Precision Farming in Nepal: A Machine Learning Perspective

Anam Giri[1*]; Rabin Sapkota[2]; Raunak Shrestha[3]; Prathama Shrestha[4]; Darshan Paudel[5]; Ashwina Pokharel[6]

1301 Aquarena Springs, Dr Apt 16307A, Texas, United States
*Corresponding Author

Received:- 15 December 2024/ Revised:- 23 December 2024/ Accepted:- 27 December 2024/ Published: 31-12-2024

Abstract— *This paper encompasses three different machine learning models that we built to help Nepali farmers in selecting ideal crops for their land, using the right fertilizers, and predicting plant diseases. We tried about five models each for crop recommendation and fertilizer recommendation and a single model for plant disease prediction. We chose "Decision Trees" for both our Crop Recommendation and Fertilizer Recommendation and "Convolutional Neural Networks (CNN)" for Plant Disease Prediction. All models achieved over 95% accuracy. Our GitHub repository houses all the code, making it accessible for future researchers and ML developers working on related tasks.*

(https://github.com/anamgiri/uunchai).

Keywords— Machine learning, Algorithms, Nepal, Agriculture, Plant disease, fertilizers, crop, recommendation, Plant Disease Prediction Nepal, Decision Tree, Random Forest, Convolutional Neural Networks (CNN), Deep Learning.

I. INTRODUCTION

Most of the foodstuffs in Nepal are still imported from foreign countries, despite the fact that agriculture is the primary occupation for most people. While technology has significantly advanced in other sectors in Nepal, the agricultural sector still relies on traditional farming methods, which are more time-consuming and less productive. Farmers are uneducated about various modern farming practices that could be very beneficial to them. Therefore, we developed machine learning models to assist Nepali farmers in integrating technology into their farming methods. Our focus is solely on the agricultural sector, including both professional farmers and individuals growing crops at home. Through this paper, we aim to demonstrate how our model can identify the right crops to plant under optimal conditions, recommend appropriate fertilizers, and accurately predict plant diseases in a timely manner, which will greatly assist in precision farming for Nepali farmers.

II. LITERATURE REVIEW:

In Hema MS, Niteesha Sharma, and Ch. Santoshini's research paper on "Plant Disease Prediction Using Convolutional Neural Networks," they emphasized that many people pursue agriculture in India, but many crops die due to unidentified diseases. Furthermore, employing professionals costs a lot for small and medium-scale farms, so using machine learning and deep learning models is a more viable alternative for this sector.

In Mahendra N's research on "Crop Prediction Using Machine Learning Approaches," he explains how to design a model that recommends crops and fertilizers based on soil to solve soil and fertilizer-related problems. He used Decision Tree algorithms for crop prediction and SVG for rainfall prediction.

Musanase, Vodacek, Hanyurwimfura, Uwitonze, and Kabandana, in their paper, explain their crop recommendation system that utilizes different machine learning models, where Random Forest performed the best. Soil nutrient parameters (N, P, K) were used to predict the ideal crop for soil. The research provides insight into the benefits of using ML for crop recommendations, showcasing a critical development in precision agriculture.

Lili Li, Shujuan Zhang, and Bing Wang developed an image recognition disease model using deep learning models. This model overviewed the diseases through images and facilitated farmers in using the right fertilizers. Various imaging techniques like Support Vector Machines (SVM), K-means clustering, and K-nearest neighbors (KNN) were employed. A comparison of deep learning models was conducted, and accurate plant leaf disease recognition was achieved using deep learning.

In the research paper "Crop Recommendation System to Maximize Crop Yield Using Machine Learning Technique" by Rohit, Ankit, Mitalee, Pooja, Suresh, and Avinash, they explained their system for recommending crops based on soil conditions to enhance productivity. They addressed the need for precision agriculture, particularly in small, rain-fed farms, by utilizing machine learning models like SVM, ANN, Random Forest, and Naïve Bayes.

In the research paper "Crop Prediction and Fertilizer Recommendation Using Machine Learning" by Prof. Kiran, Priyanka, Pooja, Tushar, and Mayuri, their study focuses on using machine learning, specifically the Support Vector Machine (SVM) algorithm, to predict crop yields and recommend fertilizers based on soil and environmental data. The research involves collecting and preprocessing a dataset with parameters like calcium, magnesium, potassium, and nitrogen, crucial for assessing crop suitability and yield potential. By employing SVM, the study creates a model integrated into a web application where farmers can input soil details to receive tailored crop and fertilizer suggestions. The results indicate high accuracy in predictions, showcasing the significant impact of data-driven insights on improving agricultural productivity and practices. Overall, the study demonstrates the transformative potential of machine learning in agriculture, enhancing decision-making and crop management. It shows that integrating technology with agriculture can significantly enhance crop yield predictions and guide better decision-making for farmers.

III. MATERIALS AND METHODS

3.1 Datasets:

We used secondary sources of data for our research and modeling purposes. We used datasets publicly available on Kaggle for this purpose. Due to the lack of a dataset specifically focused on Nepal, we chose datasets from other countries that contain parameters closely resembling those of Nepal. The fertilizer recommendation dataset contained 100 rows and 8 features (temperature, humidity, moisture, soil type, crop type, nitrogen, potassium, phosphorus) with the fertilizer name as the label. The crop recommendation dataset contained 2,200 rows with 7 features (nitrogen, phosphorus, potassium, temperature, pH, rainfall) and plant name as the label. The plant disease dataset consists of about 87K RGB images of healthy and diseased crop leaves categorized into 38 different classes. For the fertilizer recommendation and crop recommendation systems, we used a 60/20/20 ratio for train, validation, and test sets, whereas for disease prediction, we used an 80/20 ratio for train and validation sets and a separate test images directory.

3.2 Proposed Models:

For the Crop recommendation system: We tried four different models for this system. Decision tree and Random forest both gave us pretty good results. Results for Decision tree were: Test Accuracy: 0.95, Test Precision: 0.96, Test Recall: 0.95, Test F1 Score: 0.95 Results for Random Forest were: Test Precision: 0.98, Test Recall: 0.97, Test F1 Score: 0.97 We also tried SVC and gradient boosting but they couldn't show better results than the previous two because we believe our data also contained non-linear relationships, so SVC couldn't perform as well as other models. Gradient boosting requires more hyperparameters and is sensitive to overfitting, so decision trees and random forest performed slightly better than these models for our data. We finalized the decision tree as our model as it had a slightly better classification report than random forest.

For the Fertilizer recommendation System: We tried four different models for both of these systems. Decision tree and Random forest both gave us pretty good results. Results for decision tree were: Test Accuracy: 1.0, Test Precision: 1.0, Test Recall: 1.0, Test F1 Score: 1.0 Results for Random Forest were: Test Accuracy: 0.95, Test Precision: 1.0, Test Recall: 0.95, Test F1 Score: 0.97 We also tried SVC and Naive Bayes but they couldn't show better results than the previous two because we believe our data also contained non-linear relationships, so SVC couldn't perform as well as other models. Naives Bayes, although with high accuracy, usually assumes data to be independent which is untrue in our case. So, we finalized the decision tree as our model as it had a slightly better classification report than random forest.

For the Plant Disease Prediction System: We used CNN for disease classification as it uses filters and multilayers to detect patterns and edges which is very convenient for image processing. Also, the max pooling layer reduces the spatial dimensions of the image which helps the model become invariant to small translations and distortions.

3.3 Training Process:

3.3.1 Fertilizer Recommendation and Crop Recommendation:

We used ordinal encoder for encoding in all 4 models. The decision tree and random forest models encountered overfitting problems. So, we observed the accuracy for their depths and selected ideal depths for solving the problem.

3.3.2 For the Plant Disease Prediction System:

We used multiple layers for CNN. The first layers were Conv2D, Conv2D, MaxPool2D. These layers were repeated 5 times with different filter sizes (32, 64, 128, 256, 512). We used multiple Conv2D layers with increasing filter sizes to allow the network to learn features at various levels of abstraction. The initial layers with smaller filter sizes (e.g., 32, 64) detect simple, low-level features like edges and textures. As you move deeper into the network, the larger filter sizes (e.g., 128, 256, 512) help the network capture more complex, high-level patterns and structures, such as shapes and objects relevant to plant diseases. The activation function used was 'relu'. Then the layers were Dropout, Flatten, Dense and Dropout respectively. Then, the final layer was Dense with activation as softmax. The optimizer we used was Adam. We did 10 epochs on the training set.

IV. CONCLUSION AND FUTURE WORK:

In summary, we successfully developed three ML models to assist Nepali farmers in choosing the right crops, selecting appropriate fertilizers, and predicting plant diseases. For the Plant Disease Prediction System, our future work includes using different CNN architectures with early stopping to address potential overfitting. Moving forward, we plan to utilize more complex and pre-trained models to analyze their accuracies. For the Crop and Fertilizer Recommendation Systems, our future efforts will focus on finding datasets with more parameters. We aim to collect larger and updated datasets specific to Nepal, which could revolutionize precision agriculture with technology. Additionally, we plan to integrate these three systems into our website. Currently, the models are only hyperlinked to our site, but in the future, we intend to develop an interactive UI where users can input their parameters and receive recommendations or predictions directly.

V. DISCUSSION

5.1 Potential Impact:

- **Increased Productivity:** By optimizing crop selection, fertilizer use, and disease management, these models can significantly boost agricultural yields in Nepal.
- **Reduced Costs:** Precision farming minimizes resource wastage (fertilizers, pesticides) and reduces labor costs by automating certain tasks.
- **Improved Food Security:** Increased productivity can contribute to greater food security for Nepal, potentially reducing reliance on imports.
- **Sustainability:** Optimized resource use can lead to more sustainable agricultural practices, minimizing environmental impact.

5.2 Technological Advancement:

- The successful implementation of these models demonstrates the potential of machine learning in modernizing Nepalese agriculture.
- This can encourage further research and development in this area, leading to more sophisticated and impactful solutions.

5.3 Farmer Empowerment:

By providing farmers with data-driven insights and decision-making tools, these models can empower them to make informed choices and improve their livelihoods.

VI. LIMITATIONS

6.1 Data Limitations:

- **Data Availability:** Relying on datasets from other countries can introduce biases and limit the model's accuracy in the specific context of Nepal.
- **Data Quality:** The quality of available data significantly impacts model performance. Inaccurate or incomplete data can lead to unreliable predictions.
- **Data Collection:** Collecting high-quality, real-time data from Nepalese farms can be challenging due to limited infrastructure and resources.

6.2 Model Limitations:

- **Generalization:** Models trained on limited datasets may not generalize well to new, unseen situations or variations in environmental conditions.
- **Interpretability:** Some complex models, like deep neural networks, can be difficult to interpret, making it challenging to understand the rationale behind their predictions.
- **Maintenance:** Machine learning models require ongoing maintenance, including retraining with new data and adapting to changing conditions.

6.3 Implementation Challenges:

- **Technology Access:** Ensuring access to technology (smartphones, internet connectivity) for all farmers in Nepal can be a significant hurdle.
- **Digital Literacy:** Farmers may require training and support to effectively use and understand the outputs of these models.
- **Trust and Adoption:** Building trust among farmers in the use of technology and convincing them to adopt new practices can be challenging.

6.4 Addressing Limitations:

- **Data Collection:** Invest in initiatives to collect high-quality, location-specific data on soil, weather, and crop conditions in Nepal.
- **Model Development:** Explore more robust and interpretable models, such as explainable AI techniques.
- **Technology Access:** Improve digital infrastructure in rural areas and provide affordable access to smartphones and internet connectivity.
- **Farmer Education:** Conduct workshops and training programs to educate farmers on the use of these technologies and their benefits.
- **Continuous Improvement:** Regularly monitor model performance, gather feedback from farmers, and continuously refine models based on real-world experience.

REFERENCES

[1] Hema, M. S., Sharma, N., Sowjanya, Y., Santoshini, C., Sri Durga, R., & Akhila, V. (2021). Plant disease prediction using convolutional neural network. EMITTER International Journal of Engineering Technology, 9(2), 283–293. https://doi.org/10.24003/emitter.v9i2.640

[2] Emitter Mahendra, N., Vishwakarma, D., Nischitha, K., Ashwini, & Manjuraju, M. R. (2020). Crop prediction using machine learning approaches. International Journal of Engineering Research & Technology, 9(8). https://doi.org/10.17577/IJERTV9IS080029

[3] IJERT Rohit, A., Ankit, M., Mitalee, P., Pooja, S., Suresh, A., & Avinash, K. (2017). Crop recommendation system to maximize crop yield using machine learning technique. International Research Journal of Engineering and Technology, 4(12), 1036–1040. Retrieved from https://www.irjet.net/archives/V4/i12/IRJET-V4I12179.pdf

[4] IRJET Lohar, T. S., & Prof. Kiran, S. (2023). Crop prediction and fertilizer recommendation using machine learning. International Journal of Engineering Research and Applications, 13(3), 28–32. Retrieved from https://www.ijera.com/papers/vol13no3/D13032832.pdf Ijera

[5] Musanase, C., Vodacek, A., Hanyurwimfura, D., Uwitonze, N., & Kabandana, I. (2023). Data-driven analysis and machine learning-based crop and fertilizer recommendation system for revolutionizing farming practices. Agriculture, 13(11), 2141. Retrieved from https://www.mdpi.com/2077-0472/13/11/2141 Li, L., Zhang, S., & Wang, B. (2021).

[6] Plant disease detection and classification by deep learning—A review. IEEE Access, 9, 56683–56698. https://doi.org/10.1109/ACCESS.2021.3069646.