IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

A Comparative Analysis Of Crop Yield Prediction System Based On Machine Learning And Expert Knowledge Integration

Dr. G. Murali*1, K .Saketh Reddy*2,Y. Charitha Sree*3, P. Nikhitha*4, G. Vyshnavi*5

1*Professor, Department of CAI, KKR & KSR Institute of Technology and Sciences, Guntur,Andhra Pradesh

2*Co-founder, Codegnan , Vijayawada, Andhra Pradesh

3*Students, Department of CSE-Artificial Intelligence, KKR & KSR Institute of Technology and Sciences, Guntur

ABSTRACT:

This article aims to forecast future outcomes using machine learning as a crucial tool for supporting decisions in crop yield prediction, such as choosing crops and managing them throughout their growth. Agricultural planning is crucial for the economic growth and food security of countries relying on agriculture. Despite its significance, the agricultural sector faces numerous challenges. Farmers invest considerable time, money, and effort into farming activities, with crop selection being a particularly daunting task. Choosing the wrong crop can lead to financial hardships for farmers. Given the imperative to sustainably feed the global population, advancements in agricultural production are essential. While past research has explored methods like statistical analysis and data mining for crop and soil classification, predicting the most suitable crop for a given soil remains challenging. Machine learning techniques have emerged as a promising solution to this issue. The proposed work aims to predict the optimal crop and recommend suitable fertilizers by analyzing soil composition and properties using machine learning algorithms. This approach seeks to address existing challenges and improve prediction accuracy, ultimately assisting farmers in selecting the right crops to maximize yields and earnings.

KEYWORDS: Crop Yield Prediction,Machine Learning,Expert Knowledge Integration,Agricultural DataAnalytics,Remote Sensing,Climate Change Adaptation,Precision Agriculture.

1. INTRODUCTION

The fusion of Machine Learning (ML) with specialized knowledge in the realm of crop yield forecasts represents a significant evolution in farming techniques, with the objective of enhancing decision-making capabilities and agricultural productivity. This analytical comparison investigates the performance of these advanced systems relative to traditional crop yield forecasting methods, focusing on their effectiveness, reliability, and their capacity to drive innovation in agricultural planning and sustainability. The intention is to provide a detailed examination of how the synergy between ML algorithms and the comprehensive, practical knowledge of agronomy professionals could lead to enhanced accuracy, resilience, and adaptability in agricultural practices.

The Advent of Machine Learning Models :
Machine learning has brought forward potent tools for addressing the complexities of agriculture. These ML models have the capability to sift through intricate, multi-dimensional data sets, including satellite imagery, soil composition, weather patterns, and the genetic data of crops, to predict results with a level of precision that was previously unachievable. They excel in detecting patterns and relationships that might not be obvious to human observers, offering new insights into the factors that affect crop yields.

Blending Expert Knowledge:
Notwithstanding the advancements in ML, the value of specialized expertise cannot be overstated, particularly in understanding the nuances of local environments,agricultural practices, and environmental
specialized knowledge with ML models involves utilizing in-depth, field-specific expertise to refine, alter, or influence the predictions generated by the algorithms. This blend may involve creating hybrid models that incorporate expert knowledge into rule-based systems alongside ML, or ensemble models that utilize expert opinions in conjunction with machine-generated predictions.

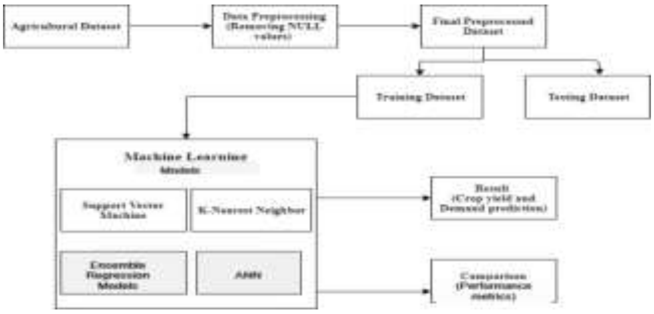
FIG :

II. MATERIAL AND METHODS

The methodology of the "Crop Yield Prediction" project involves a systematic approach to data collection, preprocessing, model training, and prediction. This study introduces a method known as CSM for determining the optimal sequence of crops for planting throughout a season. The CSM approach aims to enhance the overall net yield of seasonal crops. This method makes decisions on crop selection based on anticipated yield rates, which are affected by various factors such as weather conditions, soil type, water availability, and crop species. Data for the crop sowing schedule were collected from farmers in the Patna District of Bihar, India. The method considers the type of crop, its planting time, growth duration, and forecasted yield rate for the season to identify a crop sequence that maximizes daily production throughout the season. Accuracy: The effectiveness and precision of the CSM technique rely on the accuracy of the predicted values for the influencing parameters. Future Work: There is a necessity for employing a more accurate and efficient forecasting technique.

Analysis of Soil Behaviour and Prediction of Crop Yield Authors:

Monali Paul, Santosh K. Vishwakarma, and Ashok Verma Title: Analysis of Soil Behaviour and Prediction of Crop Yield using a Data Mining Approach Published in: IEEE (2015) Methodology: In this research, experiments were conducted using RapidMiner 5.3 software. Two prominent classification algorithms, K-Nearest Neighbor (KNN) and Naive Bayes (NB), were utilized on a soil dataset obtained from a soil testing laboratory in Jabalpur, M.P., India. The aim was to classify soil into low, medium, and high categories to facilitate crop yield predictions using the data available. This investigation aids soil analysts and farmers in making informed decisions about which plots of land could yield better crop production. Accuracy: The accuracy varied between the two classification methods. The KNN algorithm categorized 30 plots as having low-quality soil, 45 as medium, and 25 as high. In contrast, the Naive Bayes algorithm identified 15 plots as low, 40 as medium, and 45 as high-quality soil. Future Work: Due to certain complexities, a small dataset was utilized in this study. Future research could benefit from a larger dataset of 1GB or mored 45 as high-quality soil. Future Work: Due to certain complexities, a small dataset was utilized in this study. Future research could benefit from a larger dataset of 1GB or more.



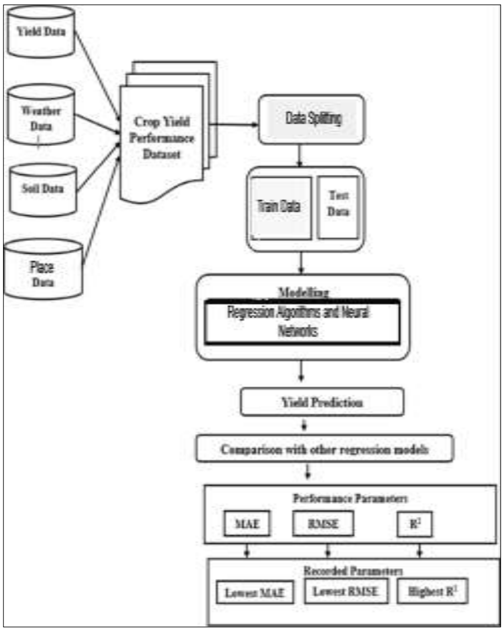


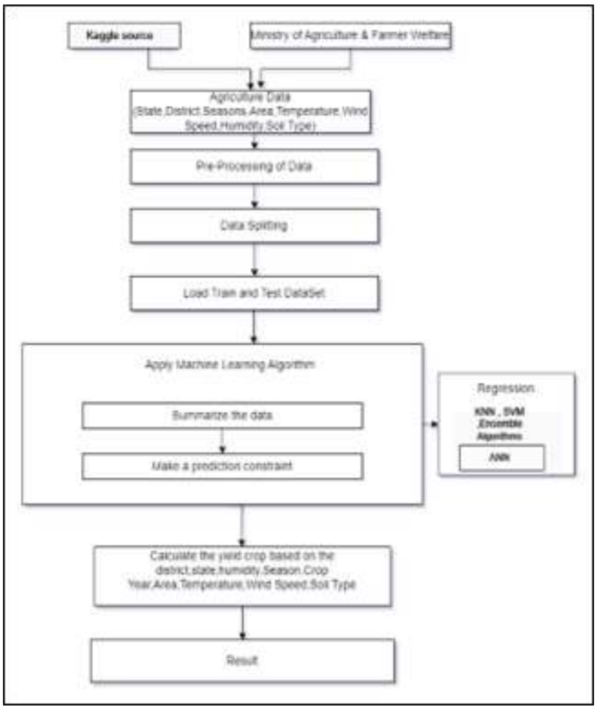
Fig :

DATA FLOW DIAGRAM:

Data Flow Diagram offer a visual depiction of how information circulates within a system, showing the movement of data among processes, data stores, and external entities. Standard symbols represent components like processes, data stores, data flows, and external entities in DFDs. They serve to elucidate the scope and boundaries of a system, as well as the processes involved in transforming data, facilitating system analysis, design, and communication.

Activity/Process Diagram:

Visualizes the sequence of actions or processes involved in each stage of the project, providing a high-level overview of project execution.The process flow diagram delineates the steps for forecasting crop yield by integrating data inputs from various origins. These inputs encompass yield, weather, soil, and location data, amalgamated into a unified dataset. Subsequently, the dataset undergoes division into train and test sets for model training and evaluation purposes.



Machine learning models, encompassing regression algorithms and neural networks, are employed to establish correlations between input variables and crop yield. Predictions are then generated utilizing these models, with their accuracy assessed using performance metrics like MAE, RMSE, and R^2. The identification of the best-performing model is based on documented parameters, ensuring precise crop yield foreca

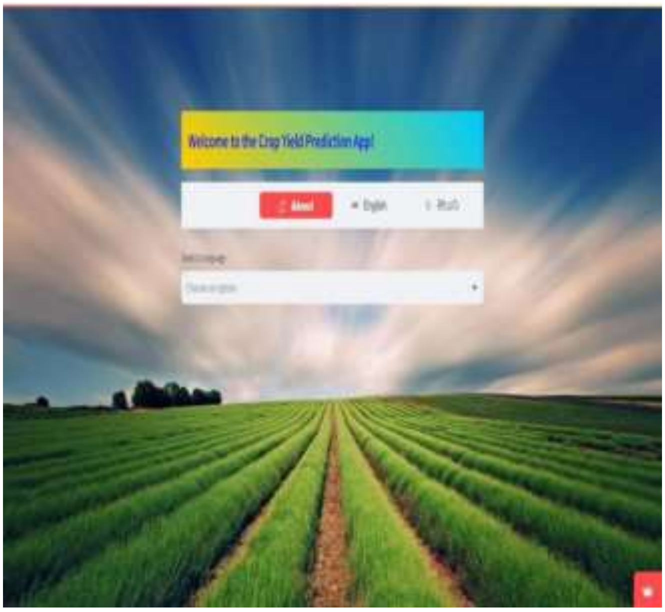


Fig a : Process Diagram

III. SIMULATION & RESULTS

MAPPING OCCUPATION THE LAND BY ASSISTED IMAGE CLASSIFICATION

The activity flow diagram depicts the step-by-step path users take to predict crop yield efficiently.

Users provide input on different factors influencing crop growth, which the system verifies. Following validation, the system analyzes these inputs and determines if a trained model exists. If not, it proceeds to train one using regression algorithms or neural networks. After training, the system forecasts crop yield using the provided parameters and delivers the results to users, accompanied by confidence intervals and explanations, facilitating informed decision-making.Predicting crop yield through machine learning entails using past data on factors like weather, soil quality, crop types, and farming methods to anticipate future yields. Through training on this data, machine learning algorithms detect patterns and connections, improving the accuracy of predictions. These forecasts aid farmers in making well-informed choices concerning planting times, irrigation,

fertilization, and pest control, thereby enhancing crop output. By embracing sophisticated technologies such as machine learning, those in agriculture can reduce risks, enhance productivity, and bolster food security.

Fig b : Homepage

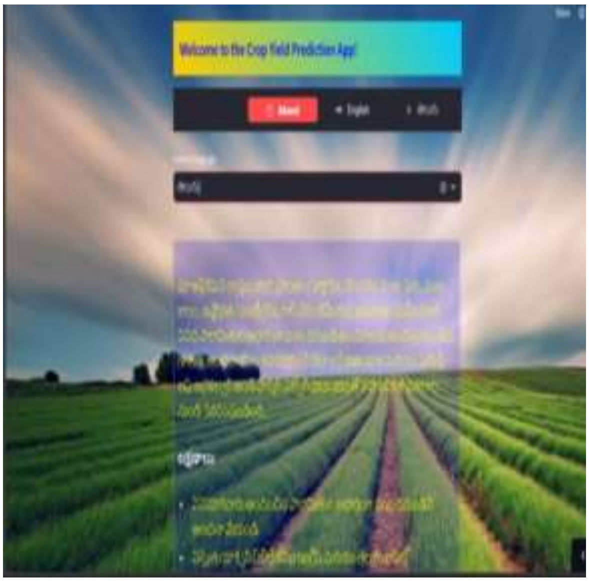


Fig c: Homepage with Telugu Language

The information provided is to be input data for a crop yield prediction model using machine learning. It includes details such as the location (Andaman and Nicobar Islands), district (Nicobars), season (Kharif), crop (Arecanut), area (2.43 hectares), and environmental factors like wind speed, precipitation, humidity, and soil type (clay). This data will be used by the machine learning model to predict the yield of Arecanut based on these parameters,for a more accurate prediction, additional data such as temperature, sunlight duration, pest infestation, fertilization practices, and historical yield data for the region could be included. These factors play crucial roles in determining crop yield and would enhance the performance of the machine learning model.The below diagram is in English Language.

ig d: Home page with English Language

Fig e :prediction page

The information provided is to be input data for a crop yield prediction model using machine learning. It includes details such as the location (Andaman and Nicobar Islands), district (Nicobars), season (Kharif), crop (Arecanut), area (2.43 hectares), and environmental factors like wind speed, precipitation, humidity, and soil type (clay). This data will be used by the machine learning model to predict the yield of Arecanut based on these parameters,for a more accurate

prediction, additional data such as temperature, sunlight duration, pest infestation, fertilization practices, and historical yield data for the region could be included. These factors play crucial roles in determining crop yield and would enhance the performance of the machine learning model.The below diagram is in Telugu Language.

Fig f : Telugu Language Page

Fig g : Prediction page

Output:

The primary outcome of the "Crop Yield Prediction" initiative is the forecasted crop yield, depicted as a decimal number. Utilizing historical data and advanced machine learning techniques, the model will analyze input parameters to generate a precise estimation of crop yield. This estimation serves as a valuable tool for farmers, enabling them to make informed decisions regarding crop planning, resource allocation, and farming techniques optimization. The anticipated crop yield will be presented in a clear and comprehensible manner, likely in numerical form.

Fig h : Output Screen

fig :

Fig

Location: NICOBARS, Andaman and Nicobar Islands
Season: Kharif
Soil Type: Clay
Weather Conditions:
Wind Speed: 0.27
Precipitation: 1009
Humidity: 37%
Crop Details:
Crop: Arecanut
Area Cultivated: 2.43 hectares
Yield Prediction:
Predicted Yield: 4.625 quintals per hectare
Total Predicted Production: 41825 quintals.

IV. PERFORMANCE METRICS:

Fig :

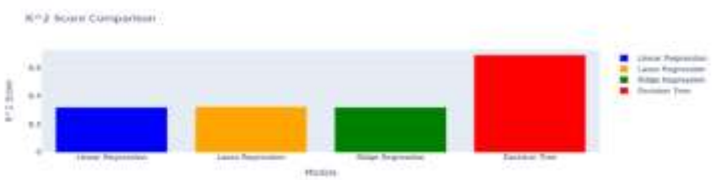


Fig :

V.DISCUSSIONS

This systematic literature review (SLR) on machine learning (ML) for crop yield prediction scrutinizes the breadth of research methodologies, with a focus on algorithm usage, feature selection, evaluation parameters, and challenges. The SLR's robust search strategy and clear methodology ensure validity and reliability, capturing a wide array of studies. It highlights Linear Regression's role as a benchmark rather than a top performer, showing promise due to its superior capabilities like automatic feature extraction. Key features for prediction include soil type, rainfall, and temperature,emerges as the primary evaluation metric. Challenges remain in data collection, model implementation, and achieving precision in predictions, underscoring the need for further research and interdisciplinary collaboration to enhance the practical application of ML agriculture.

VI. CONCLUSION

Farmers can leverage machine learning to determine the most lucrative crops for planting, considering market demand and environmental conditions. Through analyzing past market trends and weather data, machine learning algorithms forecast crop demand and recommend ideal planting schedules and locations.This study highlights the diversity of features and machine learning algorithms utilized in crop yield prediction research. While no consensus exists on the best model, random forest, neural networks, linear regression, and gradient boosting tree are commonly employed. Despite Neural Networks predominance, deep learning methods like CNN, LSTM, and DNN show promise. This research sets the stage for future endeavors in DL-based crop yield prediction models. Moving forward, our focus will be on developing and refining such models to enhance accuracy and applicability in agricultural settings.

REFERENCES

[1] Swati Vashisht, Praveen Kumar, Praveen Kumar, Improvised Extreme Learning Machine for Crop Yield Prediction, 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM).

[2] Ranjani J, V.K.G Kalaiselvi, A. Sheela, Deepika Sree D, Janaki G, Crop Yield Prediction Using Machine Learning Algorithm, 2021 4th International Conference on Computing and Communications Technologies (ICCCT).

[3] Fariha Shahrin; Labiba Zahin; Ramisa Rahman; ASM Jahir Hossain; Abdulla Hil Kaf; A.K.M Abdul Malek Azad , Agricultural Analysis and Crop Yield Prediction of Habiganj using Multispectral Bands of Satellite Imagery with Machine Learning , 2020 11th International Conference on Electrical and Computer Engineering (ICECE)

[4] Yotsaphat Kittichotsatsawat, Nakorn Tippayawong & Korrakot Yaibuathet Tippayawong, Prediction of arabica coffee production using artificial neural networkand multiple linear regression techniques, Scientific Reports volume 12, Article number: 14488 (2022)

[5] Glorot, X., and Bengio, Y. (2010). “Understanding the difficulty of training deep feedforward neural networks,” in Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics (Sardinia), 249–256.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). Deep Learning, Vol. 1. Cambridge: MIT Press Cambridge.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). “Deep residual learning for image recognition,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (Las Vegas), 770–778.

Heslot, N., Akdemir, D., Sorrells, M. E., and Jannink, J.-L. (2014). Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. Theor. Appl. Genet. 127, 463–480. doi: 10.1007/s00122-013-2231-5

Horie, T., Yajima, M., and Nakagawa, H. (1992). Yield forecasting. Agric. Syst. 40, 211–236. doi: 10.1016/0308-521X(92)90022-G

Hornik, K., Stinchcombe, M., and White, H. (1990). Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. Neural Netw. 3, 551–560. doi: 10.1016/0893-6080(90)90005-6

Hu, Q., and Buyanovsky, G. (2003). Climate effects on corn yield in missouri. J. Appl. Meteorol. 42, 1626–1635. doi: 10.1175/1520-0450(2003)042<1626:CEOCYI>2.0.CO;2

Ioffe, S., and Szegedy, C. (2015). Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv [preprint] arXiv:1502.03167.

Kaur, A., Sharma, J., and Agrawal, S. (2011). Artificial neural networks in forecasting maximum and minimum relative humidity. Int. J. Comput. Sci. Netw. Secur. 11, 197–199.

Khaki, S. (2019). Source Code. Available online at: https://github.com/saeedkhaki92/Yield-Prediction-DNN.

Kingma, D. P., and Ba, J. (2014). Adam: a method for stochastic optimization. arXiv [preprint] arXiv:1412.6980.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature 521, 436. doi: 10.1038/nature14539

LeCun, Y., Bottou, L., Orr, G. B., and Müller, K.-R. (1998). “Efficient backprop,” in Neural Networks: Tricks of the Trade, eds G. Montavon, G. B. Orr, and K. R. Müller (Berlin; Heidelberg: Springer), 9–50.

Liu, J., Goering, C. E., and Tian, L. (2001). A neural network for setting target corn yields. Trans. ASAE 44, 705. doi: 10.13031/2013.6097

Maqsood, I., Khan, M. R., and Abraham, A. (2004). An ensemble of neural networks for weather forecasting. Neural Comput. Appl. 13, 112–122. doi: 10.1007/s00521-004-0413-4

Marko, O., Brdar, S., Panic, M., Lugonja, P., and Crnojevic, V. (2016). Soybean varieties portfolio optimisation based on yield prediction. Comput. Electron. Agric. 127, 467–474. doi: 10.1016/j.compag.2016.07.009

Marko, O., Brdar, S., Panić, M., Šašić, I., Despotović, D., Knežević, M., et al. (2017). Portfolio optimization for seed selection in diverse weather scenarios. PLOS ONE 12:e0184198. doi: 10.1371/journal.pone.0184198

Montesinos-López, O. A., Montesinos-López, A., Crossa, J., Montesinos-López, J. C., Mota-Sanchez, D., Estrada-González, F., et al. (2018). Prediction of multiple trait and multiple environment genomic data using recommender systems. G3 (Bethesda) 8, 131–147. doi: 10.1534/g3.117.300309

PubMed Abstract | CrossRef Full Text | Google Scholar