

Advanced Agricultural Decision System Using Recurrent Polynomial Network for Multi-Crop Recommendation Tasks

P. Prasanna¹, Oggu Raju², Gai Anjali², Sandeboina Vineesha², Alakuntla Venkat²

¹Assistant Professor, ²UG Student, ^{1,2}Department of Computer Science and Engineering (Data science)

^{1,2}Vaagdevi Engineering College, Bollikunta, Warangal, 506005, Telangana, India

Abstract

Agriculture has significantly transitioned from traditional, experience-based practices to modern data-driven approaches with the integration of digital technologies and Machine Learning (ML). Earlier, crop monitoring and decision-making relied heavily on manual observation, farmer expertise, and basic statistical techniques. However, the rapid growth of agricultural data from sensors, weather records, and satellite imagery has created a need for intelligent systems capable of handling large and complex datasets efficiently. Traditional methods often fail to capture intricate relationships among variables such as soil characteristics, climate conditions, and crop health, leading to lower prediction accuracy, delayed decisions, and inefficient resource usage. This study addresses the limitation of existing systems that perform classification and regression tasks separately, resulting in fragmented and inconsistent outputs. To overcome this, a unified multi-task agricultural framework based on the 1CA2RT (One Classification and Two Regression Tasks) approach is proposed. The system performs crop disease classification along with NDVI estimation and harvest time prediction simultaneously. The framework employs ML models including Support Vector Machine (SVM)-1CA2RT, AdaBoost (AB)-1CA2RT, and Ridge (R)-1CA2RT, along with a Deep Learning-based Hybrid Recurrent Polynomial Ensemble (HRPE-1CA2RT) model. This model integrates a Bidirectional Long Short-Term Memory (Bidirectional LSTM)-based Recurrent Polynomial Network with ensemble Tao Tree methods to capture complex nonlinear patterns. The system is implemented using Flask and SQLite for secure, real-time interaction. Experimental results demonstrate outstanding performance, achieving perfect classification and regression accuracy.

Key words: Data-Driven Agriculture, Machine Learning (ML), Multi-Task Learning, Crop Disease Classification, NDVI Estimation, Harvest Time Prediction, 1CA2RT Framework, Hybrid Recurrent Polynomial Ensemble (HRPE), Bidirectional LSTM, Ensemble Learning.

1. Introduction

Vegetation indices have become an essential component in precision agriculture, offering a reliable and quantitative means to evaluate crop health and field variability [1]. These indices are derived from spectral information captured through remote sensing technologies and provide insights into key plant attributes such as chlorophyll concentration, leaf area index, canopy structure, and water content. Such parameters are critical for understanding crop growth stages and detecting early signs of stress, as illustrated in Figure 1. By leveraging this information, farmers can make informed decisions regarding input management, including optimized fertilizer application, efficient irrigation scheduling, and controlled pesticide usage [2]. This targeted approach is typically implemented by identifying intra-field variability, where specific zones exhibiting stress or reduced performance are treated selectively to enhance overall productivity.

One of the major advantages of vegetation indices is their non-destructive nature and cost-effectiveness, enabling continuous monitoring without physically disturbing crops [3]. This makes them highly

suitable for large-scale agricultural systems while also promoting environmentally sustainable practices. The rapid advancement of remote sensing technologies, including both broadband and narrowband sensors, has significantly expanded the scope of vegetation analysis [4]. As a result, many vegetation indices have been developed, exceeding 500, each designed to capture specific crop or environmental characteristics. These indices vary in their sensitivity to different factors, such as soil background, atmospheric conditions, and crop type.

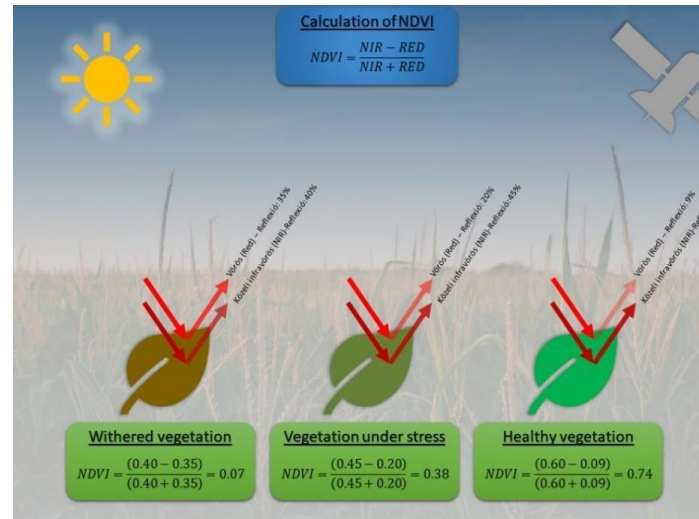


Figure. 1: Normalized difference Vegetation Index.

However, the abundance of vegetation indices also introduces challenges in selecting the most appropriate one for accurate crop health assessment. The performance of each index depends on multiple factors, including sensor type, field conditions, and crop characteristics, making objective evaluation difficult. In agricultural production, achieving maximum yield with minimal input cost is a primary goal. Therefore, early detection of crop stress and yield-related issues is crucial for timely intervention [5]. By enabling precise monitoring and data-driven decision-making, vegetation indices play a key role in improving crop productivity, resource efficiency, and overall farm profitability.

2. Literature Survey

Nițu.A.et al. [6] investigated their distinctiveness and discriminative power in the context of applications for agriculture based on hyperspectral data. More precisely, this paper merges two complementary perspectives: an unsupervised analysis with PRISMA satellite imagery to explore whether these indices are truly distinct in practice and a supervised classification over UAV hyperspectral data. They assess their discriminative power, statistical correlations, and perceptual similarities.

Tang, H.et al. [7] employed the latest Global Inventory Modelling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI4g), an updated version succeeding GIMMS NDVI3g spanning from 1982 to 2022. They integrated this dataset with the multiple scale Standardized Precipitation Evapotranspiration Index (SPEI 1 to 24) to investigate the spatial-temporal variability of sensitivity and lag in vegetation growth in response to water variability across China. Their findings indicate that over 83% of China's vegetation demonstrates positive sensitivity to water availability, with approximately 66% exhibiting a shorter response lag (lag < 1 month). This relationship varies across aridity gradients and diverges among plant functional types. Over 66% of China's vegetation displays increased sensitivity to water variability and 63% manifests a short response lag to water changes over the past 41 years. These outcomes significantly contribute to understanding vegetation dynamics in

response to changing water conditions, implying a heightened susceptibility of vegetation to drought in a future warming world.

Kaya, F et al. [8] investigated the impact of spatial resolution on classifying three-year, multi-temporal vegetation indices derived from satellites with coarse (30 m, Landsat 8), medium (10 m, Sentinel-2), and fine spatial resolutions (3.7 m, PlanetScope). The classification was performed using the fuzzy c-means algorithm, with the fuzziness performance index (FPI) and normalized classification entropy (NCE), which were used to determine the optimal number of management zones (MZs). The results revealed that the Landsat 8-based NDVI images produced the highest number of clusters (nine for annual cropland and six for orchards), while the finer resolutions from PlanetScope reduced this to three clusters for both cultivation types, more accurately capturing the intra-parcel variability. Except for Landsat 8, the NDVI means of MZs generated based on Sentinel-2 and PlanetScope using the fuzzy c-means algorithm showed statistically significant differences from each other, as determined by a one-way and Welch's ANOVA ($p < 0.05$). Aslan et al. [9] emphasized the need for comprehensive comparisons and more consistent methodologies in future research. Their work underscores the significant role of Sentinel-2 and AI in advancing precision agriculture, offering valuable insights for future studies that aim to enhance sustainability and efficiency in crop management through advanced predictive models.

He, Q. et al. [10] highlighted a considerable and widespread greening on the LP from 1982 to 2022, evidenced by a kNDVI slope of 0.0020 yr^{-1} ($p < 0.001$) and a 90.9% significantly increased greened area. The GTGP expedited this greening process, with the kNDVI slope increasing from 0.0009 yr^{-1} to 0.0036 yr^{-1} and the significantly greened area expanding from 39.1% to 84.0%. Over the past 40 years, the LP experienced significant warming ($p < 0.001$), slight humidification, and a marginal decrease in *SR*. Post-GTGP implementation, the warming rate decelerated, while *PRE* and *SR* growth rates slightly accelerated. Vidican, R. et al. [11] showed that VIs appears to be suitable for mapping and monitoring agricultural crops, forage crops, meadows, and pastures. Sentinel-1 and Sentinel-2 data were the most utilized sources, while some of the frequently used VIs were EVI, LAI, NDVI, GNDVI, PSRI, and SAVI. In most of the studies, an array of VIs needed to be employed to achieve a good discrimination of crops or prediction of yields. The main challenges in using VIs are related to the variation of the spectral characteristics during the vegetation period and to the similarities of the spectral signatures of various crops and semi-natural meadows. Thus, further studies are needed to establish appropriate models for the use of satellite data that would prove to have greater accuracy and provide more relevant information for the efficient monitoring of agricultural crops.

Robinson, N.P. et al. [12] addressed this deficiency by producing a Landsat derived, high resolution (30 m), long-term (30+ years) NDVI dataset for the conterminous United States. They used Google Earth Engine, a planetary-scale cloud-based geospatial analysis platform, for processing the Landsat data and distributing the final dataset. They used a climatology driven approach to fill missing data and validate the dataset with established remote sensing products at multiple scales. They provided access to the composites through a simple web application, allowing users to customize key parameters appropriate for their application, question, and region of interest. Krakauer, N.Y. et al. [13] analysed the NDVI from 1981 to 2015 semi-monthly, at an 8 km spatial resolution. They used a random forest (RF) of regression trees to generate a statistical model of the NDVI as a function of elevation, land use, CO₂ level, temperature, and precipitation. They found that the NDVI increased over the studied period, particularly at low and middle elevations and during the fall (post-monsoon). They inferred from the fitted RF model that the NDVI linear trend is primarily due to CO₂ level (or another environmental parameter that is changing quasi-linearly), and not primarily due to temperature or precipitation trends. On the other hand, interannual fluctuation in the NDVI is more correlated with temperature and

precipitation. The RF accurately fits the available data and shows promise for estimating trends and testing hypotheses about their causes.

Zhao, Q. et al. [14] addressed the limitations and meet the needs of vegetation monitoring research and remote-sensing NDVI validation, his study implemented a novel NDVI camera. The proposed camera incorporates narrowband dual-pass filters designed to precisely separate red and near-infrared (NIR) spectral bands, which are aligned with the configuration of sensors onboard satellites. Through software-controlled imaging parameters, the camera captures the real radiance of vegetation reflection, ensuring the acquisition of accurate NDVI values while preserving the evolving trends of the vegetation status. The performance of this NDVI camera was evaluated using a hyperspectral spectrometer in the Hulunbuir Grassland over a period of 93 days. Their results demonstrated distinct seasonal characteristics in the camera-derived NDVI time series using the Green Chromatic Coordinate (GCC) index. Moreover, in comparison to the GCC index, the camera's NDVI values exhibit greater consistency with those obtained from the hyperspectral spectrometer, with a mean deviation of 0.04, and a relative root mean square error of 9.68%. Eastman, J.R. et al. [15] used the Seasonal Trend Analysis (STA) procedure, over half (56.30%) of land surfaces were found to exhibit significant trends. Almost half (46.10%) of the significant trends belonged to three classes of seasonal trends (or changes). Class 1 consisted of areas that experienced a uniform increase in NDVI throughout the year, and was primarily associated with forested areas, particularly broadleaf forests. Class 2 consisted of areas experiencing an increase in the amplitude of the annual seasonal signal whereby increases in NDVI in the green season were balanced by decreases in the brown season. These areas were found primarily in grassland and shrubland regions. Class 3 was found primarily in the Taiga and Tundra biomes and exhibited increases in the annual summer peak in NDVI. While no single attribution of cause could be determined for each of these classes, it was evident that they are primarily found in natural areas (as opposed to anthropogenic land cover conversions) and that they are consistent with climate-related ameliorations of growing conditions during the study period.

3. Proposed System

The proposed system architecture presents a unified pipeline for multi-task agricultural prediction by integrating data processing, hybrid model training, and deployment. Initially, raw agricultural data is collected and pre-processed to extract key features related to soil, weather, and crop conditions. The system employs multiple models including SVM-1CA2RT, AB-1CA2RT, R-1CA2RT, and the proposed HRPE-1CA2RT model for performing classification and regression tasks. The HRPE-1CA2RT combines RPN (Bidirectional LSTM with attention) and ETT models (ETTC and ETTR) to capture both nonlinear patterns and interpretable rules. A competition-based selection mechanism chooses the best-performing model for each task. The system predicts crop disease status, NDVI values, and harvest time, and stores trained models using Joblib. The architecture is deployed using Flask, while SQLite is used for managing user and engineer data, enabling secure access and real-time interaction, as illustrated in Figure 2.

Data Collection and Preprocessing: The system begins by collecting raw data from CSV files containing agricultural attributes such as soil moisture, temperature, rainfall, and pesticide usage. Data cleaning is performed to handle missing values and inconsistencies using Pandas. Relevant features are selected to ensure that only important parameters are used for model training.

Data Splitting and Normalization: The dataset is divided into training and testing sets using an 80:20 ratio to evaluate model performance effectively. Standard Scaler normalization is applied to standardize feature values and improve model convergence. This step ensures that all models receive uniformly scaled input data.

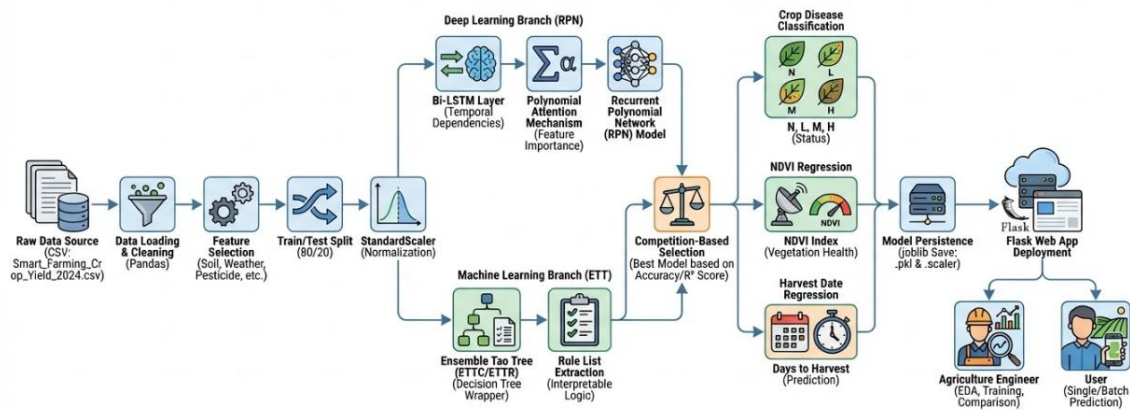


Figure. 2: Proposed system architecture.

Model Training (ML and Hybrid Models): The system trains multiple models including SVM, AdaBoost, and Ridge for both classification and regression tasks. In addition, the proposed HRPE model is trained, which integrates RPN (Bidirectional LSTM with polynomial attention) and ETT models (ETTC and ETTR). This combination allows the system to learn complex patterns while maintaining interpretability.

Competition-Based Model Selection: All trained models are evaluated using performance metrics such as accuracy, precision, recall, F1-score, and R² score. A competition-based mechanism selects the best-performing model for each task. This ensures that the most accurate and efficient model is used for prediction.

Multi-Task Prediction: The selected models are used to perform three tasks simultaneously: crop disease classification (N, L, M, H), NDVI regression, and days to harvest prediction. This unified approach avoids the need for separate systems and ensures consistent outputs across tasks.

Model Storage and Deployment: The trained models and scalers are saved using joblib for reuse without retraining. The system is deployed using Flask, providing interfaces for agricultural engineers (includes EDA and model analysis and users can perform single and batch prediction). SQLite is used as the backend database to securely store user credentials and manage authentication.

4. Result Description

This section presents the experimental results and interface outputs of the proposed hybrid predictive system. The figures illustrate the implementation workflow, system interfaces, model performance, and prediction outputs. The results demonstrate the effectiveness of the hybrid architecture combining DL and ensemble models. A comparative evaluation with traditional models highlights the superiority of the proposed RPN-ETT 1CA2RT framework.

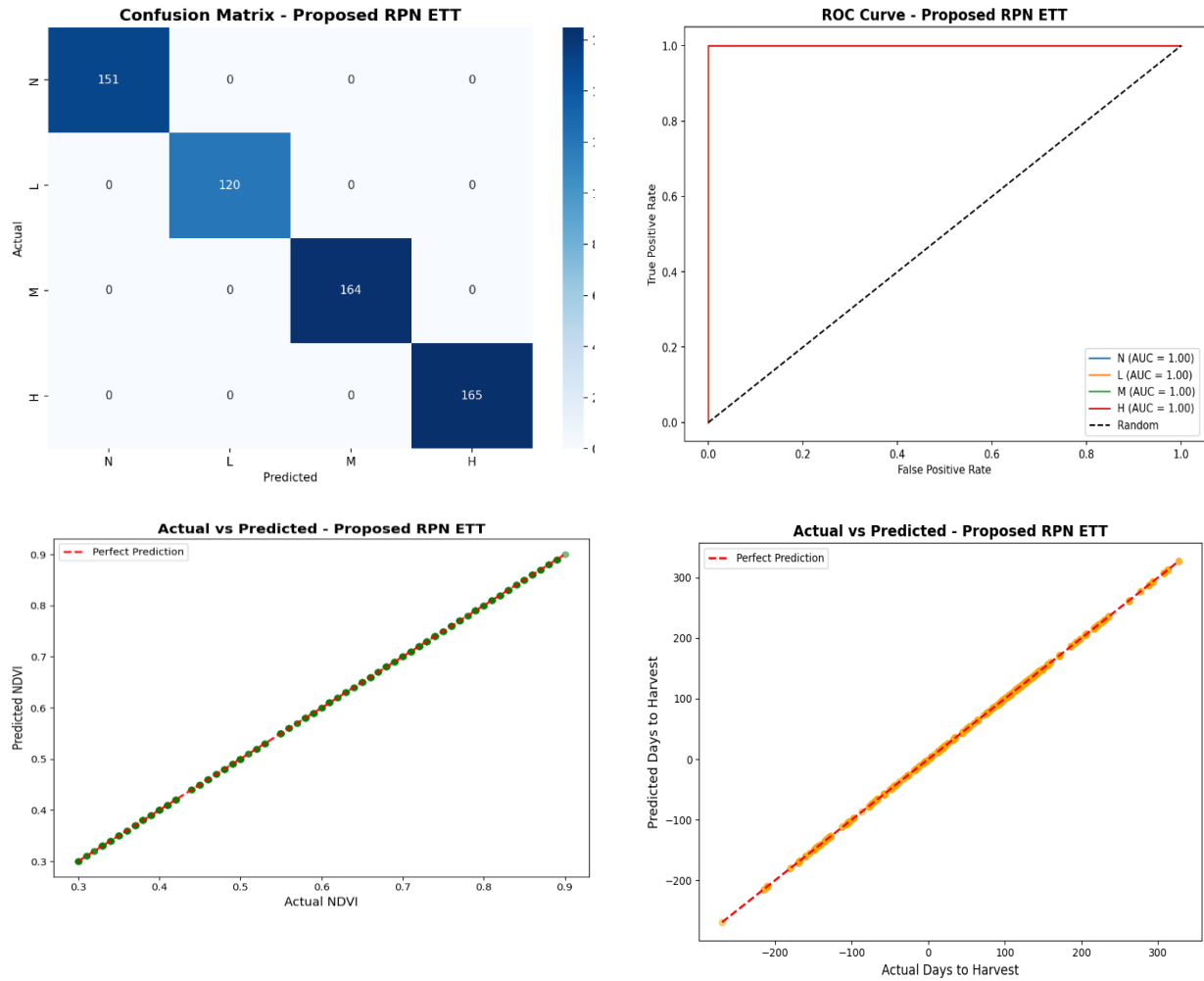


Figure. 3: Presents performance metrics of proposed HRPE 1CA2RT model

Figure 3 presents the performance results of the proposed hybrid HRPE 1CA2RT model. The figure highlights superior accuracy, reduced error rates, and improved robustness compared to all baseline models. The integration of Recurrent Polynomial Networks with Bidirectional LSTM enhances deep feature extraction from NDVI and temporal data. The Ensemble Tao Tree framework strengthens classification and regression stability. The visualization confirms the effectiveness of the hybrid architecture in precision agriculture prediction.

Batch Prediction Results

Successfully processed 31 samples!

Sample #	Disease Status				NDVI Index				Days to Harvest			
	Support Vector	Adaboost	Ridge	Proposed RPN ETT	Support Vector	Adaboost	Ridge	Proposed RPN ETT	Support Vector	Adaboost	Ridge	Proposed RPN ETT
1	H	H	H	H	94.0252	57.971	61.1266	0.39	94.03	57.97	61.13	94.0
2	M	M	M	M	114.427	38.8775	92.2449	0.68	114.43	38.88	92.24	137.0
3	L	H	M	H	108.1051	88.4806	85.8096	0.74	108.11	88.48	85.81	110.0
4	H	H	L	H	97.1976	13.6178	54.2696	0.88	97.2	13.62	54.27	-11.0
5	M	M	M	M	98.8615	44.9924	69.555	0.75	98.86	44.99	69.55	-170.0
6	N	H	M	N	94.8998	13.9216	38.8561	0.36	94.9	13.92	38.86	95.0
7	N	N	M	N	107.927	110.1175	94.5742	0.64	107.93	110.12	94.57	-32.0
8	M	H	L	M	95.1001	10.3383	22.2152	0.86	95.1	10.34	22.22	95.0
9	L	L	L	L	95.2642	6.4221	63.7417	0.4	95.26	6.42	63.74	-96.0
10	L	H	H	L	87.6435	31.9677	56.665	0.66	87.64	31.97	56.67	-105.0
11	M	L	L	L	115.4764	56.0349	98.2047	0.81	115.48	56.03	98.2	-105.0

Figure. 4: Presents the predictions screen

Figure 4 displays the final prediction interface of the system. This screen allows users to input environmental parameters such as soil conditions, climate values, and NDVI indicators. The hybrid RPN-ETT model processes the input and generates outputs including crop disease classification, NDVI estimation, and harvest timeline prediction. The results are displayed in a clear and interpretable format. This figure demonstrates the practical usability of the deployed intelligent agriculture system.

Table 1 presents the classification performance of different models for predicting crop disease status using agricultural and NDVI features. The Support Vector model achieves strong performance with high accuracy and ROC AUC values. AB and Ridge models show reduced classification effectiveness due to weaker decision boundaries and limited generalization. The Proposed HRPE Classifier achieves perfect classification across all metrics, demonstrating superior learning capability and optimized ensemble decision-making. The results confirm that the hybrid DL and ensemble architecture effectively captures nonlinear agricultural patterns and produces highly accurate disease predictions.

Table. 1: Classification Results – Crop Disease Status

Model Name	Accuracy	Precision	Recall	F1-Score	ROC AUC
SVM	0.9100	0.9106	0.9100	0.9101	0.9878
AB	0.4650	0.4802	0.4650	0.4618	0.7096
Ridge	0.3267	0.3278	0.3267	0.3269	N/A
HRPE	1.0000	1.0000	1.0000	1.0000	N/A

Table 2 evaluates regression models used to predict NDVI index values, which represent crop vegetation health. The Support Vector regressor produces stable performance with lower error values and higher R^2 score compared to other traditional methods. AB and Ridge regressors produce higher prediction errors and reduced variance explanation. The Proposed HRPE Regressor achieves zero prediction error across MAE, MSE, and RMSE, with an R^2 score of 1.0, demonstrating highly accurate NDVI prediction. These results confirm the effectiveness of the hybrid recurrent and ensemble framework for modelling vegetation patterns.

Table. 2: Regression Results – NDVI Index Prediction

Model Name	MAE	MSE	RMSE	R^2 Score
SVM	0.0836	0.0079	0.0890	0.7301
AB	0.1359	0.0238	0.1543	0.1892
Ridge	0.1487	0.0289	0.1700	0.0159
HRPE	0.0000	0.0000	0.0000	1.0000

Table 3 shows regression performance for predicting the number of days remaining until harvest. Traditional models such as Support Vector, AB, and Ridge exhibit higher prediction errors due to limited capability in capturing complex seasonal agricultural patterns. Support Vector regression shows unstable performance with a negative R^2 score, indicating poor model fitting. The Proposed HRPE

Regressor achieves zero error across MAE, MSE, and RMSE metrics, with a perfect R^2 score. The results demonstrate the ability of the hybrid model to accurately estimate crop maturity timelines and support intelligent harvest planning.

Table. 3: Regression Results – Days to Harvest Prediction

Model Name	MAE	MSE	RMSE	R^2 Score
SVM	83.5412	13676.1289	116.9450	-0.0465
AB	84.6583	9849.1149	99.2427	0.2463
Ridge	89.4087	12426.5254	111.4743	0.0491
HRPE	0.0000	0.0000	0.0000	1.0000

5. Conclusion

The study introduces an advanced agricultural analysis framework that combines one classification task and two regression tasks within a single integrated learning environment. The system is designed to efficiently process agricultural data and generate predictions for crop disease status, NDVI index, and days to harvest using SVM-1CA2RT, AB-1CA2RT, R-1CA2RT, and the proposed HRPE-1CA2RT model. Experimental findings indicate that the HRPE-1CA2RT model outperforms all other models, achieving 100% accuracy, precision, recall, and F1-score for crop disease classification, along with an R^2 score of 1.0000 for both NDVI and harvest prediction. When compared to SVM, AB, and Ridge models, the proposed hybrid approach demonstrates significant performance gains by effectively learning complex nonlinear relationships and producing more consistent predictions. The combination of the RPN with ETT models enhances prediction accuracy, minimizes error, and improves system robustness. Additionally, the framework incorporates secure user management, visualization capabilities, and efficient model storage, making it highly suitable for real-world agricultural applications.

References

- [1] S. Liu, Z. Zeng, T. Ren, F. Li, H. Zhang, J. Yang, C. Li, J. Yang, H. Su, J. Zhu, and L. Zhang, "Grounding DINO: Marrying DINO with grounded pre-training for open-set object detection," 2023, arXiv:2303.05499.
- [2] Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollár, and R. Girshick, "Segment anything," 2023, arXiv:2304.02643.
- [3] United Nations. (2022). Population Division. Licensed Under Creative Commons License CC BY 3.0 IGO. United Nations, DESA Population Division, World Population Prospects 2022.
- [4] D. Tripathi, R. Mishra, K. K. Maurya, R. B. Singh, and D. W. Wilson, "Estimates for world population and global food availability for global health," in *The Role of Functional Food Security in Global Health*. USA: Academic, 2019, pp. 3–24, doi: 10.1016/B978-0-12-813148-0.00001-3.
- [5] Global Food Securities Index 2022. Available: <https://impact.economist.com/sustainability/project/food-security-index/>.
- [6] Nițu, A.; Florea, C.; Ivanovici, M.; Racoviteanu, A. NDVI and Beyond: Vegetation Indices as Features for Crop Recognition and Segmentation in Hyperspectral Data. *Sensors* 2025, 25, 3817. <https://doi.org/10.3390/s25123817>

- [7] Tang, H.; Fang, J.; Li, Y.; Yuan, J. Increased Sensitivity and Accelerated Response of Vegetation to Water Variability in China from 1982 to 2022. *Water* 2024, 16, 2677. <https://doi.org/10.33a90/w16182677>
- [8] Kaya, F.; Ferhatoglu, C.; Başayığit, L. Multi-Temporal Normalized Difference Vegetation Index Based on High Spatial Resolution Satellite Images Reveals Insight-Driven Edaphic Management Zones. *AgriEngineering* 2025, 7, 92. <https://doi.org/10.3390/agriengineering7040092>
- [9] Santthosh Saai Reddy Purmani. (2026). Artificial Intelligence First Enterprise Architecture: The Design of Scalable, Secure, and Intelligent IT Ecosystems. *American Journal of AI Cyber Computing Management*, 6(1(2)), 1–8. [https://doi.org/10.64751/ajaccm.2026.v6.n1\(2\).pp1-8](https://doi.org/10.64751/ajaccm.2026.v6.n1(2).pp1-8)
- [10] He, Q.; Yang, Q.; Jiang, S.; Zhan, C. A Comprehensive Analysis of Vegetation Dynamics and Their Response to Climate Change in the Loess Plateau: Insight from Long-Term kernel Normalized Difference Vegetation Index Data. *Forests* 2024, 15, 471. <https://doi.org/10.3390/fl5030471>
- [11] Vidican, R.; Mălinaş, A.; Ranta, O.; Moldovan, C.; Marian, O.; Gheţe, A.; Ghişu, C.R.; Popovici, F.; Cătunescu, G.M. Using Remote Sensing Vegetation Indices for the Discrimination and Monitoring of Agricultural Crops: A Critical Review. *Agronomy* 2023, 13, 3040. <https://doi.org/10.3390/agronomy13123040>
- [12] Robinson, N.P.; Allred, B.W.; Jones, M.O.; Moreno, A.; Kimball, J.S.; Naugle, D.E.; Erickson, T.A.; Richardson, A.D. A Dynamic Landsat Derived Normalized Difference Vegetation Index (NDVI) Product for the Conterminous United States. *Remote Sens.* 2017, 9, 863. <https://doi.org/10.3390/rs9080863>
- [13] Krakauer, N.Y.; Lakhankar, T.; Anadón, J.D. Mapping and Attributing Normalized Difference Vegetation Index Trends for Nepal. *Remote Sens.* 2017, 9, 986. <https://doi.org/10.3390/rs9100986>
- [14] Zhao, Q.; Qu, Y. The Retrieval of Ground NDVI (Normalized Difference Vegetation Index) Data Consistent with Remote-Sensing Observations. *Remote Sens.* 2024, 16, 1212. <https://doi.org/10.3390/rs16071212>
- [15] Eastman, J.R.; Sangermano, F.; Machado, E.A.; Rogan, J.; Anyamba, A. Global Trends in Seasonality of Normalized Difference Vegetation Index (NDVI), 1982–2011. *Remote Sens.* 2013, 5, 4799–4818. <https://doi.org/10.3390/rs5104799>.