental conditions. The system addresses the limitations of conventional fertilizer application methods, which often rely on fixed rules and lack adaptability to changing soil and climate conditions leading to nutrient imbalances, reduced crop yields, and environmental degradation. Ferti Forecast utilizes supervised machine learning algorithms, specifically Decision Tree and Random Forest classifiers, trained on real-world datasets containing Nitrogen (N), Phosphorus (P), Potassium (K) values, along with temperature and rainfall data. Among the tested models, Random Forest demonstrated superior accuracy and generalizability due to its ensemble learning capability and resistance to overfitting. The software implementation is carried out using Python, Scikit-learn, and Pandas within the Google Collab environment, allowing for scalable training and cloud-based prediction. On the hardware side, the system employs a Node MCU microcontroller connected to an NPK soil sensor, a temperature sensor, and a rainfall sensor for collecting live environmental and nutrient data. The Node MCU transmits this data via Wi-Fi to the trained model hosted online or locally. Based on the model's prediction, an LCD provides instant fertilizer recommendations to the user directly in the field.

Key words: Soil Nutrient Analysis, NPK Sensor, Fertilizer Recommendation System, Machine Learning, Random Forest Algorithm, Precision Agriculture, IoT-based Smart Farming, Real-Time Soil Monitoring, Sustainable Agriculture, Embedded Systems

I. INTRODUCTION

Agriculture is a major contributor to economic development, especially in developing countries. Proper fertilizer usage is essential for improving crop yield and maintaining soil fertility, as plants require key nutrients such as Nitrogen (N), Phosphorus (P), and Potassium (K) for healthy growth. However, inaccurate fertilizer application caused by limited soil nutrient knowledge often results in poor productivity, increased costs, and long-term soil degradation. This highlights the need for efficient and reliable fertilizer management practices.

Conventional fertilizer recommendation methods are largely based on manual soil analysis, fixed guidelines, or expert opinion, which may not adapt well to changing soil and environmental conditions. These approaches are time-consuming and can lead to overuse or underuse of fertilizers. With the growing availability of agricultural data, machine learning techniques have emerged as effective tools for analysing soil parameters and providing accurate, data-driven recommendations.

Recent advancements in IoT and sensor technology have enabled real-time monitoring of soil nutrients. Integrating NPK sensors with machine learning algorithms such as Decision Tree and Random Forest allows the development of intelligent fertilizer recommendation systems. Such systems assist farmers in making timely and precise decisions, improving crop productivity while supporting sustainable and eco-friendly agricultural practices.

II. LITERATURE SURVEY

Kumar R. & Mehta S. (2021) proposed a smart fertilizer recommendation system using Decision Tree classification based on soil NPK values and pH levels. The system analyzed soil test data to suggest suitable fertilizers for different crops. Their approach demonstrated improved recommendation accuracy compared to manual methods. However, the model relied on laboratory soil reports and did not support real-time sensor-based soil data collection.

Patel A. et al. (2022) developed an IoT-based soil monitoring system integrated with machine learning algorithms for nutrient analysis. Sensors were used to collect soil moisture and NPK values, which were processed using supervised learning models. The system enabled continuous monitoring and automated decision support. Despite its effectiveness, the implementation required multiple sensors, increasing system cost and maintenance complexity.

Sharma N. & Verma P. (2020) introduced a fertilizer prediction framework using Naïve Bayes classification to recommend nutrients based on soil fertility levels. The study focused on reducing fertilizer overuse by categorizing soil into nutrient-deficient classes. While the model was computationally efficient, it showed lower accuracy when handling complex soil datasets with non-linear relationships.

Rao S. et al. (2023) proposed a precision agriculture system combining IoT sensors and Random Forest algorithm for soil nutrient analysis. The system utilized real-time NPK sensor data to improve fertilizer selection accuracy. Their results showed that Random Forest outperformed single classifiers in terms of robustness and prediction reliability. However, the system required stable connectivity for data transmission, limiting deployment in remote farming regions.

Chakrabarty A. et al. (2018) implemented a deep learning-based crop prediction model using multiple soil and fertilizer attributes. The proposed system achieved high accuracy using a J48 classifier and highlighted the potential of neural networks in agricultural analytics. Despite promising results, the system required high computational resources and did not emphasize low-cost hardware deployment.

Vrushali Bhuyar (2019) conducted a comparative study on soil fertility classification using Random Forest, Naïve Bayes, and J48 algorithms. The results showed that Random Forest provided superior accuracy and robustness for fertility index prediction. The study emphasized decision support for fertilizer selection but did not integrate sensor-based real-time data collection.

Chandana S. & Parthasarathy R. (2022) proposed a regression-based fertilizer optimization model using the Naïve Bayes algorithm. The system aimed to improve fertilizer efficiency and crop yield by analyzing soil nutrient patterns. Although the approach reduced fertilizer misuse, it depended on pre-existing datasets and lacked IoT-based soil sensing mechanisms.

Islam M. et al. (2018) presented a deep neural network-based system for crop selection and yield prediction in Bangladesh. The model considered extensive soil and fertilizer attributes to enhance prediction reliability. While the study demonstrated the effectiveness of deep learning in agriculture, real-time implementation and sensor-level integration were not addressed.

III. EXISTING METHODOLOGY

Existing fertilizer recommendation systems are primarily based on traditional agricultural practices and predefined nutrient guidelines. In most cases, fertilizer selection is carried out using manual soil testing reports or fixed threshold values for soil nutrients such as Nitrogen, Phosphorus, and Potassium. These methods depend heavily on expert knowledge and farmer experience, making them time-consuming and prone to human error. As a result, fertilizer application is often generalized and does not accurately address the specific nutrient requirements of different soil conditions.

Some modern systems have attempted to improve fertilizer management using basic computational or statistical approaches. These systems utilize historical soil data and rule-based logic to recommend fertilizers; however, they lack adaptability to dynamic soil and environmental variations. Since the input data is usually static and not collected in real time, the recommendations may become outdated or inaccurate. Additionally, such systems often require laboratory-based soil analysis, which increases cost and limits frequent monitoring.

Although recent studies have introduced machine learning techniques for agricultural decision support, many existing implementations are limited to software simulations or offline datasets. Hardware integration for real-time soil nutrient sensing is often absent, reducing their practical applicability in field conditions. Furthermore, several existing systems rely on single classification models, which may suffer from overfitting and reduced accuracy. These limitations highlight the need for an integrated, real-time, and intelligent fertilizer recommendation system that combines sensor-based data collection with robust machine learning algorithms.

IV. PROPOSED SYSTEM AND WORKING METHODOLOGY:

The proposed fertilizer recommendation system is designed to intelligently suggest suitable fertilizers by continuously analyzing real-time soil nutrient conditions and environmental parameters. The system uses a NedeMCU microcontroller as the central processing unit to collect soil data through an NPK sensor along with auxiliary sensors for temperature, humidity, and soil moisture. Instead of relying on static threshold-based rules, the system follows a machine learning-based decision-making approach, where historical soil nutrient data is learned to predict appropriate fertilizer recommendations. A trained Random Forest model evaluates Nitrogen, Phosphorus, Potassium levels, temperature, humidity, and moisture values to determine the most suitable fertilizer for the given soil condition. The predicted fertilizer recommendation is displayed instantly on an LCD module, enabling farmers to make timely and accurate decisions. This methodology ensures precision fertilization, reduced chemical overuse, improved soil health, and minimal manual intervention.

A) DATA ACQUISITION

Soil and environmental data collection forms the foundation of the proposed system. The NPK sensor is used to measure real-time Nitrogen, Phosphorus, and Potassium content present in the soil. In addition, temperature, humidity, and soil moisture sensors continuously monitor surrounding environmental conditions that influence fertilizer efficiency. All sensor readings are collected by the NodeMCU controller and transmitted to the processing unit either through serial communication or wireless connectivity. The data is updated periodically, ensuring accurate and real-time representation of soil conditions.

B) DATA PRE-PROCESSING AND FEATURE FORMATION

The raw sensor data obtained from the field may contain noise and varying scales. Therefore, preprocessing techniques such as normalization and data cleaning are applied before feeding the data to the machine learning model. Feature vectors are formed using soil nutrient parameters (N, P, K), temperature, humidity, and moisture values. This structured dataset improves prediction accuracy and ensures consistency across different soil samples. The processed data is then formatted into a tabular structure suitable for machine learning model inference.

C) MACHINE LEARNING TRAINING MODEL

The fertilizer recommendation model is developed using supervised machine learning techniques. Historical datasets containing soil NPK values and corresponding fertilizer labels are used to train multiple classifiers such as Decision Tree, Naïve Bayes, and Random Forest. Among these, the Random Forest algorithm demonstrated superior performance due to its ability to handle nonlinear relationships, reduce overfitting, and improve generalization accuracy. The trained model is saved using serialization techniques and deployed for real-time fertilizer prediction based on incoming sensor data.

D) DECISION EXECUTION AND OUTPUT DISPLAY

Once the processed soil data is passed to the trained machine learning model, the system predicts the most suitable fertilizer type. The NodeMCU receives the prediction output and displays the recommended fertilizer name on an LCD display for user reference. This real-time output allows farmers to immediately apply the recommended fertilizer. The system ensures reliability by continuously updating predictions as soil conditions change, supporting accurate and sustainable fertilizer management.

E) BLOCK DIAGRAM

The block diagram of the proposed system represents the complete data flow from soil sensing to fertilizer recommendation. The NodeMCU microcontroller acts as the central control unit, receiving real-time input from the NPK sensor and environmental sensors. Soil nutrient values and environmental parameters are processed and forwarded to the embedded machine learning model. The trained Random Forest model analyzes the inputs and generates a fertilizer recommendation. The final output is displayed on an LCD module for user interaction. A stable power supply ensures uninterrupted operation of the system. The block diagram clearly illustrates the integration of hardware sensing, machine learning-based prediction, and real-time output generation.

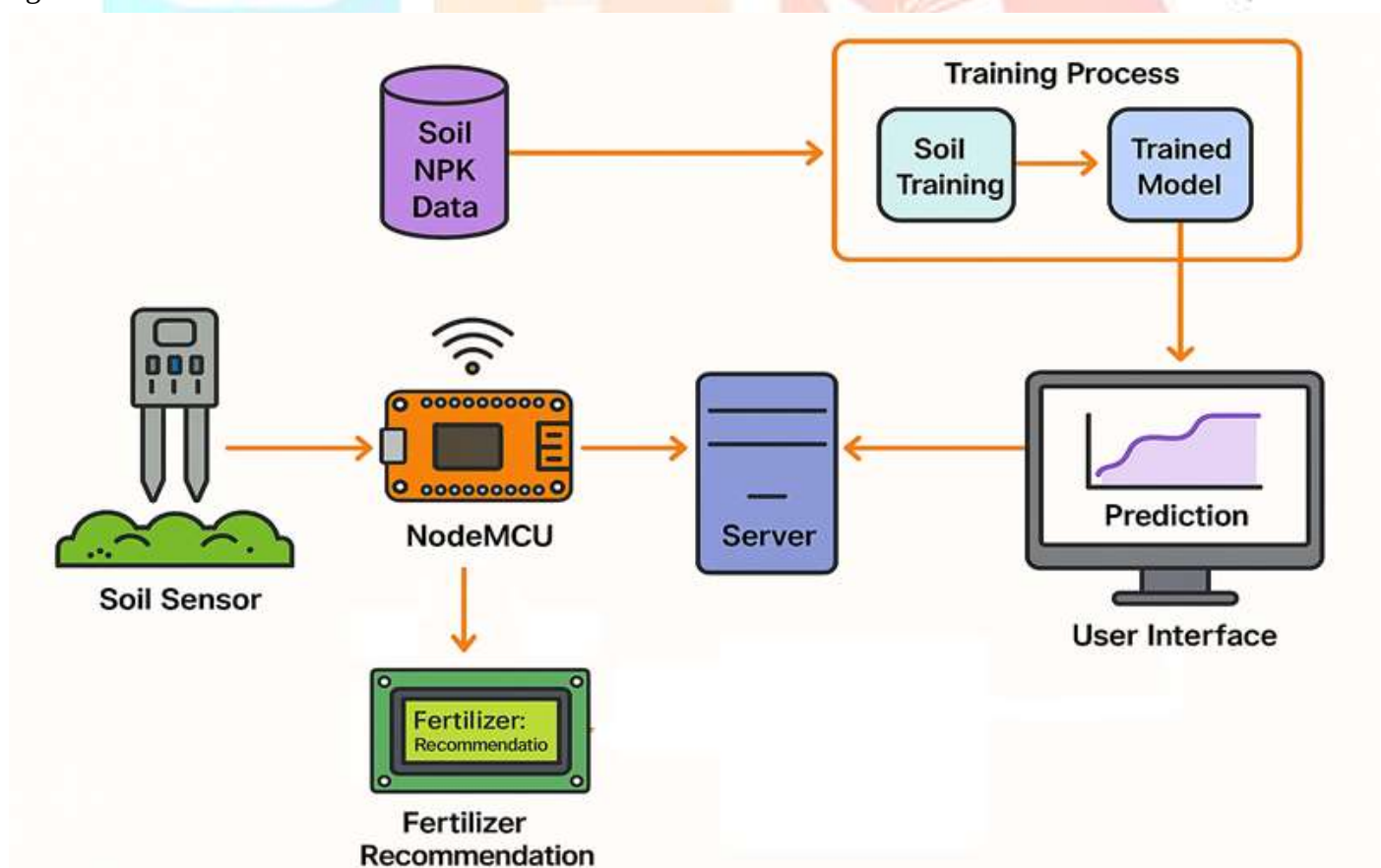


Fig. 1. Block Diagram of Proposed ferti forecast

F) WORKFLOW DIAGRAM

The workflow diagram illustrates the sequential execution of the proposed fertilizer recommendation system, beginning with system initialization followed by continuous monitoring of soil nutrient and environmental conditions. The process starts once the NodeMCU controller is powered ON. All connected sensors are initialized and begin capturing real-time values of Nitrogen, Phosphorus, Potassium, temperature, humidity, and soil moisture. These parameters are collected continuously to reflect the current soil condition accurately. The acquired sensor data is then forwarded to the machine learning inference module, where the trained Random Forest model evaluates the input parameters in real time. Based on the analyzed NPK levels and environmental conditions, the model predicts the most suitable fertilizer required for the soil. If the soil nutrient levels indicate deficiency, the system generates a specific fertilizer recommendation to restore nutrient balance. In cases where the nutrient levels are already optimal, the system avoids unnecessary fertilizer application. After prediction, the recommended fertilizer information is displayed on the LCD module, providing immediate feedback to the user. The display is updated dynamically after each execution cycle, ensuring that farmers receive real-time guidance based on changing soil conditions. This workflow repeats continuously throughout system operation, enabling accurate, adaptive, and sustainable fertilizer management with minimal manual intervention.

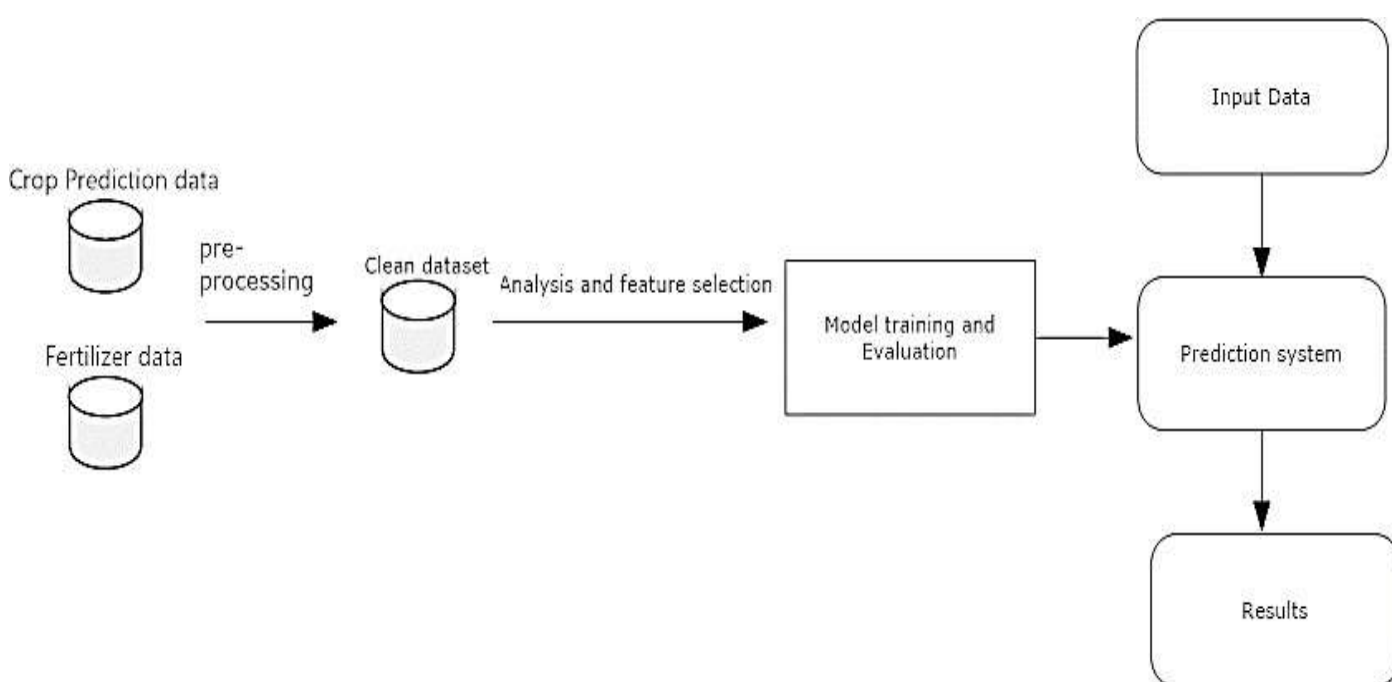


Fig. 2. Workflow Diagram of ferti forecast

V. RESULT

The developed fertilizer recommendation prototype was tested under various real-time soil and environmental conditions to evaluate its prediction accuracy and system reliability. The NodeMCU controller successfully acquired soil nutrient values from the NPK sensor along with temperature, humidity, and moisture readings. The collected data was processed by the trained machine learning model, which generated accurate fertilizer recommendations in real time. The system responded dynamically to changing soil conditions, ensuring precise and efficient fertilizer guidance with minimal manual intervention.

A) HARDWARE RESULT – PROTOTYPE IMPLEMENTATION

The complete working setup of the proposed system is shown below. The NodeMCU microcontroller is interfaced with an NPK sensor, temperature–humidity sensor, soil moisture sensor, and an LCD display module. Proper wiring, stable sensor placement, and consistent power supply ensured reliable data acquisition and uninterrupted system operation. The successful integration of sensors and controller validates the feasibility of deploying the system in real agricultural environments.

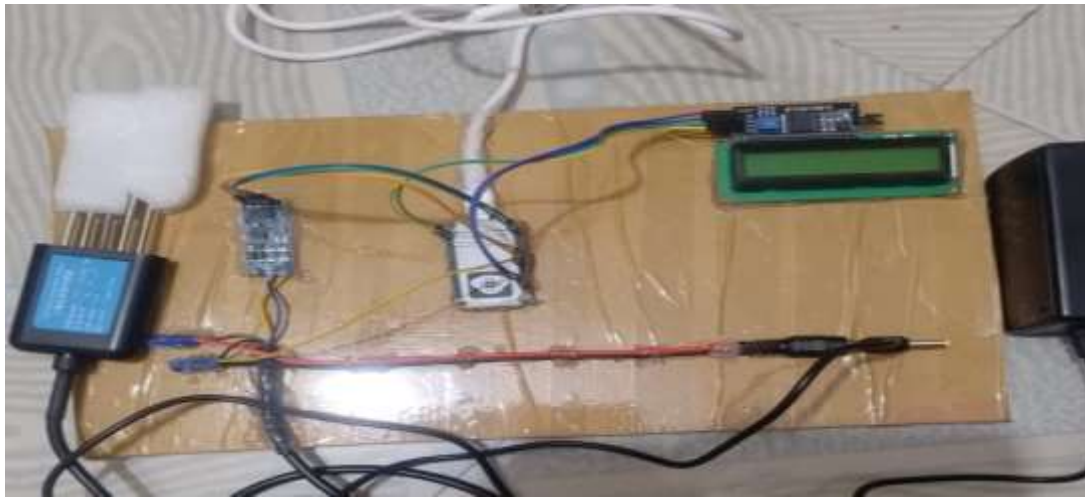


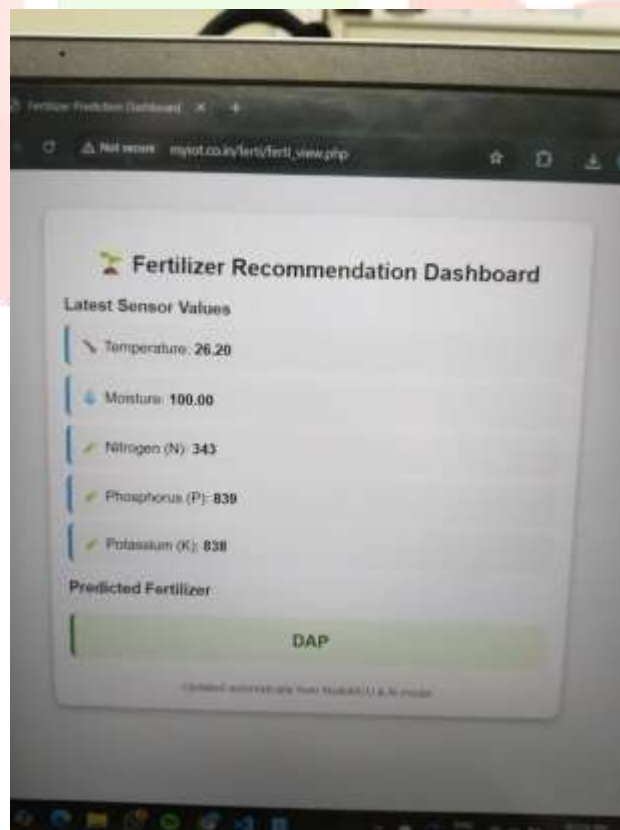
Fig. 3. Prototype setup with NodeMCU, NPK sensor, environmental sensors, and LCD display.

B) PRACTICAL WORKING OUTPUT – REAL-TIME FERTILIZER RECOMMENDATION

The system was tested using multiple soil samples with varying nutrient compositions. When the soil showed Nitrogen deficiency, the system correctly recommended Urea, while balanced NPK conditions resulted in the suggestion of compound fertilizers such as 17-17-17. The fertilizer recommendation was updated automatically whenever soil conditions changed, demonstrating the system's real-time responsiveness and prediction accuracy.

Fig. 4. Live fertilizer recommendation output displayed on LCD for different soil conditions.

C) REAL-TIME SENSOR DATA DISPLAY



The LCD interface displayed live values of Nitrogen, Phosphorus, Potassium, temperature, humidity, and soil moisture, along with the corresponding fertilizer recommendation. This real-time visualization provided transparency to the user and enabled immediate understanding of soil health. Continuous updates ensured that farmers could monitor soil conditions without relying on laboratory testing.



Fig. 5. LCD output showing live NPK values and recommended fertilizer.

D) MODEL PREDICTION LOG OUTPUT

The Python console displayed real-time prediction logs generated by the machine learning model. For each input dataset received from the sensors, the model output a fertilizer label, which was then converted into a readable fertilizer name. The prediction logs confirmed consistent model behaviour and accurate decision-making across different test scenarios.

```
import pandas as pd

sample = pd.DataFrame([[25, 80, 40, 50, 20, 30]],
                      columns=['Temperature', 'Humidity', 'Moisture',
                              'Nitrogen', 'Potassium', 'Phosphorous'])

pred = loaded_model.predict(sample)
print("Predicted Label:", pred)

# If you want fertilizer name:
fert_name = label.inverse_transform(pred)[0]
print("Recommended Fertilizer:", fert_name)
```

[14]

```
... Predicted Label: [6]
Recommended Fertilizer: DAP
```

Fig. 6. Console-based output displaying predicted fertilizer labels and names.

SUMMARY OF RESULTS

- Accurate real-time fertilizer recommendation based on soil NPK values
- Successful integration of hardware sensors with machine learning model
- Reduced dependency on manual soil testing and expert consultation

- Continuous real-time monitoring of soil and environmental parameters
- System supports sustainable and precision fertilizer management

VII. CONCLUSION

The smart fertilizer recommendation system developed in this project demonstrates the effective integration of soil nutrient sensing, environmental monitoring, and machine learning-based decision making to support precision agriculture. Using the NodeMCU microcontroller as the central control unit, the system continuously collects real-time data from the NPK sensor along with temperature, humidity, and soil moisture sensors. This data is processed through trained machine learning models to generate accurate and context-aware fertilizer recommendations, which are displayed instantly to the user.

By continuously analyzing soil nutrient levels and environmental conditions, the system ensures that fertilizers are recommended only when required and in appropriate proportions. This approach helps in improving crop productivity while minimizing excessive fertilizer usage, thereby preserving soil health and reducing environmental impact. The use of the Random Forest algorithm enhances prediction reliability compared to conventional rule-based methods, enabling adaptive and data-driven decision making under varying soil conditions.

The modular and scalable design of the proposed system allows easy integration of additional sensors and advanced analytical models, making it adaptable to diverse agricultural requirements. Overall, the system proved to be reliable, cost-effective, and capable of providing real-time fertilizer recommendations with minimal human intervention. The successful operation of the prototype under practical conditions highlights its suitability for deployment in modern farming environments and validates the role of machine learning-assisted systems in advancing sustainable agriculture practices.

VIII. FUTURE SCOPE

In the future, the fertilizer recommendation system can be further enhanced to improve accuracy, scalability, and usability. Additional soil parameters such as pH level, electrical conductivity, and micronutrients can be integrated to provide more comprehensive soil analysis. Incorporating crop-specific requirements and growth stages will allow the system to generate customized fertilizer recommendations tailored to different crops and seasons.

Cloud connectivity and IoT integration can enable remote monitoring and data storage, allowing farmers to access soil health information through mobile or web applications. Advanced machine learning techniques such as deep learning or reinforcement learning may be employed to predict long-term soil fertility trends and optimize fertilizer schedules automatically. Integration with GPS and weather forecasting systems can further improve decision-making by considering climatic variations.

Energy-efficient operation can be enhanced by integrating renewable energy sources such as solar power for field deployment. Expanding the system to support large-scale farms, multi-field monitoring, and compatibility with commercial agricultural equipment will transform it into a fully scalable and intelligent agricultural decision support system, contributing to sustainable and precision farming solutions.

IX. REFERENCES

- [1] Aruvansh Nigam, Saksham Garg, Archit Agrawal "Crop Yield Prediction using ML Algorithms", International Journal of Engineering Research & Technology (IJERT), 2019.
- [2] R. Kaur and S. K. Soni, "A review of crop prediction models and remote sensing based approaches for crop yield forecasting," International Journal of Agricultural and Environmental Information Systems, vol. 8, no. 1, pp. 1-22, 2017.
- [3] R. Garg, P. Jain, and R. K. Jain, "Crop yield prediction using machine learning algorithms: A review," Journal of Agricultural Science and Technology, vol. 21, no. 3, pp. 501-516, 2019.
- [4] S. Saha, S. Ghosh, and M. Pal, "A review on machine learning approaches for crop yield prediction," International Journal of Computer Applications, vol. 182, no. 27, pp. 23-30, 2018.
- [5] D. Kumar, A. Kumar, and S. Kumar, "Crop yield prediction using machine learning techniques: A review," Journal of King Saud University-Computer and Information Sciences, vol. 30, no. 2, pp. 163-178, 2018.

- [6] N. Dubey and D. Kumar, "A survey of machine learning techniques for crop yield prediction," *International Journal of Computer Sciences and Engineering*, vol. 7, no. 4, pp. 416-423, 2019.
- [7] S. K. Roy, S. C. Paul, M. S. Islam, and M. A. Hasan, "Crop yield prediction using machine learning algorithms: A comprehensive review," *IEEE Access*, vol. 7, pp. 78907-78926, 2019.
- [8] P. Patil, A. Shinde, and M. Kadam, "A survey on crop yield prediction using data mining techniques," *International Journal of Advanced Research in Computer Science*, vol. 8, no. 1, pp. 421-425, 2017.
- [9] S. Kumar, M. R. Gupta, and M. S. Bhatia, "Crop yield prediction using machine learning algorithms: A comprehensive review," *International Journal of Computer Applications*, vol. 179, no. 6, pp. 1-6, 2018.
- [10] M. Kamble, D. Kumar, and P. Kumar, "A review on crop yield prediction using machine learning techniques," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 7, no. 4, pp. 260-267, 2017.
- [11] H. Yu and S. Liu, "A survey on crop yield prediction models and methods," *Journal of Northeast Agricultural University*, vol. 25, no. 1, pp. 1-10, 2018.
- [12] H. Ma, X. Zhang, Y. Ma, and Y. Wang, "A survey on crop yield prediction based on deep learning," in *2018 IEEE International Conference on Big Data and Smart Computing (Big Comp)*, Shanghai, China, 2018, pp. 335-340.
- [13] S. T. K. Das and S. S. Mohanty, "Prediction of crop yield using machine learning techniques: A review," *International Journal of Agricultural and Environmental Information Systems*, vol. 10, no. 3, pp. 1-16, 2019.
- [14] V. N. Nguyen, Q. N. Nguyen, and T. N. Huynh, "A comparative study of machine learning algorithms for crop yield prediction," in *2020 International Conference on Advanced Technologies for Communications (ATC)*.
- [15] S. R. Singh, A. Singh, and V. P. Singh, "Crop yield prediction using machine learning: A review," *Journal of Big Data*, vol. 6, no. 1, pp. 1-27, 2019.
- [16] D. K. Pal and B. Datta, "Crop yield prediction: A review," *Journal of Agrometeorology*, vol. 16, no. 2, pp. 121-130, 2014.
- [17] S. Kumar, N. Kumar, and D. Kumar, "Crop yield prediction using machine learning algorithms: A comprehensive review," in *2018 IEEE 4th International Conference on Computational Intelligence and Communication Networks (CICN)*, pp. 68-72.
- [18] N. K. Gupta and V. G. S. Kumar, "Crop yield prediction using machine learning algorithms: A review," in *2018 International Conference on Smart Computing and Informatics (SCI)*, pp. 647-652.
- [19] R. Dubey and S. Jain, "A review of crop yield prediction models and methods," *Journal of Engineering Science and Technology*, vol. 13, no. 9, pp. 2720-2737, 2018.
- [20] A. Sahay and P. N. V. N. N. Murthy, "Crop yield prediction using machine learning: A review," in *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, pp. 1-6.
- [21] S. G. S. Saini, S. S. Rajput, and S. S. Khillare, "Crop yield prediction using data mining techniques: A review," in *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, pp. 1953-1957.
- [22] M. Manasa, S. S. Kumar, and A. Kumar, "A review on crop yield prediction using machine learning techniques," in *2018 International Conference on Computing, Power and Communication Technologies (GUCON)*, pp. 901-905.