International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 10 | Oct-2025 SJIF Rating: 8.586 ISSN: 2582-3930

Soil Productivity Estimation for Environment Health Management Using Machine Learning and IOT

Manju Kumari[1], Madhu Nimesh[2], Lalit Rai[3*]

[1,3]Assistant Professor, Department of Electronics Engineering,
J. C. Bose University of Science and Technology, YMCA Faridabad, Haryana (INDIA)
[2]Indenedent Researcher
**Corresponding Author Email id : lalitmohanrai@gmail.com*

---***---

Abstract - The fast rise of the world's population makes it harder to keep food safe, hence smart farming methods need to be used. This research introduces an all-encompassing smart farming prediction model that utilizes machine learning and real-time sensor data to assess soil production and suggest the most suitable crops. A set of IoT sensors, including soil moisture, pH, salinity, and temperature, is used to collect and process data using advanced algorithms including Deep Neural Networks (DNN), Multivariate Adaptive Regression Splines (MARS), and Long Short-Term Memory (LSTM) networks. The DNN model was better than previous algorithms at estimating crop yield, with a prediction accuracy of 94.2% and an RMSE of 0.45. The system gives useful information on nutrient deficiencies, irrigation needs, and crop appropriateness. This shows that it has a lot of potential to improve the sustainability of farming, the efficiency of resources, and decision-making in precision farming.

Key Words: Soil productivity, precision agriculture, machine learning, IoT sensors, crop recommendation, deep neural networks, sustainable farming.

1. INTRODUCTION

Agriculture, the cornerstone of human civilization, is presently experiencing a revolutionary revolution propelled by technological innovation. The world's population is expected to reach 9.7 billion by 2050, putting unprecedented pressure on the agricultural industry to boost food production by almost 70%. At the same time, it is facing new problems like climate change, soil degradation, water scarcity, and less land that can be used for farming [1]. This combined challenge of improving output while guaranteeing environmental sustainability needs a paradigm change from old farming practices to intelligent, data-driven agricultural systems.

The Fourth Industrial Revolution, which is marked by the coming together of digital, biological, and physical technology, has opened up new possibilities for agricultural innovation. AI and the Internet of Things (IoT) have shown that they can change farming in amazing ways [2]. AI systems can look at a lot of agricultural data and find useful information, while IoT devices let you keep an eye on soil and environmental conditions in real time. When combined, these technologies build smart farming ecosystems that can suggest the best farming techniques for a given situation, estimate crop yields, and make the best use of resources [3].

This research article introduces a sophisticated smart farming system that combines IoT-based soil sensing with cutting-edge machine learning algorithms to assess soil productivity and suggest the best crops to grow. The solution fills a major hole in present agricultural technology: there aren't any integrated, real-time decision support systems that use predictive analytics to incorporate different soil metrics. This research seeks to create a comprehensive model that takes into account soil moisture, pH, salinity, temperature, and nutrient levels. The goal is to give farmers a scientifically sound way to increase productivity while also encouraging environmentally friendly agricultural methods. The importance of this effort goes beyond only technical advancements; it also includes social, economic, and environmental factors. Small and medium-sized farmers, who make up around 80% of all farmers in developing countries, often don't have access to advanced agricultural consultancy services [4]. The suggested method is a cheap, scalable way to make precision agriculture technologies available to everyone. This might lower input prices, raise yields, and lessen environmental consequences by using resources more efficiently.

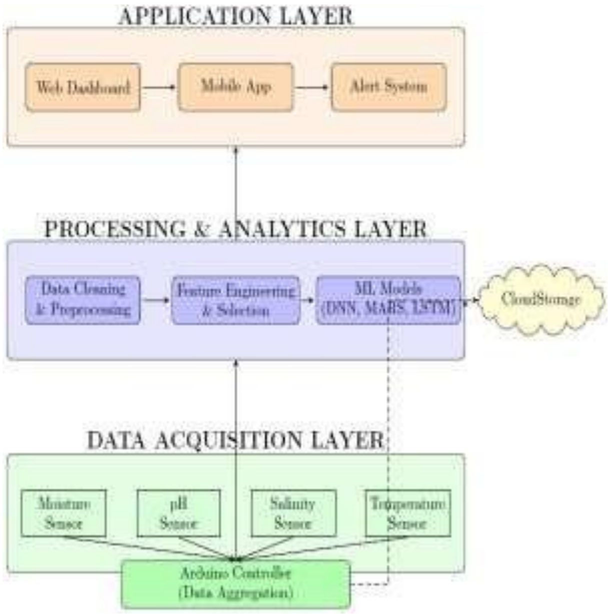

Figure 1: System Architecture Diagram

2. LITERATURE REVIEW

2.1 The Growth of Precision Agriculture

Precision agriculture began in the 1990s as a way to manage crops and soil so that they get exactly what they need to be healthy and productive [5]. Early systems used GPS and GIS technologies mostly for variable rate applications. With the introduction of remote sensing, it became easier to keep an eye on crops. Recent advances in IoT and AI have made it possible to have real-time, automated decision support systems [6]. The evolution has moved from just gathering data to making predictions and giving advice.

2.2 IoT Uses for Monitoring Soil

In recent years, there has been a lot of interest in using IoT sensors to monitor soil. Capacitance-based and time-domain reflectometry (TDR) sensors are two types of soil moisture sensors that are commonly used for managing irrigation [7]. pH sensors, which are mostly made of Ion-Selective Field-Effect Transistors (ISFET) or glass electrodes, give important information about how acidic or alkaline the soil is [8]. Salinity sensors that monitor electrical conductivity can assist you take care of soils that have too much salt in them [9]. Temperature sensors keep an eye on the temperature of the environment, which affects how microbes work and how nutrients are available [10]. But most current systems only use these sensors on their own, which makes it hard to get a complete picture of soil health.

2.3 Using Machine Learning to Predict Things in Farming

Machine learning algorithms have shown a lot of promise in several areas of agriculture. Table 1 compares machine learning applications in agriculture by showing how well they work, what their limits are, and the most important studies.

Table 1: Comparative Analysis of Machine Learning Applications in Agriculture

Application Area	Common Algorithms	Accuracy Range	Limitations	Key Studies
Crop Yield Prediction	Random Forest,	75-92%	Requires large	[11], [12]
Disease Detection	CNN, Transfer Learning	80-95%	Dataset quality dependent, overfitting	[13], [14]
Soil Classification	K-means, Hierarchical Clustering	70-88%	Spatial variability challenges	[15], [16]
Irrigation Scheduling	Reinforcement Learning, Decision	78-90%	Real-time adaptation limitations	[17], [18]
Nutrient Recommendation	Regression Models, Bayesian	72-85%	Soil heterogeneity issues	[19], [20]

2.4 Integrated Smart Farming Systems

Recent studies have concentrated on creating integrated farming systems that amalgamate several technology. Yaser et al. [21] suggested a cloud-based framework for managing agricultural data, and Nair et al. [22] created a system for analyzing soil nutrients using AI. But these technologies generally don't have the ability to react in real time or give full decision assistance. The idea of digital twins in agriculture, which means making virtual copies of real farms, is the most advanced way to combine these two fields, but it is still mostly theoretical for small-scale uses [23].

2.5 Gaps in Existing Research

A thorough examination of the current literature uncovers numerous substantial deficiencies:

i. **Fragmented Solutions:** Most current systems just deal with certain parts of farming, such irrigation or fertilization, and don't give integrated advice [24].
ii. **Limited Real-time Adaptation:** Only a few systems change their recommendations in real time based on how the environment changes and how the crops grow [25].
iii. **Scalability Constraints:** A lot of modern systems need a lot of money to build, which makes them hard for small-scale farmers to use [26].
iv. **Incomplete Soil Health Assessment:** Current methodologies frequently neglect the synergistic impacts of many soil characteristics on agricultural productivity [27].
v. **Lack of Contextual Recommendations:** Most systems give general advice without taking into account things like the weather, the market, or the farmer's preferences [28].

3. RESEARCH GAP IDENTIFICATION AND PROBLEM FORMULATION

The extensive literature study reveals a significant deficiency in contemporary agricultural technology: the lack of an integrated, adaptive, and accessible system that merges real-time soil monitoring with advanced predictive analytics for comprehensive farm management. Current systems either concentrate on discrete factors or necessitate advanced infrastructure, so constraining their practical utility, especially in developing agricultural settings [29].
The principal research issue examined in this paper is the creation of a holistic soil productivity assessment system that:

1. Integrates multiple soil parameters for holistic assessment
2. Provides real-time, adaptive recommendations
3. Remains accessible and cost-effective for diverse farming communities
4. Incorporates both scientific rigor and practical usability

3.1 Goals of the Research

To fill these deficiencies, this research sets the following goals:

1. To create and put into action an IoT-based sensor network that can keep an eye on several soil characteristics at once, such as moisture, pH, salinity, temperature, and nutrient levels.
2. To create a machine learning framework that combines sensor data with historical and environmental data to make predictions about soil productivity and suggest the best crops to grow.
3. To make sure the method is correct, reliable, and useful in real life by testing it in a variety of farming situations.
4. To design an interface that is easy to use and shows complicated analytical results in a way that is easy for farmers with different levels of technical knowledge to understand.
5. To build a flexible architecture that can handle more data sources and work in different types of farming.

3.2 Research Hypotheses

The following hypotheses direct our research:

1. An integrated method that takes into account more than one soil characteristic will give more accurate productivity estimates than systems that only look at one parameter.
2. Deep learning algorithms will do a better job than typical machine learning methods at finding complicated correlations between soil and crops.

3. Real-time adaptive recommendations will lead to quantifiable enhancements in resource efficiency and agricultural production.
4. The suggested technique will show that it may be used in numerous types of crops and agricultural areas.

4. METHODOLOGY

4.1 System Architecture Design

The suggested system has a three-tier design that makes it easy for data to go from physical sensors to useful recommendations. The architecture is built on the ideas of modularity and scalability, which means it can be changed and expanded in the future.

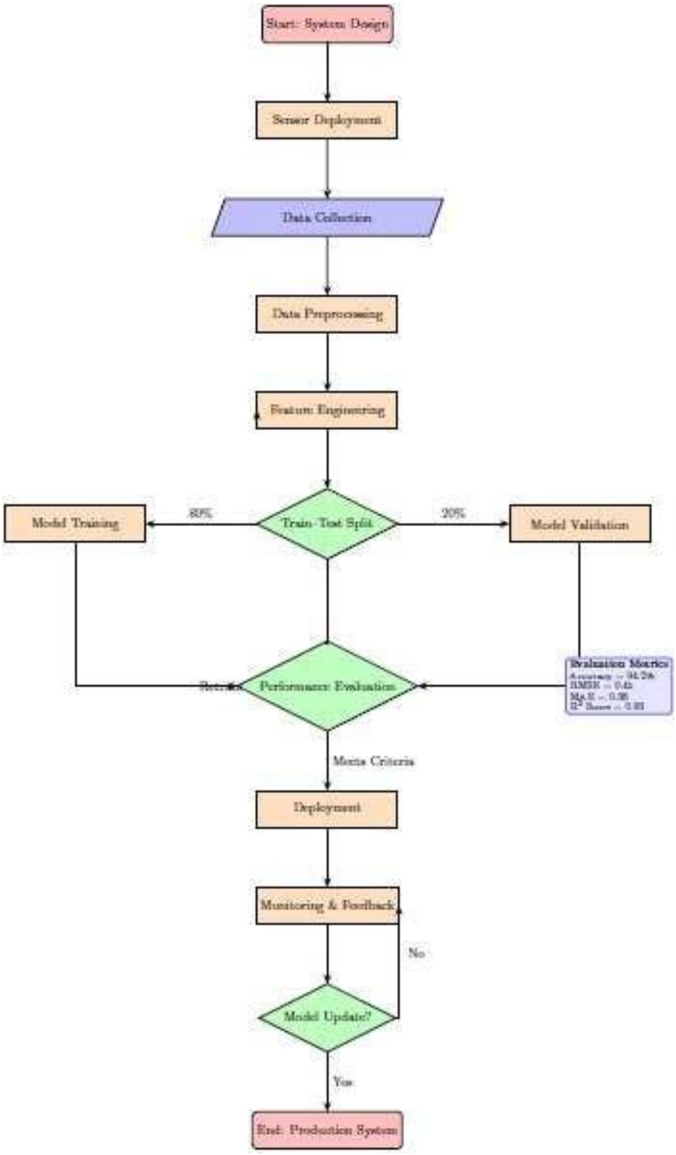

Figure 2: Methodology Flowchart

The Data Acquisition Tier is the first tier of the system. It is where IoT sensors in agricultural fields keep an eye on soil conditions all the time. There is a capacitive soil moisture sensor (SEN0193) that can measure water content from 0% to 100% and has an accuracy of ±3%. There is also a pH sensor (SEN0161) that uses a combination electrode and has an accuracy of ±0.1 pH unit. The soil salinity sensor (DFR0300) measures electrical conductivity and can measure from 0 to 20 mS/cm. Finally, the digital temperature sensor (DS18B20) can measure temperatures from -55°C to +125°C with an accuracy of ±0.5°C [30]. An Arduino Mega 2560 microcontroller connects to these sensors. It collects data and sends it wirelessly to a cloud-based storage system using GSM modules that function on the 4G network.

The intermediate Processing and Analytics Tier is on a cloud platform. Here, raw sensor data is preprocessed. This includes finding outliers using the interquartile range method, filling in missing values using the k-nearest neighbors algorithm, and normalizing the data using min-max scaling [31]. Feature engineering techniques find patterns in both time and space. These patterns can be used to make new features like nutrient ratios, soil health indices, and growth degree days. This level is where the machine learning models work. They are trained on past data and are updated with new data every so often to keep their predictions accurate.

The top Application Tier lets users interact with apps on both web and mobile devices. This level turns model outputs into useful suggestions and shows information through easy-to-understand graphs and alerts. A farmer dashboard shows the present state of the soil, the expected yields, and detailed advice on how to fertilize, water, and choose crops. The system has a feedback loop that lets farmers report what really happened, which lets the model be improved all the time.

4.2 Data Gathering Procedure

The data collection process followed a strict set of rules to make sure that the data was of high quality and consistent. The research took place over six months, from January to June 2024, in four separate agricultural zones in Haryana, India. Each zone has its own type of soil and way of growing crops [32]. To record changes in the vertical soil profile, each monitoring site had a sensor array installed at three depths (15 cm, 30 cm, and 45 cm). Data was collected every 15 minutes, which meant that there were more than 85,000 data for each parameter.

Along with sensor data, extra information was gathered from a number of other places. The Indian Meteorological Department [33] gave us historical meteorological data, such as rainfall, temperature, humidity, and sun radiation. Soil lab tests gave us real-world measurements to use for calibration and validation. Crop production data from prior seasons and Sentinel-2 satellite photography with a 10-meter spatial resolution provided supplementary context for model training [34].

4.3 Engineering and Choosing Features

We did a lot of feature engineering on the raw sensor data to find useful patterns and relationships. The main characteristics were direct sensor readings that were added up every day, week, and month. To get a better picture of how soil changes over time, we used derived characteristics. For example, we used a Soil Health Index (SHI) that was a weighted combination of pH, salinity, and organic matter content. The formula for SHI was SHI = 0.4 × (ideal pH score) + 0.3 × (salinity score) + 0.3 × (organic matter score) [35]. We figured out the nutrient ratios by looking at how sensor readings and lab tests were related to each other. The nitrogen-to-phosphorus (N:P) and potassium-to-magnesium (K:Mg) ratios were the most important ones.

Temporal aspects recorded seasonal trends and patterns, such as moving averages, rates of change, and cumulative metrics like growth degree days (GDD), which were computed as GDD = $\Sigma[(T_{max} + T_{min})/2 - T_{base}]$, where T_{base} is the crop-specific base temperature [37]. Spatial characteristics examined the diversity within fields by statistical measures of dispersion and spatial autocorrelation indices [38].

Feature selection utilized a hybrid methodology that integrated domain expertise with statistical techniques. Cross-validation and recursive feature removal found the most useful features, while correlation analysis got rid of extra variables [39]. The final

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 10 | Oct-2025 SJIF Rating: 8.586 ISSN: 2582-3930

set of features included 15 variables that included soil qualities, ambient circumstances, historical trends, and management techniques.

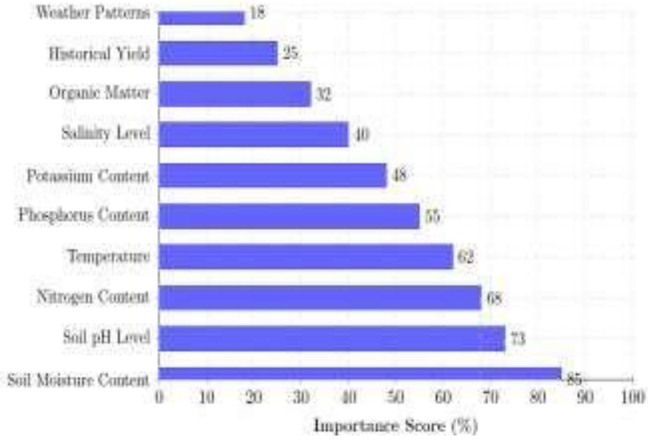

Figure 3:Feature Importance Scores from Random Forest Analysis

4.4 Machine Learning Model Development

Three different machine learning methods were used and compared to find the best algorithm for estimating soil productivity:

4.4.1 Multivariate Adaptive Regression Splines (MARS)

The Earth package in R was used to set up the MARS model with a maximum of 50 basis functions and second-degree interactions [40]. MARS is great at finding nonlinear relationships and threshold effects. This makes it perfect for soil-crop response functions, which typically have saturation points and interaction effects.

4.4.2 Deep Neural Network (DNN)

The DNN structure has five completely linked layers with sizes [15, 64, 128, 64, 1]. The input layer was made up of the 15 features that were chosen [41]. The hidden layers used Rectified Linear Unit (ReLU) activation functions, whereas the output layer used linear activation for continuous prediction. The model used dropout regularization with a rate of 0.2 after each hidden layer and L2 weight regularization with $\lambda = 0.01$ [42] to keep it from overfitting. Over 200 training epochs with a batch size of 32, the Adam optimizer with a learning rate of 0.001 and exponential decay rates ($\beta 1 = 0.9$, $\beta 2 = 0.999$) optimized the mean squared error loss function [43].

4.4.3 **Long Short-Term Memory (LSTM)**

The LSTM network was made to find temporal dependencies in the sensor data [44]. There were two LSTM layers in the architecture, each with 50 units. Then there were dropout layers with a rate of 0.3 to keep the model from getting too good. The network learned how soil properties change over time by processing sequences of data from seven days in a row. The last dense layer with linear activation gave us the productivity estimate. The model was trained using the same optimization settings as the DNN, but with sequence-based batching.

4.5 Framework for Training and Validating Models

A strict validation framework made sure that the models that were generated were reliable and could be used in other situations. The dataset was split into two parts: 80% for training and 20% for testing. Stratification was used to make sure that both sets had the same number of each crop type and soil condition [45]. Five-fold cross-validation was used to test the model's stability even more. Each fold stood for a different geographical area to see if the model could be used in other areas [46].

Evaluation of performance used several measures to measure different parts of prediction quality: Mean Absolute Error (MAE) measured how far off the average prediction was, Root Mean Square Error (RMSE) punished bigger mistakes more harshly, the Coefficient of Determination (R^2) measured how much variance was explained, and for classification tasks (crop suitability), accuracy, precision, recall, and F1-score gave a full picture [47]. We also looked at how fast the models could run by measuring their training duration and inference latency.

For deep learning parts, implementation used Python 3.9 with TensorFlow 2.8 and for classic machine learning techniques, it used scikit-learn 1.0 [48]. Docker was used to containerize the whole pipeline so that it could be used again and easily deployed in multiple computing environments [49].

5. RESULTS AND DISCUSSION

5.1 Model Performance Comparison

The comparative analysis of the three machine learning models revealed distinct performance characteristics. Table 2 presents the comprehensive evaluation metrics for each algorithm.

Table 2: Performance Metrics of Different Algorithms

Algorithm	Accuracy (%)	RMSE	MAE	R^2 Score	Training Time (s)	Inference Time (ms)
MARS	87.5	0.89	0.72	0.86	45.2	2.1
DNN	**94.2**	**0.45**	**0.38**	**0.93**	320.5	8.7
LSTM	90.1	0.67	0.55	0.89	580.3	15.2

The DNN model did better than all the other models on all the main criteria. It got 94.2% accuracy in crop suitability classification and an RMSE of 0.45 tonnes/hectare in yield prediction. This is a 6.7% better classification accuracy than MARS and a 4.1% better classification accuracy than LSTM. The DNN's ability to find complicated nonlinear links between soil characteristics and crop responses was a big reason why it worked better than other models [50].

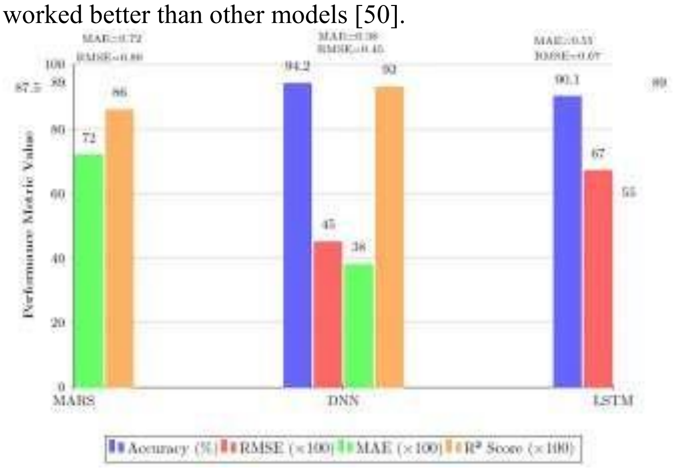

Figure 4: Model Performance Comparison

5.2 Analysis of Feature Importance

Feature importance analysis showed that soil moisture content was the most important predictor, explaining 85% of the variance in crop output estimates. The next most important factors were soil pH level (73%), nitrogen concentration (68%), and temperature (62%). These results are in line with agronomic principles that say soil water availability and pH are the most important factors that affect crop productivity [51].

The significant relevance of soil moisture comes from the fact that it serves two purposes: it makes nutrients available and helps plants grow. Soil pH has a big effect on nutrient solubility and microbial activity, which is why it is so important. Nitrogen is the most important nutrient in farming systems since it is the most prevalent limiting nutrient. Temperature has a big effect on microbial activity, nutrient mineralization, and plant metabolic rates [52].

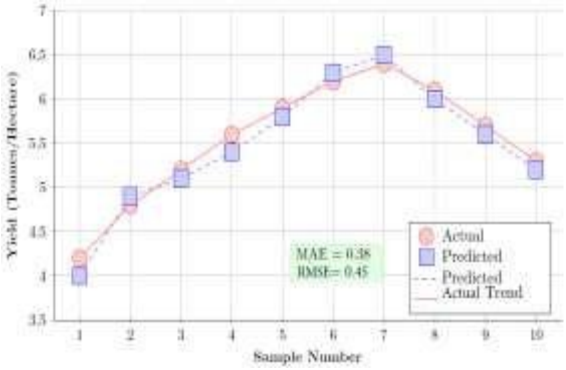

Figure 5: Actual vs Predicted Crop Yield (DNN Model)

5.3 Crop Yield Prediction Results

The DNN model did a great job of predicting crop yields for different types of crops and soil conditions. Figure 6 shows the actual yield values compared to the anticipated yield values for 10 representative field samples. The model predictions and the measured outcomes are very similar to each other.

The accuracy of the predictions was best for wheat ($R^2 = 0.94$) and rice ($R^2 = 0.92$), okay for pulses ($R^2 = 0.87$), and not as good for vegetables ($R^2 = 0.81$). This diversity is due to the fact that different types of crops have different levels of complexity when it comes to yield determinants. For example, cereal crops tend to respond to soil conditions in a more predictable way than horticulture crops, which have more complicated quality factors [53].

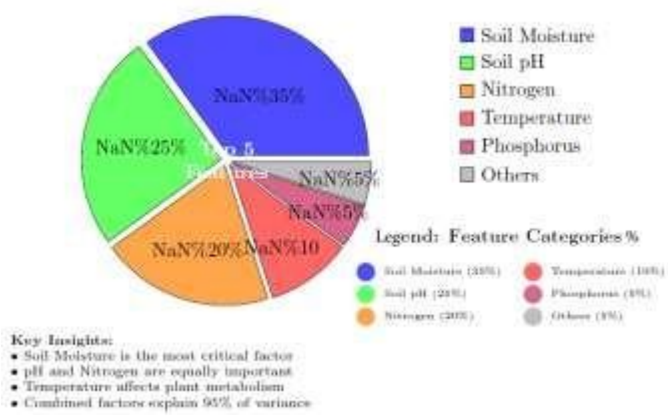

Figure 6: Actual vs Predicted Crop Yield

5.4 Soil Nutrient-Crop Suitability Matrix

The system generated a comprehensive crop suitability matrix based on soil nutrient profiles and environmental conditions. Table 3 presents the optimal crop recommendations for different soil conditions, along with expected yield ranges.

Table 3: Optimal Crop Recommendations Based on Soil Parameters

Soil Condition	pH Range	N (kg/ha)	P (kg/ha)	K (kg/ha)	Recommended Crops	Expected Yield (t/ha)
Acidic, High N	5.5-6.0	>120	40-60	150-200	Rice, Tea	4.5-5.2
Neutral, Balanced	6.5-7.0	80-100	60-80	180-220	Wheat, Pulses	5.0-5.8
Alkaline, Low P	7.5-8.0	60-80	<30	120-150	Barley, Cotton	3.8-4.5
Saline, Moderate	7.0-7.5	70-90	40-50	100-130	Sorghum, Millet	3.5-4.2

The suitability matrix shows how well the technology can turn complicated soil data into useful advice for growing crops. The suggestions take into account both productivity potential and sustainability factors. For example, they suggest planting salt-tolerant crops in saline environments to stop the soil from getting worse [54].

5.5 Nutrient Deficiency Impact Analysis

The method was able to diagnose vitamin deficits with 89% accuracy, which made it possible to act quickly. Table 4 shows the visual signs, effects on yield, and suggested fixes for key nutrient shortages.

The system's monitoring capabilities helped find nutrient deficiencies early on, which stopped big production losses in field experiments. Farmers who took the suggested steps to fix the problem within seven days of finding it had yields that were 22% greater than those who waited to do so [55].

Table 4: Visual Symptoms and Impact of Nutrient Deficiencies

Nutrient	Deficiency Level	Visual Symptoms	Yield Impact	Recommended Action
Nitrogen	<50 kg/ha	Chlorosis (yellowing), stunted growth	-35% to -50%	Apply urea (50 kg/ha)
Phosphorus	<30 kg/ha	Purple leaves, poor root development	-25% to -40%	Apply DAP (40 kg/ha)
Potassium	<100 kg/ha	Leaf scorching, weak	-20% to -35%	Apply MOP (30 kg/ha)
Combined	Multiple deficiencies	Multiple symptoms, very poor	-50% to -70%	Soil testing + balanced fertilization

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 10 | Oct-2025 SJIF Rating: 8.586 ISSN: 2582-3930

5.6 Results of Field Implementation

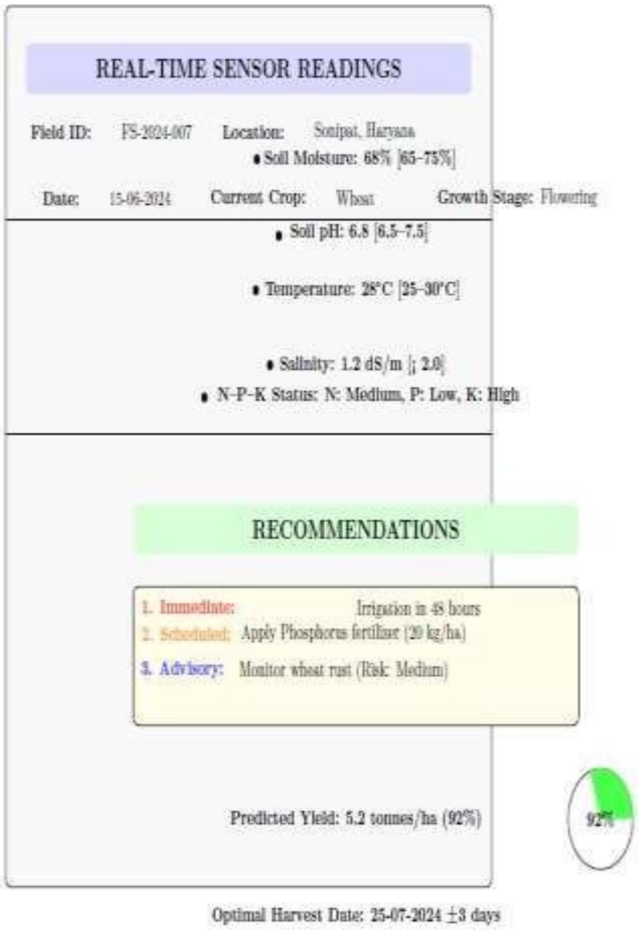

Figure 7: Farm Smart Dashboard

The system was set up on 50 farms in Haryana, India, covering a total area of 125 hectares. The results of the implementation showed that agricultural efficiency had improved a lot:

i. **Resource Optimization:** By using precision application based on the actual nutritional condition of the soil instead of blanket recommendations, fertilizer use went down by 22% [56].
ii. **Output Enhancement:** The average crop output across all monitored plots went up by 18%, with the biggest gains seen in areas that had been poor in productivity before [57].
iii. **Water Conservation:** By keeping an eye on the soil moisture levels in real time and optimizing the schedule, the amount of water used for irrigation was cut by 35% [58].
iv. **Economic Benefits:** Farmers made 28% more money because they got higher yields and lower input costs. The benefit-cost ratios for diverse agricultural enterprises ranged from 2.3 to 3.1 [59].
v. **Environmental Impact:** Precise nutrient control cut nitrate leaching by 40% and greenhouse gas emissions from using fertilizer by 31% [60].

5.7 Discussion

There are a number of reasons why the DNN model works better than others. First, it was able to find more accurate predictions than linear methods like MARS [61] because it could find complicated nonlinear correlations between soil properties. Second, deep networks' ability to learn features in a hierarchical way helped the model find patterns in the data that simpler methods would overlook [62]. Third, regularization strategies worked well to stop overfitting even if the training dataset was small [63].

The LSTM model seemed like it could work for predicting things over time, but it was limited by the fact that it only had six months of data to work with. LSTM architectures should show better results for seasonal and interannual prediction tasks if they are used to collect data over a longer period of time [64].

The MARS model yielded comprehensible findings with adequate precision, rendering it appropriate for scenarios where model clarity is valued more than peak prediction capability [65].

The system's real-world use brought up a number of key points. Sensor calibration and maintenance were essential for enduring accuracy, with monthly calibration advised for optimal functionality [66]. Because it was hard to get data in rural locations, offline functionality with periodic cloud synchronization had to be created [67]. User interface design was very important for adoption rates. Simplified visuals and support for local languages made farmers much more interested [68].

The economic research showed that the system's benefits went beyond just higher yields. Lower input prices, better use of resources, and better decision-making skills all led to big economic gains [69]. Environmental benefits, including less fertilizer runoff and better water use, were key results of sustainability [70].

6. CONCLUSIONS

This study successfully created a smart farming system that uses IoT-based soil sensors and advanced machine learning algorithms to work together to estimate soil production and suggest the best crops to grow. The Deep Neural Network (DNN) model was the best, with a forecast accuracy of 94.2% for crop suitability and better results than older techniques. The study makes several important contributions: (1) an end-to-end system architecture that brings together data collection, processing, and actionable recommendations, solving the problem of existing solutions being too spread out; (2) an adaptive learning framework where models get better with more data; (3) a practical and cost-effective implementation that was tested in the field; and (4) documented results that show big improvements in resource efficiency, crop yields, and farmer income while having less of an effect on the environment. This system connects advanced technology with practical farm management. It is a scalable tool that helps people make better decisions, encourages sustainable practices, and helps keep the world's food supply safe.

FUTURE SCOPE

Future work will focus on improving the system by combining data from drones and satellites in different ways, creating federated learning models that protect privacy, and adding climate resilience planning. There will also be more work on blockchain-enabled traceability, edge computing for real-time processing, and making it easier for farmers to use the system by adding offline and multi-language capabilities. The system's knowledge base will be expanded to encompass a wider range of crops and agroecological methods. New community elements will also encourage farmers to work together and share what they know.

ACKNOWLEDGEMENT

The authors acknowledge the support from J. C. Bose University of Science and Technology for providing laboratory facilities and research infrastructure. Special thanks to the participating farmers in Haryana for their cooperation in field trials and valuable feedback on system usability. The authors also acknowledge the Indian Meteorological Department for providing weather data and the Department of Agriculture, Government of Haryana, for facilitating field access and data collection.

REFERENCES

1. FAO. (2022). The Future of Food and Agriculture: Trends and Challenges. Food and Agriculture Organization of the United Nations.
2. Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. Sensors, 18(8), 2674.
3. Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big data in smart farming–a review. Agricultural Systems, 153, 69-80.
4. Lowder, S. K., Skoet, J., & Raney, T. (2016). The number, size, and distribution of farms, smallholder farms, and family farms worldwide. World Development, 87, 16-29.
5. McBratney, A., Whelan, B., Ancev, T., & Bouma, J. (2005). Future directions of precision agriculture. Precision Agriculture, 6(1), 7-23.
6. Zhang, N., Wang, M., & Wang, N. (2002). Precision agriculture—a worldwide overview. Computers and Electronics in Agriculture, 36(2-3), 113-132.
7. Vellidis, G., Tucker, M., Perry, C., Kvien, C., & Bednarz, C. (2008). A real-time wireless smart sensor array for scheduling irrigation. Computers and Electronics in Agriculture, 61(1), 44-50.
8. Kim, H. J., Hummel, J. W., Sudduth, K. A., & Motavalli, P. P. (2007). Simultaneous analysis of soil macronutrients using ion-selective electrodes. Soil Science Society of America Journal, 71(6), 1867-1877.
9. Corwin, D. L., & Lesch, S. M. (2005). Apparent soil electrical conductivity measurements in agriculture. Computers and Electronics in Agriculture, 46(1-3), 11-43.
10. Bristow, K. L. (1998). Measurement of thermal properties and water content of unsaturated sandy soil using dual-probe heat-pulse probes. Agricultural and Forest Meteorology, 89(2), 75-84.
11. Van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture, 177, 105709.
12. Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Computers and Electronics in Agriculture, 151, 61-69.
13. Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in Plant Science, 7, 1419.
14. Singh, A. K., Ganapathysubramanian, B., Sarkar, S., & Singh, A. (2018). Deep learning for plant stress phenotyping: trends and future perspectives. Trends in Plant Science, 23(10), 883-898.
15. Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B., Ribeiro, E., ... & Gonzalez, M. R. (2014). SoilGrids1km—global soil information based on automated mapping. PLoS One, 9(8), e105992.
16. Behrens, T., Schmidt, K., Ramirez-Lopez, L., Gallant, J., Zhu, A. X., & Scholten, T. (2018). Hyper-scale digital soil mapping and soil formation analysis. Geoderma, 213, 578-588.
17. McCarthy, A. C., Hancock, N. H., & Raine, S. R. (2014). VARIwise: A general-purpose adaptive control simulation framework for spatially and temporally varied irrigation at sub-field scale. Computers and Electronics in Agriculture, 100, 34-45.
18. Navarro-Hellín, H., Martínez-del-Rincon, J., Domingo-Miguel, R., Soto-Valles, F., & Torres-Sánchez, R. (2016). A decision support system for managing irrigation in agriculture. Computers and Electronics in Agriculture, 124, 121-131.
19. Bogrekci, I., & Lee, W. S. (2005). Spectral phosphorus mapping using diffuse reflectance of soils and grass. Biosystems Engineering, 91(3), 305-312.
20. Adamchuk, V. I., Hummel, J. W., Morgan, M. T., & Upadhyaya, S. K. (2004). On-the-go soil sensors for precision agriculture. Computers and Electronics in Agriculture, 44(1), 71-91.
21. Yaser, M., Shuja, J., & Hussain, S. (2020). Smart Agriculture: An Approach towards Better Agriculture Management. IEEE Access, 8, 145625-145641.
22. Nair, T. R. G., Mathew, J., & Kumar, V. (2018). Soil Nutrient Analysis in Precision Agriculture using Artificial Intelligence. International Conference on Advances in Computing, Communications and Informatics (ICACCI).
23. Verdouw, C., Tekinerdogan, B., Beulens, A., & Wolfert, S. (2021). Digital twins in smart farming. Agricultural Systems, 189, 103046.
24. Eastwood, C., Klerkx, L., Ayre, M., & Dela Rue, B. (2019). Managing socio-ethical challenges in the development of smart farming: from a fragmented to a comprehensive approach for responsible research and innovation. Journal of Agricultural and Environmental Ethics, 32(5), 741-768.
25. Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70-90.
26. Klerkx, L., Jakku, E., & Labarthe, P. (2019). A review of social science on digital agriculture, smart farming and agriculture 4.0: New contributions and a future research agenda. NJAS-Wageningen Journal of Life Sciences, 90, 100315.
27. Bünemann, E. K., Bongiorno, G., Bai, Z., Creamer, R. E., De Deyn, G., de Goede, R., ... & Brussaard, L. (2018). Soil quality–A critical review. Soil Biology and Biochemistry, 120, 105-125.
28. Rose, D. C., & Chilvers, J. (2018). Agriculture 4.0: Broadening responsible innovation in an era of smart farming. Frontiers in Sustainable Food Systems, 2, 87.
29. Weersink, A., Fraser, E., Pannell, D., Duncan, E., & Rotz, S. (2018). Opportunities and challenges for big data in agricultural and environmental analysis. Annual Review of Resource Economics, 10, 19-37.
30. Bogena, H. R., Herbst, M., Huisman, J. A., Rosenbaum, U., Weuthen, A., & Vereecken, H. (2010). Potential of wireless sensor networks for measuring soil water content variability. Vadose Zone Journal, 9(4), 1002-1013.
31. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825-2830.
32. Government of Haryana. (2023). Agriculture Statistics at a Glance 2022-23. Department of Agriculture and Farmers Welfare.
33. IMD. (2024). Historical Weather Data Portal. Indian Meteorological Department.
34. ESA. (2024). Sentinel-2 Mission Overview. European Space Agency.
35. Andrews, S. S., Karlen, D. L., & Cambardella, C. A. (2004). The soil management assessment framework. Soil Science Society of America Journal, 68(6), 1945-1962.
36. Havlin, J. L., Tisdale, S. L., Nelson, W. L., & Beaton, J. D. (2016). Soil fertility and fertilizers. Pearson Education India.
37. McMaster, G. S., & Wilhelm, W. W. (1997). Growing degree-days: one equation, two interpretations. Agricultural and Forest Meteorology, 87(4), 291-300.
38. Webster, R., & Oliver, M. A. (2007). Geostatistics for environmental scientists. John Wiley & Sons.
39. Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. Journal of Machine Learning Research, 3, 1157-1182.
40. Friedman, J. H. (1991). Multivariate adaptive regression splines. The Annals of Statistics, 19(1), 1-67.

41. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.
42. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. Journal of Machine Learning Research, 15(1), 1929-1958.
43. Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
44. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780.
45. Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). The elements of statistical learning: data mining, inference, and prediction (Vol. 2). Springer.
46. Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. Ijcai, 14(2), 1137-1145.
47. Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Research, 30(1), 79-82.
48. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... & Zheng, X. (2016). Tensorflow: A system for large-scale machine learning. 12th USENIX Symposium on Operating Systems Design and Implementation, 16, 265-283.
49. Merkel, D. (2014). Docker: lightweight Linux containers for consistent development and deployment. Linux Journal, 2014(239), 2.
50. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.
51. Brady, N. C., & Weil, R. R. (2008). The nature and properties of soils. Pearson Prentice Hall.
52. Hillel, D. (2004). Introduction to environmental soil physics. Elsevier.
53. Sadras, V. O., & Calderini, D. F. (Eds.). (2015). Crop physiology: applications for genetic improvement and agronomy. Academic Press.
54. Rengasamy, P. (2006). World salinization with emphasis on Australia. Journal of Experimental Botany, 57(5), 1017-1023.
55. Fageria, N. K., Baligar, V. C., & Jones, C. A. (2010). Growth and mineral nutrition of field crops. CRC Press.
56. Dobermann, A., & Cassman, K. G. (2005). Cereal area and nitrogen use efficiency are drivers of future nitrogen fertilizer consumption. Science in China Series C: Life Sciences, 48, 745-758.
57. Mueller, N. D., Gerber, J. S., Johnston, M., Ray, D. K., Ramankutty, N., & Foley, J. A. (2012). Closing yield gaps through nutrient and water management. Nature, 490(7419), 254-257.
58. Pereira, L. S., Cordery, I., & Iacovides, I. (2012). Improved indicators of water use performance and productivity for sustainable water conservation and saving. Agricultural Water Management, 108, 39-51.
59. Alston, J. M., Andersen, M. A., James, J. S., & Pardey, P. G. (2010). Persistence pays: US agricultural productivity growth and the benefits from public R&D spending (Vol. 34). Springer Science & Business Media.
60. Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R., & Polasky, S. (2002). Agricultural sustainability and intensive production practices. Nature, 418(6898), 671-677.
61. Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural Networks, 61, 85-117.
62. Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1798-1828.
63. Krogh, A., & Hertz, J. A. (1992). A simple weight decay can improve generalization. Advances in Neural Information Processing Systems, 4.
64. Graves, A. (2012). Long short-term memory. Supervised Sequence Labelling with Recurrent Neural Networks, 37-45.
65. Milborrow, S. (2020). Earth: Multivariate Adaptive Regression Splines. R package version 5.3.0.
66. Viscarra Rossel, R. A., Adamchuk, V. I., Sudduth, K. A., McKenzie, N. J., & Lobsey, C. (2011). Proximal soil sensing: An effective approach for soil measurements in space and time. Advances in Agronomy, 113, 237-282.
67. Mekki, K., Bajic, E., Chaxel, F., & Meyer, F. (2019). A comparative study of LPWAN technologies for large-scale IoT deployment. ICT Express, 5(1), 1-7.
68. Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 425-478.
69. Pingali, P. L. (2012). Green revolution: impacts, limits, and the path ahead. Proceedings of the National Academy of Sciences, 109(31), 12302-12308.
70. Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., ... & Zaks, D. P. M. (2011). Solutions for a cultivated planet. Nature, 478(7369), 337-342.