**Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control**
**Journal homepage: http://kinetik.umm.ac.id**
**ISSN: 2503-2267**
**Vol. 11, No. 1, February, Pp. 1-8**


# Enhancing plant recommendation through IoT-integrated LLM systems

**Panji Maulana Al Latief*[1], Cutifa Safitri[1]**
Faculty of Computer Science, President University, Indonesia[1]

## Article Info

**Keywords:**
LLaMA, Large Language Model, Plant Recommendation, Prompt Engineering, TALLRec

**Article history:**
Received: February 02, 2025
Accepted: December 25, 2025
Published: February 01, 2026

**Cite:**
P. Maulana and Cutifa Safitri, "Enhancing Plant Recommendation through IoT-integrated LLM Systems ", *KINETIK*, vol. 11, no. 1, Feb. 2026.
https://doi.org/10.22219/kinetik.v11i1.2241

*Corresponding author.
Panji Maulana Al Latief
E-mail address:
panji.latief@student.president.ac.id

## Abstract

*Over the past decade, artificial intelligence has experienced phenomenally rapid and extensive expansion across a wide range of industries. Alongside these developments, the agricultural sector stands to benefit significantly from the integration of technology. A significant challenge encountered by farmers is selecting the appropriate crop to plant. The selection of crops is influenced by various factors. Despite advancements in agricultural technology, a considerable gap remains in the integration of IoT with large language models (LLM) for delivering context-specific and data-driven plant recommendation. This study evaluates the reliability of plant recommendations produced by Internet of Things (IoT) devices utilizing the Llama 3.2 model. The model leverages real-time environmental data, including soil pH, altitude, and temperature, to recommend appropriate plant. The recommendations from the base model and a fine-tuned model were compared using precision, recall and F1-score metrics, and were further assessed against established agricultural literature on plant compatibility and growth requirements through human evaluation. The results show substantial performance improvements. The proposed approach achieved an AUC value 59% higher than that of the base model. Precision increased by 40%, recall improved by 105%, and the F1 score rose by 80% compared to the base model.*

## 1. Introduction

Over the past decade, artificial intelligence (AI) has experienced rapid and extensive expansion across various industries. This growth is predicted to significantly impact sectors such as agriculture, where AI can enhance productivity, quality, and efficiency [1]. According to the findings of a study conducted by Ahmad Ali and colleagues, the application of artificial intelligence (AI) in the agricultural sector is highly feasible, which would lead to improvements in terms of productivity, quality, and quantity of products [2].

Farmers face significant challenges in selecting appropriate crops due to factors like temperature, soil composition, and market demand [3]. Recommender systems have demonstrated their effectiveness in supporting decision-making across diverse domains, including agriculture [4]. For example, Venkata et al. developed a chatbot utilizing natural language processing that interprets IoT sensors and delivers information pertinent to agricultural land, offering plant choices based on seasonal or environmental factors [5]. Similarly, Archana et al. proposed a fertilizer recommendation system using a voting-based ensemble classifier [3]. Despite these advancements, previous Deep Neural Network-based models (e.g., CNN and LSTM) and pre-trained language models (e.g., BERT) face constraints in adequately assimilating textual knowledge regarding users and items, revealing deficiencies in natural language comprehension, which results in weak predictive performance across diverse recommendation contexts [6].

The implementation of AI and IoT technology is being explored by farmers to optimize limited land, enhance plant quality, mitigate crop failure, and reduce labor requirements. These technologies can also be used for real-time climate forecasting and hydroponics integration, allowing farmers to monitor and irrigate remotely [7], [8], [9].

Artificial intelligence technology further assists farmers in making informed decisions about planting, management, and harvesting, thereby increasing productivity and profitability [10]. Large Language Models (LLMs) provide effective assistance in informing farmers [11]. LLMs can generalize across unseen tasks and domains, allowing them to adapt to new tasks with suitable instructions or limited demonstrations [6].

Several studies have utilized ChatGPT for agricultural text classification [12] and for enhancing agricultural meteorological recommendations [13]. LLMs have also been applied to improve decision-making in agricultural machinery operations [14] and to support plant care through multimodal dialogue systems [15]. Specialized models like AgriBERT [16] and PLLaMa [17] have been developed to address particular challenges in agriculture and plant science. Bao et al. proposed the TALLRec framework, which shows significant improvement for utilizing LLMs to generate recommendations, enhancing AUC scores by more than 60% [18].

*Table 1. Previous Research*

| Feature | Archana et al. [3] | Venkata et al. [5] | Proposed Method |
|---|---|---|---|
| Core Technology | Ensemble Classifier | Chatbot / NLP | LLM (Llama 3.2) |
| Data Integration | Static Data | IoT Sensors | IoT Sensors |
| Recommendation Type | Fertilizer | General Info | Specific Plant Selection |
| Tuning Method | Traditional ML Training | Standard NLP | LoRA + TALLRec |

Based on the analysis of previous studies, this research enhances plant recommendations through the application of IoT for environmental data collection and employs LLM as a recommendation system. This study employs LLaMA 3.2, incorporating insights from previous research. LLaMA 3.2 is fine-tuned through LoRA and the application of prompt engineering to enhance accuracy and minimize hallucinations in the recommendation outcomes. The recommendation results are evaluated against the baseline model using precision, recall, partial match, average similarity metrics, and human evaluation. Based on the explanation presented in this section, the research problem is formulated into the following research question: *How is the accuracy result on LLM after finetune and prompt engineering?*

This study evaluates the reliability of plant recommendations produced by Internet of Things (IoT) devices utilizing the Llama 3.2 model. The system incorporates real-time environmental data, including soil pH, altitude, and temperature, to recommend appropriate crops. The recommendations performance of the base model and the fine-tuned model is compared using precision, recall and F1-score metrics and further assessed in relation to established agricultural literature concerning plant compatibility and growth requirements through human evaluation. This study seeks to determine the effectiveness of an IoT-based approach utilizing Llama 3.2 in supporting farmers' informed crop selection decisions.

Furthermore, this study aims to develop and implement an Internet of Things (IoT) solution integrated with the advanced large language model (LLM) Llama 3.2, providing recommendations for plant selection informed by environmental data. The scope of the study includes: (1) utilizing Llama to analyze environmental data and produce appropriate plant recommendations grounded in existing agricultural knowledge and data regarding plant growth prerequisites, and (2) assessing the model performance by comparing the fine-tuned model with the base model using precision, recall and F1-score metrics.

## 2. Research Method

### 2.1 Literature Review

This section provides a comprehensive review of fundamental concepts and technologies relevant to the development of an IoT-integrated LLM system for plant recommendations. The review encompasses several key areas including generative models, large language models (LLMs), LLaMA, prompt engineering, the TALLRec framework, LoRA fine-tuning techniques, and evaluation metrics. Each of these components plays a crucial role in creating an effective plant recommendation system that integrates real-time environmental data with advanced language models. Understanding these concepts and their interrelations is essential for developing a system that can provide accurate and contextually relevant plant recommendations based on environmental conditions.

#### 2.1.1 Generative Model

Transformers were introduced in 2017 [19], focusing exclusively on the attention mechanism while eliminating the recurrent network. In contrast to RNNs, transformers can be trained in parallel. In transformers, the encoding and decoding components are stacked vertically. This architecture is autoregressive, requiring the consumption of each generated word prior to the creation of a new word [20].

Generative LLMs function as sequence-to-sequence models, receiving input text and striving to complete it autonomously. A crucial component of these completion models is referred to as the context length or context window. The context length denotes the maximum quantity of tokens that the model is capable of processing. A substantial context window enables the entire documents to be included and processed by the LLMs. Due to the autoregressive nature of these models, the effective context length increases as new tokens are generated [21].

#### 2.1.2 Large Language Model

The field of natural language processing (NLP) has evolved from statistical language modeling to neural language modeling, progressing from pre-trained language models (PLMs) to large language models (LLMs). Pre-trained Language Models (PLMs) are developed through self-supervised learning on extensive text corpora to acquire a universal representation applicable to a wide range of Natural Language Processing (NLP) tasks. Through the enhancement of these models, PLMs have surpassed conventional language models. LLMs have developed into sophisticated artificial intelligence systems proficient in analyzing and creating text with coherent communication and generalizing across diverse tasks [21].

#### 2.1.3 LlaMA

In 2023, Facebook Meta launched Llama as a research program. In recent years, Llama has undergone significant enhancement. The most recent variation, Llama 3.2, shows a significant enhancement in artificial intelligence (AI) capabilities. This is the first publicly available model that can compete with leading AI models about advanced proficiency in general knowledge, guidance, mathematics, tool utilization, and multilingual translation [22].

#### 2.1.4 Prompt Engineering

According to *The Art of Asking ChatGPT for High-Quality Answers* by Ibrahim Jhon [23], prompt engineering involves the formulation of prompts or instructions that guide the output of a language model such as ChatGPT. It allows users to control the model's output and generate text tailored to specific needs. Although the model is capable of producing human-like text, it may not consistently provide the desired output without sufficient guidance. In this context, prompt engineering plays a crucial role, as explicit and well-structured instructions can effectively guide the model's output and ensure its relevance.

#### 2.1.5 TALLRec Framework

The TALLRec framework is a specialized tuning methodology designed to align Large Language Models (LLMs) with recommendation tasks by addressing the inherent disparity between general language generation and the specific requirements of recommendation systems. Although LLMs exhibit strong generative capabilities, they often face challenges in ranking tasks that are critical for effective recommendations. TALLRec addresses this limitation by employing a two-stage instruction-tuning process: (1) *Alpaca Tuning*, which enhances generalization across varied tasks, and (2) *Rec-Tuning*, which optimizes the model specifically for recommendation datasets [18]. Within this framework, recommendation data is structured as instruction-input-output pairs, where user interaction histories and target items serve as inputs, and the model predicts a binary preference (e.g., "Yes" or "No"). Previous studies utilizing TALLRec have reported significant performance gains, with AUC improvements of over 60% compared to traditional methods, thereby validating its effectiveness for complex decision-making tasks like crop selection in this study [18].

#### 2.1.6 Low-Rank Adaptation (LoRA)

Fine-tuning massive Large Language Models (LLMs) typically requires substantial computational resources, which can be prohibitive for IoT-integrated applications operating on consumer-grade hardware. Low-Rank Adaptation (LoRA) addresses this challenge by freezing the pre-trained model weights and injecting trainable rank decomposition matrices into each layer of the Transformer architecture [24]. This technique allows for the adaptation of the model to domain-specific tasks, such as agricultural plant recommendation, by updating only a small subset of the parameters. Prior studies indicates that LoRA can reduce the number of trainable parameters by up to 10,000 times while maintaining model quality comparable to full fine-tuning. In this study, LoRA is critical as it enables the deployment of a sophisticated plant recommendation system on edge devices or standard computers without requiring enterprise-grade GPU clusters.

#### 2.1.7 Evaluation Metrics

Precision and recall are widely utilized metrics for assessing performance in diverse information retrieval and pattern recognition applications. Scores are provided in the range of 0 to1 for a set of predicted items against the ground truth [25]. Precision is defined as the ratio of true positives to the total number of predicted positive results. while recall is defined as the ratio of true positive predictions to the total number of actual positive results.

Receiver operating characteristic (ROC) curves are important in structuring classifiers and illustrating their efficacy. They are frequently employed in medical decision-making and have increasingly been applied in machine learning and data mining research in recent years. ROC curves have traditionally been employed in signal detection theory to illustrate the trade-off between hit rates and false alarm rates of classifiers [26].

The Area Under the Curve (AUC) represents the area beneath the ROC curve and serves as a quantitative measure of the model's ability in differentiating between positive and negative classifications. The AUC quantitatively represents the proportion of the area beneath the ROC curve. The AUC is often quantified as a percentage or decimal, represented by a value between 0 and 1 [26].

### 2.2 Methodology

This section outlines the methodology employed to improve agricultural plant recommendations through the integration of large language models (LLMs) fine-tuned with LoRA and using TALLRec frameworks with prompt engineering. The proposed approach highlights the creation of self-consistency prompts to guide the LLM in generating accurate, contextually relevant, and practical recommendations for farmers based on real-time environmental data.

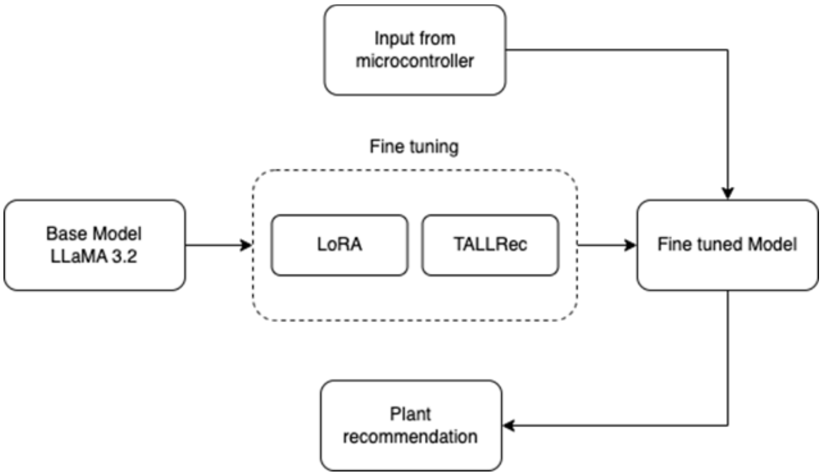

*Figure 1. Proposed Method*

The proposed research approach is necessary for several reasons:
- Consistent Recommendations: Employing the self-consistency prompt technique enables the LLM system to deliver more suitable and precise recommendation by effectively incorporating environmental data.
- Enhanced Decision-Making: Comprehensive and contextually relevant outputs facilitate superior decision-making, resulting in increased yield potential and more effective resource utilization.

In conclusion, the proposed approach leverages the strengths of the LLM to provide timely feedback while addressing the complexity of the recommendations given. By formulating suitable prompts, this study enables LLMs to produce meaningful and contextually relevant insights that can yield substantial advantages for farmers.

#### 2.2.1 Experimental Setup
Experiments were conducted using a MacBook Pro equipped with an Intel Core i5 CPU, 8GB RAM, and Google Colab. IoT devices included a pH sensor, BME280 sensor (for temperature and altitude), and a photoresistor (for light intensity). The base model was LLaMA 3.2, which was compared against a fine-tuned version.

#### 2.2.2 Prompt Design
The proposed technique primarily involves the formulation of precise and contextually detailed instructions that leverage the capabilities of Llama 3.2. These guidelines seek to address several aspects of agricultural decision-making in plant recommendations. An example prompt used to generate recommendations is as follows:
*“Generate a recommendation that is consistent with the following environment information pH 5.88, temperature 30.6°C, altitude 29.87 meters under sea level, humidity 56.75%, and low light, recommend fruit plants that are suitable for a low-maintenance garden.”*
This prompt applies the self-consistency prompting approach as described in *The Art of Asking ChatGPT for High-Quality Answers* by Ibrahim Jhon [23].

#### 2.2.3 Input and Output Specifications
The system receives the following inputs:
- Soil conditions: moisture levels and pH values.
- Environmental condition: altitude level, temperature, humidity, and light intensity.

The results comprise practical recommendations for farmers, including:
- Three plant recommendations.
- Brief descriptive information about each recommended plant.

#### 2.2.4 Data Collection
The dataset was manually collected from the book *Grow Your Own Fruits* [27], containing information on 28 fruit plants, including their pH, temperature, altitude, humidity, and light requirements. The dataset was preprocessed into JSON format and augmented with user-generated scenarios to create training data.

### 2.2.5 Data Preprocessing

The dataset was processed by converting its format from CSV to JSON. Using the TALLRec framework, this dataset was used as training data to fine-tune LLaMA 3.2 with LoRA. The training dataset contains 1388 data, generated and rated by user scenario, with each fruit associated with more than ten scenarios. Each scenario provides a rating for the corresponding fruit, and these ratings are used as scores for the final plant recommendation process.

*Table 2. Plant Dataset*

| Plant | pH | Temperature | Altitude | Humidity | Light |
|---|---|---|---|---|---|
| Grape | 7 | 23-31 | 5-1000 | 75-80 | Full |
| Apple | 5.5-7 | 16-27 | Any | 78-85 | Medium |
| Orange | 6-6.8 | 25-30 | 100-1200 | Any | Medium |
| Melon | 6-6.8 | 12-27 | 200-2000 | 70-80 | Full |
| Soursop | 5.5-6.5 | 22-28 | 0-1000 | 60-80 | Any |

### 2.2.6 Fine Tune

Fine-tuning denotes the process of additional training of pre-trained models using domain-specific data. The purpose of instruction fine-tuning is to train a pre-trained model to respond according to prompts and user input by utilizing a dataset prepared with instructions and corresponding input-output pairs [28].

In this study, fine-tuning was performed using LoRA, with hyperparameters optimized through experimentation. Table 3 summarizes the fine-tuning parameters.

*Table 3. Fine Tune Parameter*

| | R | Alpha | Dropout | Modules | Bias | Use gradient | Learning Rate | Batch size |
|---|---|---|---|---|---|---|---|---|
| Model 1 | 8 | 16 | 0.05 | Q_proj, Y_proj | None | - | $2\times10^{-4}$ | 2 |
| Model 2 | 8 | 16 | 0.05 | All Linear | None | Unsloth | $2\times10^{-4}$ | 2 |
| Model 3 | 16 | 32 | 0.05 | All Linear | None | Unsloth | $2\times10^{-4}$ | 2 |
| Model 4 | 8 | 16 | 0.05 | All Linear | None | Unsloth | $5\times10^{-5}$ | 2 |

The models presented in Table 3 were selected to evaluate specific optimization strategies and perform a hyperparameter sensitivity analysis. Model 1 represents the initial experimental setup, targeting only the query (Q_proj) and value (V_proj) projection modules. Model 2 serves as the optimized baseline, utilizing the *Unsloth* library for efficient gradient checkpointing and expanding the trainable parameters to include all linear modules, thereby enhancing the model's adaptability.

To further investigate the impact of model capacity and training stability, two additional variations were developed. Model 3 explores the effect of increased representational capacity by doubling the LoRA rank (r) to 16 and the alpha parameter to 32. Model 4 tests the impact of a more conservative learning rate to assess convergence stability compared to the standard learning rate. The hyperparameters were selected based on established heuristics and experimental considerations, as follows:

- Rank (r) and Alpha: The rank determines the dimensionality of the low-rank adaptation matrices. A higher rank preserves more information but increases computational demand. The scaling factor *alpha* was set to twice the rank to stabilize feature learning [24].
- Target Modules: While Model 1 targeted specific attention components, Models 2, 3, and 4 targeted all linear layers. This comprehensive approach enables the model to adapt weights across the entire Transformer architecture, potentially capturing more complex patterns in the agricultural dataset.
- Dropout: A consistent dropout rate of 0.05 was applied across all models to prevent overfitting without hindering the learning process.
- Gradient Checkpointing: Models 2, 3, and 4 employed the *Unsloth* framework for gradient checkpointing, which significantly optimizes memory usage during backpropagation compared to the standard approach used in Model 1.
- Batch Size: To ensure a fair comparison of convergence behavior, the batch size was fixed at 2 across all experimental configurations.

### 2.2.7 Evaluation Metrics

To align with standard practices in recommender system evaluation [4], precision is defined as the ratio of true positive results to the total number of predicted results, while recall is defined as the ratio of true positive results to the total number of actual expected results. The evaluation metrics—precision, recall and F1-score—were computed by assigning scores to each result generated by the LLM and comparing them with the ground truth.

ROC-AUC is an excellent statistic for assessing binary classification models. The ROC curve illustrates the model's ability to differentiate between positive and negative classes, whilst the AUC provides a metric that reflects the model's overall performance. As reported by Silveira et al. [4], this tool is particularly useful for addressing imbalanced datasets and for assessing model performance in greater detail across many thresholds.

## 3. Results and Discussion

This section presents the evaluation metrics and comparative analysis of the proposed approach against baseline models. The experiments aim to demonstrate that fine-tuning large language models (LLMs) and prompt engineering can be beneficial for plant recommendations.

### 3.1 Hyperparameter Sensitivity Analysis

In order to investigate the robustness of the proposed model, a sensitivity analysis was conducted with respect to the LoRA hyperparameters. The baseline model was evaluated against two variants: a high-capacity setting and a low-learning-rate. Table 4 shows the settings of hyperparameters.

*Table 4. Sensitivity Analysis Result*

| | R | Alpha | Learning Rate | Final Training Loss | AUC-ROC |
|---|---|---|---|---|---|
| Model 2 | 8 | 16 | $2\times10^{-4}$ | 0,4525 | 0,7600 |
| Model 3 | 16 | 32 | $2\times10^{-4}$ | 0,3602 | 0,8000 |
| Model 4 | 8 | 16 | $5\times10^{-5}$ | 1,0405 | 0,6800 |

### 3.2 Analysis of Training Convergence

In general, the lower training loss means that the model is learning structure from the training data. Figure 2 shows the final training loss for the three experimental settings. Model 3 (High Rank) achieved the lowest loss (0.3602) and demonstrated better convergence behavior than the baseline Model 2, which converged to a minimum of loss of 0.4525. This indicates that the increased number of trainable parameters resulting from the higher LoRa rank (r = 16) enabled the model to learn complex agricultural instructions more effectively. In contrast, Model 4 showed a substantially higher loss (1.0405), likely due to the excessively low learning rate, even though it was intended to prevent overfitting.

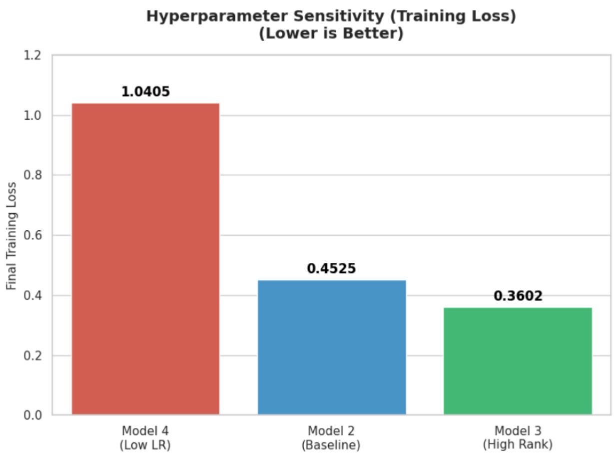


*Figure 2. Final Training Loss*

Model 3 achieved the lowest loss (0.3602), indicating better convergence performance than the baseline Model 2 (loss= 0.4525). This may indicate that the increased number of trainable parameters provided by the higher LoRA rank improved the model’s ability to fit complex agricultural instructions. In contrast, Model 4 exhibited a substantially higher loss (1.0405), indicating that the reduced learning rate hindered convergence within the available training steps. Overall, Model 3 demonstrated the most effective learning performance among all trained configurations.

### 3.3 Performance Evaluation

Figure 3 illustrates the comprehensive AUC-ROC performance, comparing the initial base model and Model 1 from previous iterations with the new experimental configurations.

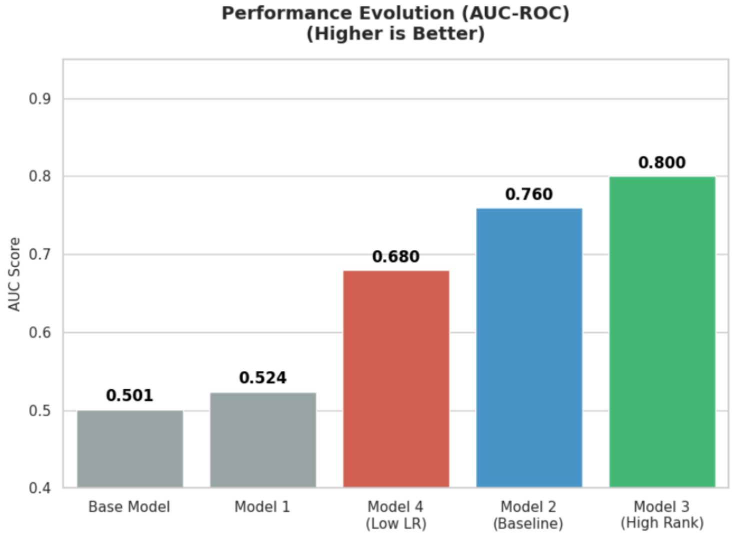


*Figure 3. Final Training Loss*

As shown in Figure 3, increasing the LoRA rank to 16 in Model 3 yielded the best performance across all metrics. Model 3 achieved an AUC-ROC score of 0.80, compared to 0.76 for Model 2 and 0.50 for the original untuned base model. This represents a total improvement of 59.6% over the base model. This confirms that higher-rank adaptations enhance accuracy for complex agricultural datasets by more effectively capturing non-linear relationships among environmental variables.

### 3.4 Detailed Evaluation Metrics

Table 5 presents a detailed comparison of precision, recall, and F1-score for the best-performing model against the base model.

*Table 5. Evaluation Metrics Comparison*

| Metrics | Base | Model 3 |
|---|---|---|
| Precision | 0.1905 | 0.2667 |
| Recall | 0.0675 | 0.1389 |
| F1 Score | 0.0996 | 0.1794 |

According to Table 5, Model 3 demonstrates significant improvements across all key evaluation metrics:

- Precision@3: Improved by 40.00%, indicating that the model is substantially more accurate in its top 3 recommendations.
- Recall@3: Increased by 105.78%, indicating that the model retrieves a significantly larger portion of relevant crops.
- F1-score@3: Improved by 80.12%, reflecting a better balance between precision and recall.

These results indicate that the fine-tuned model excels at prioritizing relevant crops and retrieving a broader range of suitable options, leading to significant gains in performance compared to the base model.

### 3.5 Quantitative Analysis

In terms of the produced output, each model exhibits certain trade-offs. The base model provides more diverse and detailed descriptions, but uses less variation in natural language terms of expression. In contrast, the fine-tuned model generates more structured responses; however, it tends to become redundant at several parts and less expressive in language variation. Additionally, the fine-tuned model often aligns its recommendations with Southeast Asian plant species indicating a focus during its fine-tuning stage.

This work exhibits promising in combining IoT technology and fine-tuned LLMs for crop recommendations. The proposed method outperforms the state-of-the-art results in AUC, Precision, Recall and F1-score. Nevertheless, there are also points to improve. The size of the dataset is too small to generalize the model. Further optimization of model parameters could also improve performance. Future plans include increasing the size of dataset, including more environmental variables, and investigating advanced fine-tuning approaches.

## 4. Conclusion

This study shows a remarkable improvement in crop recommendation accuracy using IoT-based Large Language Model framework. By integrating TALLRec framework and firm hyperparameter optimization, the fine-tuned LLM (Model 3) achieved a ROC-AUC score of 0.80, corresponding to an impressive 60% improvement over the base model. The performance of the proposed model also improved significantly in comparison with all its key metrics, with a 40% increase in Precision, a 105% increase in Recall, and an 80% improvement in F1-score compared to the baseline. These results were validated through extensive comparison with existing agricultural literature. Furthermore, the sensitivity analysis showed that the R parameter model obtained by setting the LoRA rank to 16 achieved the best trade-off between model complexity and performance.

## References


[1] A. Usardi and B. Drut, "Will Artificial Intelligence increase economic growth? Economy & Markets," Jan. 2024.

[2] A. A. AlZubi and K. Galyna, "Artificial Intelligence and Internet of Things for Sustainable Farming and Smart Agriculture," *IEEE Access*, vol. 11, pp. 78686–78692, 2023. https://doi.org/10.1109/ACCESS.2023.3298215

[3] K. Archana and Dr.K.G.Saranya, "Crop Yield Prediction, Forecasting and Fertilizer Recommendation using Voting Based Ensemble Classifier," *SSRG International Journal of Computer Science and Engineering*, vol. 7, no. 5, pp. 1-4, 2020. https://doi.org/10.14445/23488387/IJCSE-V7I5P101

[4] T. Silveira, M. Zhang, X. Lin, Y. Liu, and S. Ma, "How good your recommender system is? A survey on evaluations in recommendation," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 5, pp. 813–831, 2019. https://doi.org/10.1007/s13042-017-0762-9

[5] P. V. Reddy, K. N. Prasad, and C. Puttamadappa, "Farmer's friend: Conversational AI bot for smart agriculture," *Journal of Positive School Psychology*, vol. 6, no. 2, pp. 2541–2549, 2022.

[6] Z. Zhao et al., "Recommender Systems in the Era of Large Language Models (LLMs)," *IEEE Transactions on Knowledge and Data Engineering*, vol. 36, no. 11, pp. 6889-6907, Nov. 2024. https://doi.org/10.48550/arXiv.2307.02046

[7] R. Gunawan, T. Andhika, S. Sandi, and F. Hibatulloh, "Monitoring System for Soil Moisture, Temperature, pH and Automatic Watering of Tomato Plants Based on Internet of Things," *Telekontran : Jurnal Ilmiah Telekomunikasi, Kendali dan Elektronika Terapan*, vol. 7, no. 1, pp. 66–78, Apr. 2019. https://doi.org/10.34010/telekontran.v7i1.1640

[8] H. Husdi and Y. Lasena, "Real Time Analisys Berbasis Internet Of Things Untuk Prediksi Iklim Lahan Pertanian," *Jurnal Media Informatika Budidarma*, vol. 4, no. 3, pp. 834-840, 2020.

[9] Y. Setiawan, H. Tanudjaja, and S. Octaviani, "Penggunaan Internet of Things (IoT) untuk Pemantauan dan Pengendalian Sistem Hidroponik," *TESLA*, vol. 20, no. 2, pp. 175–182, Feb. 2019. https://doi.org/10.24912/tesla.v20i2.2994

[10] J. Li et al., "Large Language Models and Foundation Models in Smart Agriculture: Basics, Opportunities, and Challenges," *Computers and Electronics in Agriculture*, vol. 222, no. 3, pp. 1-18, Jul. 2024. https://doi.org/10.1016/j.compag.2024.109032

[11] M. T. Kuska, M. Wahabzada, and S. Paulus, "AI for crop production – Where can large language models (LLMs) provide substantial value?," *Comput Electron Agric*, vol. 221, p. 108924, Jun. 2024. https://doi.org/10.1016/J.COMPAG.2024.108924

[12] B. Zhao, W. Jin, J. Del Ser, and G. Yang, "ChatAgri: Exploring potentials of ChatGPT on cross-linguistic agricultural text classification," *Neurocomputing*, vol. 557, p. 126708, Nov. 2023. https://doi.org/10.1016/J.NEUCOM.2023.126708

[13] J. Park and S. Choi, "LLMs for Enhanced Agricultural Meteorological Recommendations," *arXiv preprint arXiv:2408.04640*, Jul. 2024. https://doi.org/10.48550/arXiv.2408.04640

[14] E. Johnson and N. Wilson, "Enhancing Agricultural Machinery Management through Advanced LLM Integration," *arXiv preprint arXiv:2407.20588*, Jul, 2024. https://doi.org/10.48550/arXiv.2407.20588

[15] M. Tomar, A. Tiwari, T. Saha, P. Jha, and S. Saha, "An EcoSage Assistant: Towards Building A Multimodal Plant Care Dialogue Assistant," In *European Conference on Information Retrieval*, pp. 318-332, 2024. https://doi.org/10.1007/978-3-031-56060-6_21

[16] S. Rezayi et al., "AgriBERT: Knowledge-Infused Agricultural Language Models for Matching Food and Nutrition," *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*, pp. 5150–5156, Jul. 2022. https://doi.org/10.24963/ijcai.2022/715

[17] X. Yang, J. Gao, W. Xue, and E. Alexandersson, "PLLaMa: An Open-source Large Language Model for Plant Science," *arXiv preprint arXiv:2401.01600*, Jan. 2024. https://doi.org/10.48550/arXiv.2401.01600

[18] K. Bao, J. Zhang, Y. Zhang, W. Wang, F. Feng, and X. He, "TALLRec: An Effective and Efficient Tuning Framework to Align Large Language Model with Recommendation," in *Proceedings of the 17th ACM Conference on Recommender Systems*, RecSys 2023, Association for Computing Machinery, Inc, Sep. 2023, pp. 1007–1014. https://doi.org/10.1145/3604915.3608857

[19] A. Vaswani et al., "Attention Is All You Need," *Advances in Neural Information Processing Systems*, Jun. 2017. https://doi.org/10.48550/arXiv.1706.03762

[20] J. Alammar and M. Grootendorst, "Hands-On Large Language Models: Language Understanding and Generation," O'Reilly Media, Inc., 2024.

[21] H. Naveed et al., "A Comprehensive Overview of Large Language Models," *arXiv preprint arXiv:2307.06435*, 2024. https://doi.org/10.48550/arXiv.2307.06435

[22] L. Team and A. @ Meta, "The Llama 3 Herd of Models," *arXiv preprint arXiv:2407.21783*, 2024. https://doi.org/10.48550/arXiv.2407.21783

[23] I. John, "The Art of Asking ChatGPT for High-Quality Answers A Complete Guide to Prompt Engineering Techniques," *Nzunda Technologies Ltd.,* 2023.

[24] E. J. Hu et al., "LoRA: Low-Rank Adaptation of Large Language Models," *arXiv preprint arXiv:2106.09685*, Jun. 2021. https://doi.org/10.48550/arXiv.2106.09685

[25] P. Fränti and R. Mariescu-Istodor, "Soft precision and recall," *Pattern Recognit Lett*, vol. 167, pp. 115–121, Mar. 2023. https://doi.org/10.1016/j.patrec.2023.02.005

[26] A. P. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms," *Pattern Recognit*, vol. 30, no. 7, pp. 1145–1159, 1997. https://doi.org/10.1016/S0031-3203(96)00142-2

[27] C. Saparinto and R. Susiana, "Grow Your Own Fruits, Panduan Praktis Menanam 28 Tanaman Buah Populer di Pekarangan" *Penerbit Andi*, 2024. ISBN: 978-979-29-5946-8.

[28] V. C. Tjokro and S. A. Sanjaya, "Methods and Applications of Fine-Tuning Llama-2 and Llama-Based Models: A Systematic Literature Analysis," *Journal of System and Management Sciences*, vol. 14, no. 10, pp. 254-266, Jun. 2024. https://doi.org/10.33168/jsms.2024.1015