AI/ML BASED MODEL FOR CROP MONITORING AND CROP YIELD PREDICTION

A LAB BASED PROJECT REPORT

Technical Report

04/05/2023

Submitted by

Murthathi Mahi Babu(20114058)

Priya(20114077)

Nimmagadda Vasavi(20114064)

Under the guidance of

Prof.Dharmendra Singh.

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

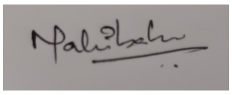
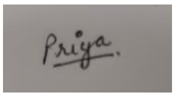
ROORKEE-247667

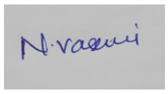
May 2023

CANDIDATE'S DECLARATION

I declare that the work carried out in this report entitled "AI/MLbased model for crop monitoring and crop yield prediction" is presented on behalf of the fulfillment of the course CSN-300 submitted to the **Department of Computer Science and Engineering**, **Indian Institute of Technology Roorkee** under the supervision and guidance of **Prof.Dharmendra Singh, Dept CSE**

I further certify that the work presented in this report has not been submitted anywhere for any kind of certification or award of any other degree/diploma.



Date: 04/05/2023

Place: Roorkee.

Signature

(Student Name, enrollment no.)

MurthatiMahibabu (20114058)

Priya (20114077)

Nimmagadda Vasavi (20114064)

CERTIFICATE

This is to certify that the above statement made by the candidates is correct to the best of my knowledge and belief.

Date: 04/05/2023

Place: Roorkee

(Signature of the Supervisor)

Tables of content:

Acronyms

- **NDVI:** Normalized Difference Vegetation Index
  NDVI = (NIR-Red)/(NIR+Red)
- **SAVI:** Soil-Adjusted Vegetation Index
  SAVI = (1+L)(NIR-Red)/(NIR+Red+L)
- **DVI:** Difference Vegetation Index
  DVI = NIR-Red
- **RVI:** Ratio Vegetation Index
  RVI = NIR/Red
- **IPVI:** Infrared Percentage Vegetation Index
  IPVI = (NIR-Blue)/(NIR+Blue)
- **MSAVI:** Modified Soil-Adjusted Index.
  MSAVI = (2*NIR+1-sqrt((2*NIR+1)^2-8*(NIR-Red)))/2
- **GNDVI:** Green Normalized Difference Vegetation Index
  GNDVI = (NIR-Green)/(NIR+Green)
- **RF:** Random Forest
- **LSTM:** Long short term Memory

Acknowledgment

We would like to extend our sincere gratitude to all those who have contributed to the success of this project. Firstly, we would like to express our thanks to our professor Dr. Dharmendra Singh Sir for his guidance, insights, and unwavering support throughout the project. His expertise and mentorship have been instrumental in shaping our project.

We also want to thank Mr. Ajay Kumar Maurya Sir (PhD student at IIT Roorkee) for his valuable insights in completing the project. Finally, we would like to acknowledge the contribution of our batchmates who helped us with the project.

We are grateful to have had the opportunity to undertake this project.

Abstract

To meet the challenges of climate change, increasing population and food demand, a timely, accurate and reliable estimation of crop yield at a large scale is more imperative than ever for crop management, food security evaluation, food trade and policy-making. In this study, taking the major wheat production regions of Haridwar(District of Uttarakhand,India) as an example, we compared a traditional machine learning method RF(random forest) and LSTM (long short-term memory networks) to predict crop yields by integrating with the modis satellite data including climate, satellite, soil properties, and spatial information data.

They all performed well for wheat field prediction at a county level from 2018 to 2020, with Rmse<=0.15. The model comparsions showed that the performance of RF was not always worse than LSTM at both the county and filed levels. Our findings demonstrated a scalable, simple and inexpensive framework for estimating crop yields at a various scales in a timely manner and with reliable accuracy, which has important implications for crop yield forecasting, agricultural disaster monitoring, food trade policy, and food security warning.

4. Introduction

It is estimated that between 2019 and 21, around 56 crore Indians, or 40.6 percent of the total population, had moderate or severe food insecurity or inadequate food supply, and by 2050, there will be two billion more people to feed and an approximately double global demand for food **[1]**.
In India the population mostly lives on wheat and paddy,but the predictions of yield for these crops is necessary to avoid inadequate food crises and also important for the farmers to produce more quantitatively. We use Satellite remote sensing data as it provides strong advantages over other monitoring techniques by providing a timely, synoptic, and up-to-date overview of large-scale crop monitoring at multiple stages (GuanSatellite-based, remote sensing crop monitoring techniques (which can acquire spatiotemporal data across large spatial scales) are used to predict the crop yield.

5. Theoretical background

The recent rapid development in artificial intelligence has led to an increase machine learning (ML) and deep learning (DL) algorithms, which have been successfully applied in various domains and obviously outperform other traditional techniques Therefore, some recent studies have applied DL and ML for yield estimation including support vector machines (SVMs), deep neural networks (DNNs), convolutional neural networks (CNNs), and long-short-term memory (LSTM).For example,**[1]** CNN and LSTM methods are adopted to estimate soybean yields in the United States.In **[2]** SVM and a two inner-product layer neural network is used to estimate maize yields The results outperformed traditional remote-sensing-based methods and USDA national-level estimation.Nevertheless, the application of ML and DL to crop yield estimation is in its infant stage. In particular, the potential of application of available datasets from satellite images and different ML and DL methods for crop yield estimation can provide better results than any other traditional methods.

6. Objectives

Following objectives have been proposed for this project:
1)Compare RF and LSTM algorithms in estimating wheat yield using multi-source data in pixel levels over large areas.
2)Understand the major factors that are essential for crop yield estimation by doing univariate analysis over each feature and yield.
3)Finally provide a scalable,simple and inexpensive operational model framework for accurately and timely estimating crop yield.

7. Motivation

Predicting crop yields is one of the most challenging tasks in agriculture. It plays an essential role in decision making at global, regional, and field levels. Soil, meteorological, environmental, and crop parameters are used to predict crop yield.Here In this model we are trying to find an appropriate AI/ML model for crop yield prediction by using timely input data for various features which are essential for yield prediction.To develop crop yield prediction model,It involves deep understanding of various ML models as 3D input data is given to predict the final output.As a CS student experimenting on various ML models and also extracting data for satellite images are also motivating factors for this project.

8. Literature review and Challenges

The first crop yield estimation using remote sensing data was proposed in the late 1970s. In **[3]**, they estimated harvests in counties of strategic interest. The existing remote sensing yield estimation methods have been mainly based on empirical relationships between vegetation indices (VIs) and field-measured yields. The main drawback to these methods is that the models are only applicable to specific crop cultivars,locations and years and not to a large-scale geographical region, which could create problems with the extrapolation of equations to another year and location. Therefore, new field-measurement yields are also needed for each new setting or crop simulation models assimilated with remote sensing data can better accommodate changes in location, weather and timing of images and estimate crop yield for each pixel.But they often need more detailed input data, such as site-specific soil and daily weather data. Both the computational costs and data requirements can hinder scaling the approach to multiple crops, regions and years without a significant investment of time, money, and labor.

9. Problem statement

An appropriate AI/Ml model for crop monitoring and yield prediction.
In this project we are Compare various ML/AI models for the crop yield prediction for the dataset obtained from satellite data.We also tried to compare the importance of various features like
NDVI (Normalized Difference Vegetation Index),
SAVI (Soil-Adjusted Vegetation Index),
DVI (Difference Vegetation Index),
RVI (Ratio Vegetation Index),
IPVI (Infrared Percentage Vegetation Index),
MSAVI (Modified Soil-Adjusted Index),
GNDVI (Green Normalized Difference Vegetation Index) for crop yield prediction and tried to optimize the no.of features in the dataset by using univariate analysis of each feature.

10. Model Architecture

The ML models which we have employed for the Crop yield prediction are – Random Forest, and LSTM(Long-Short-term memory).For tuning the models, the dataset has been divided into training and testing in 47% and 53% in random forest and LSTM model, 2018 and 2019 data has been used for training and 2020 data for testing.

11.Methods

11.1. Random Forest Regression:

Random Forest Regression is a powerful and flexible algorithm that can handle a wide range of regression problems, including those with complex and non-linear relationships between the input and output variables. It is also able to handle missing data and noisy input features. It randomly selects a subset of the training data and a subset of the input features to build each decision tree. This process is repeated to build multiple decision trees, each with different subsets of the data and features.outputs the predictions by combining decisions from a sequence of base models in the case of regression problems. Moreover, the RF model is an additive model; more formally, we can express this model as follows:

$$g(x) = f(x) + f1(x) + f2(x) + ... + fn(x)$$

where the model g(x) is the sum of simple base models fn(x). Each f (x) is a simple decision tree. All the base models are independently built using a different subsample of the data.

Basically we will train the model with input data which has features like NDVI, SAVI, DVI, RVI, IPVI, MSAVI and we will try to predict the Yield of a crop. Here, we will try to analyze each feature of an input by a Random Forest Model and also we will try to analyze each feature by plotting graph and rmse values. We also calculated yield and rmse value using all the features.

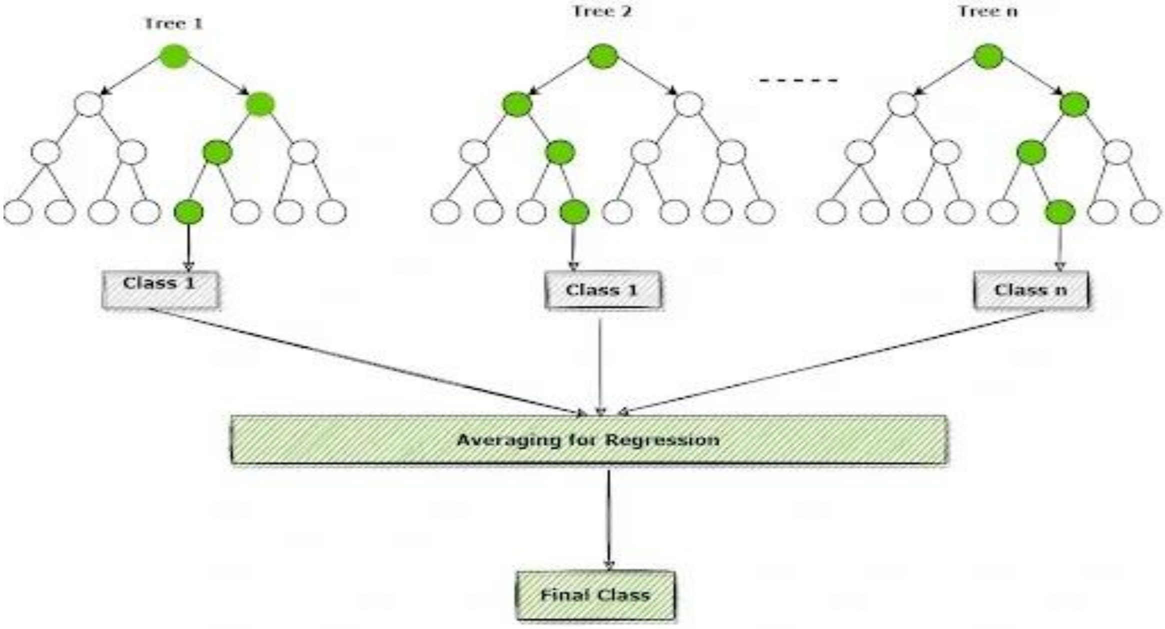


Fig 1: Random forest flow chart.

11.2. LSTM model:

An LSTM, a basic form of RNN (recursive neural network), needs to input sequential data. Each time step (ht) depends on the previous step (ht-1) and outside input (xt), then produces output (ot) at this moment and offers this time step (ht) for the next step. Finally, a fully connected layer maps the output.

The sequential data, including EVI and weather data, were dealt with using LSTMs. Finally, a fully connected layer maps the output of the last time step cell to the output node. The LSTM has time steps and three hidden layers, and each LSTM has 50 hidden units.

The non-sequential data are appended to the last hidden state,which is then fully connected to the output layer with yield.Other parameters like maximum iterations, mini-batch size, ReLU activation function, dropout, learning rate and optimizer were adjusted according to the model validation requirements.We calculate the yield for each input feature timely data and yield for all input features in timely series.

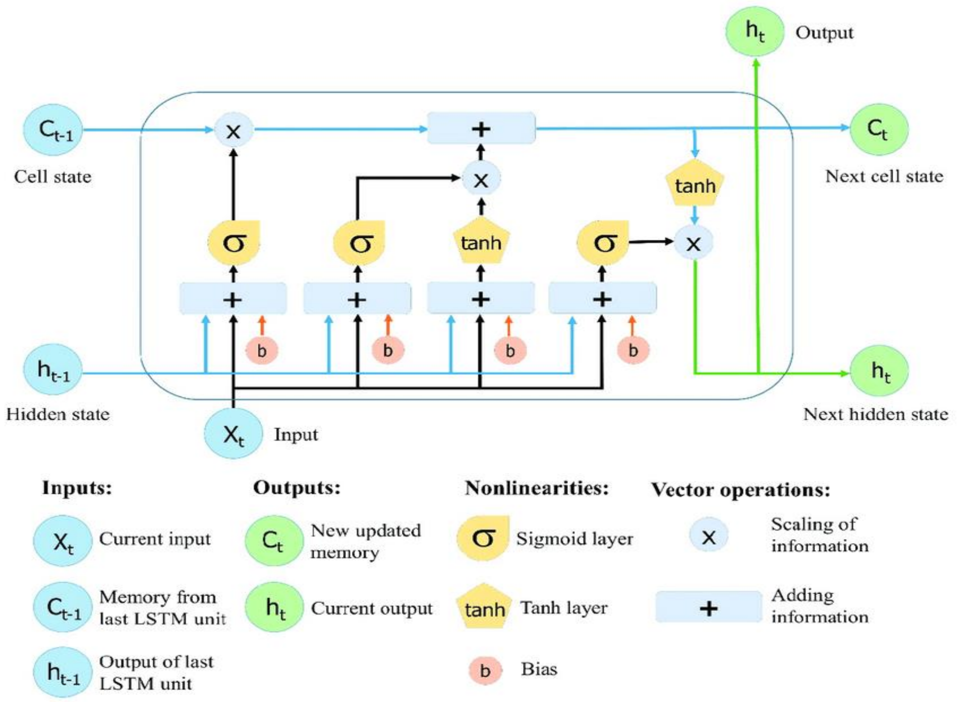


Fig 2 : Flow chart for LSTM

11.3. Adam Optimization Technique:

Adam is an optimization algorithm commonly used for training neural networks.It stands for a Adaptive Moment Estimation and is designed to be computationally efficient and perform well on problems with large datasets and or high dimensional parameter spaces.The Adam optimiser maintains a running estimate of the first and second moments of the gradients of the model parameters.These estimates are used to update the parameters at each iteration in the training process.

The update rule for the Adam optimiser isa combination of two components: a momentum term, which is similar to the momentum optimizer and a scaled gradient term, which is similar to the RMSProp optimizer.The momentum term helps the optimizer move faster in the right direction, while the scaled gradient term helps prevent oscillations and overshooting of the optimal solution.

One of the key advantages of the Adam is that it automatically adapts the learning rate based on the estimated variance of the gradients, which can be useful for the problems with sparse gradients or noisy data. Additionally Adam is well suited for the problems with non-stationary objective functions, where the optimal solution may change over time

Overall, Adam is a popular choice for deep learning tasks due to its ability to converge quickly and reliably and its robustness to hyper parameter tuning.

12. Dataset Analysis

12.1. Study area:

In this project, our center of study for the input data to test the model are mainly from the fields of Jamalpur village in Uttarakhand,Haridwar district. Latitude and longitude for the region of the study are 29°55′50°N,77°57′52°E respectively. The climate here is pleasant, gentle breeze and hot and humid during the months of may and april.

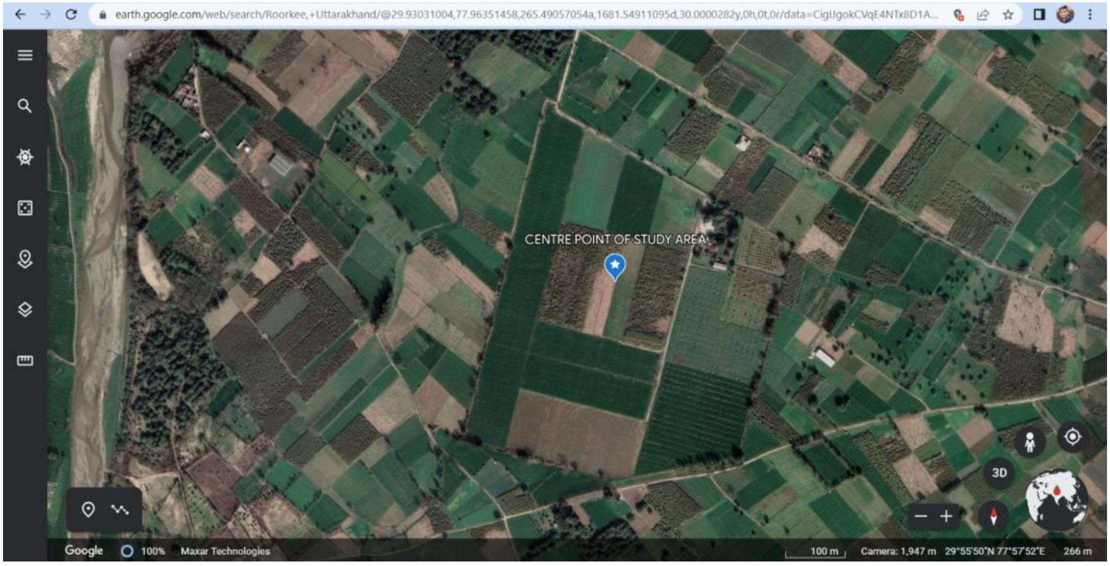


Fig 3: Study area view.

12.2. Features in Dataset and Remote sensing data:

Sentinel-2 is designed to provide high-resolution optical imagery of the Earth's surface. Sentinel-2 is captured by two identical satellites,Sentinel-2a and sentinel-2b, that orbit the earth at an altitude of 768 km. The satellites have a revisit time of five days at the equator, which means that they can capture images of the same location on earth every five days. The sensors on board the sentinel-2 satellites capture data in 13 spectral bands, ranging from the visible to infrared. The remote sensing data can be collected through various platforms like satellites, drones and ground-based sensors. Here we collected data from the sentinel-2 satellite. The value NDVI, SAVI, DVI, RVI, IPVI, MSAVI, GNDVI of a crop is collected with the help of sentinel-2 satellite.

Table 1: Full form and formula to calculate the data values of features using satellite data.

Features	**Full form**	**Formula to calculate values**
NDVI	Normalized Difference Vegetation Index.	(NIR-Red)/(NIR + Red)
SAVI	Soil-Adjusted Vegetation Index.	(1+L)(NIR-Red)/(NIR + Red +L)
DVI	Difference Vegetation Index	NIR-Red
RVI	Ratio Vegetation Index	NIR/Red
IPVI	Infrared Percentage Vegetation Index	NIR-Blue)/(NIR + Blue)
MSAVI	Modified Soil-Adjusted Index.	2*NIR+1-sqrt((2*NIR+1)^2-8*(NIR-Red)))/2
GNDVI	Green Normalized Difference Vegetation Index	(NIR-Green)/(NIR +Green)

12.3. Data Pre-Processing & Cleaning:

The collected dataset, which was intended for use in training various machine learning models, underwent a thorough cleaning and preprocessing procedure prior to utilization. This process involved removing any outliers present in the data through data cleaning techniques, as well as replacing any null values with the most appropriate corresponding values using data pre-processing methods.
The next task was to understand the correlation of input features with each other. Furthermore, to ensure optimal performance of the machine learning models during the training and validation stages, all dataset columns used for predicting crop yield using these models underwent scaler normalization. This was done to account for the differences in the ranges and units of input feature values, ensuring that they were appropriately standardized before being passed through the models.

Here the output yield is calculated in quintal/acre.

Table 2: Input dataset including all the features used in a timely series.

DATE	NDVI	SAVI	DVI	RVI	IPVI	MSAVI	GNDVI	RAINFALL	YIELD
05-01-18	0.354	0.195	0.103	2.68	0.717	0.218	0.46	1.4	40.1
09-02-18	0.622	0.39	0.117	4.68	0.808	0.436	0.613	1.4	40.1
09-03-18	0.762	0.427	0.164	5.319	0.904	0.513	0.736	13.6	40.1
13-04-18	0.272	0.149	0.087	2.297	0.621	0.126	0.303	42.9	40.1

13. Methodology

- We first extract data values of NDVI,SAVI,DVI,RVI,IPVI,MSAVI features from data available from sentinel-2 satellite.
- The extracted data is then used to train RF and LSTM models to perform the predictions of the output yield.
- After training the model we calculate the output yield and check for rmse values obtained
- If we are able to get good rmse values then we would stop else and produce results with those values.
- If unable to get good rmse values then we will try to modify the split of test and train data.

- There are many methods to split the data into test and train data. We can either do it manually by changing the ratio of test and train data.
- We can also use inbuilt sklearn methods such as ‘stratify’ to make the split random, so the essence and balance of the dataset is preserved in both the subsets.
- We can also optimize the model by modifying hyper parameters values (hyper parameter tuning).
- We adopted Adam’s optimization technique for training the model.
- After the modifications in test and train data output is predicted once again and check for rmse values.
- Even after some iterations if the rmse values don't get any better then we provide the results with best rmse values obtained using that model.
- Then in the same way the next model is trained and results of the both models are provided.

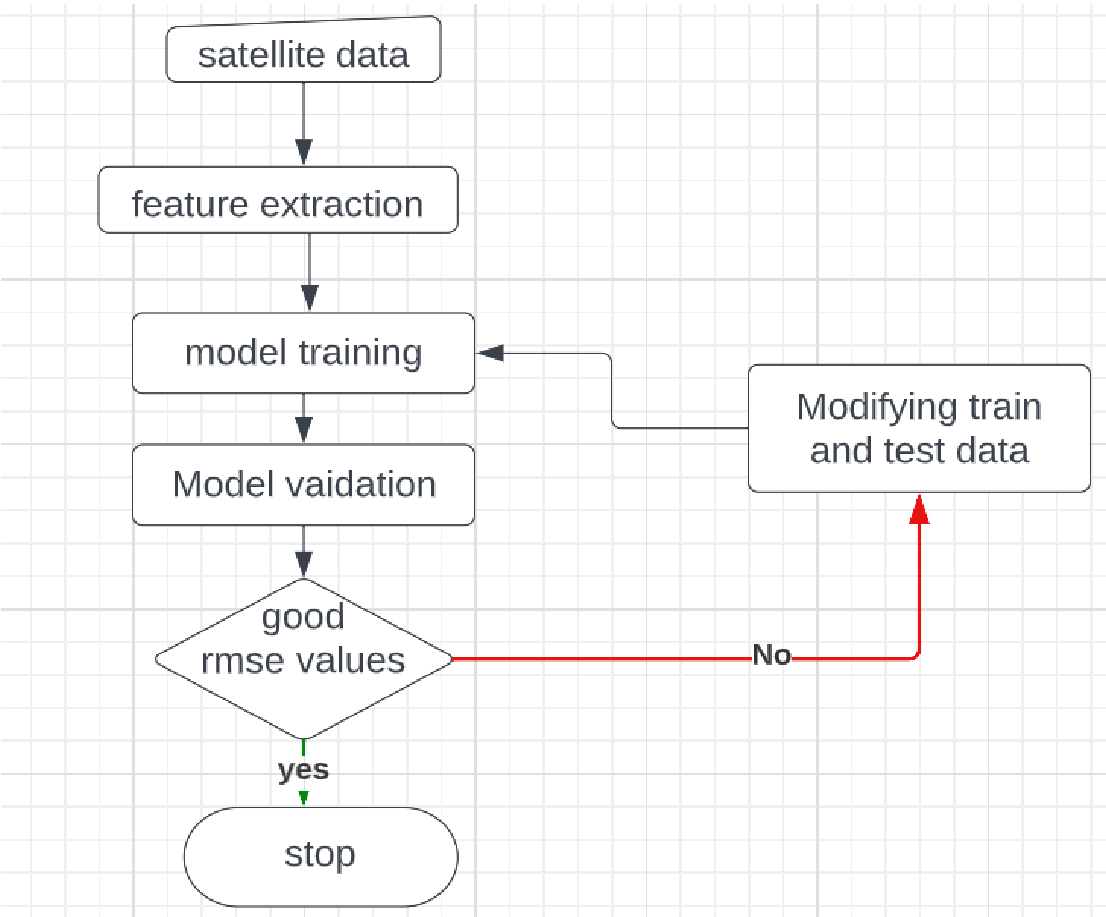


Fig 4: Flow chart of Methodology.

14. Results

After training the both RF and LSTM models with the given data, the output predicted for all features and for each feature are shown below.

- RMSE value for yield using all features by RF model: 0.1280.
- RMSE value for yield using all features by RF model: 0.1240.

(A) (B)

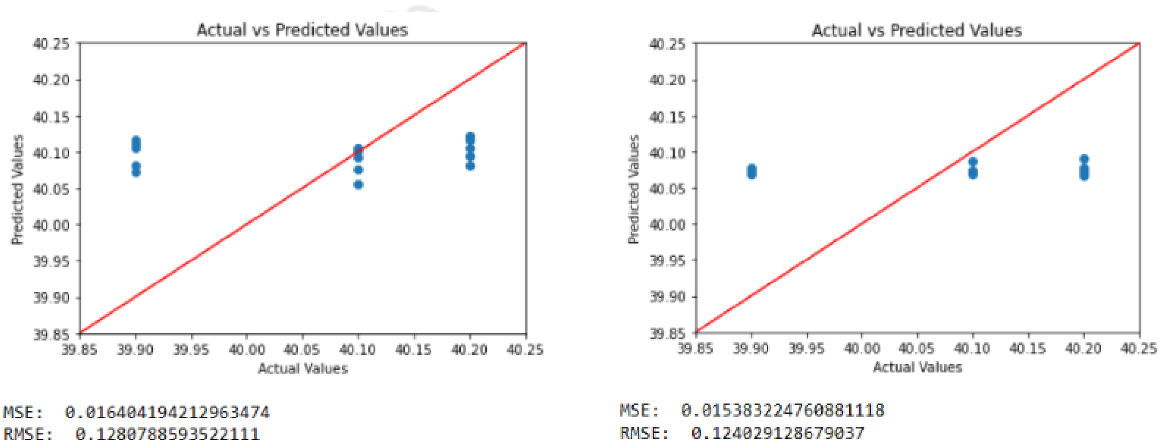


Fig 5: (A) Results for yield using all features by RF
(B) Results for yield using all features by LSTM

Yield for each feature using RF:

- RMSE value for yield using SAVI by RF model: 0.1297.
- RMSE value for yield using NDVI by RF model: 0.1357.

(A) (B)

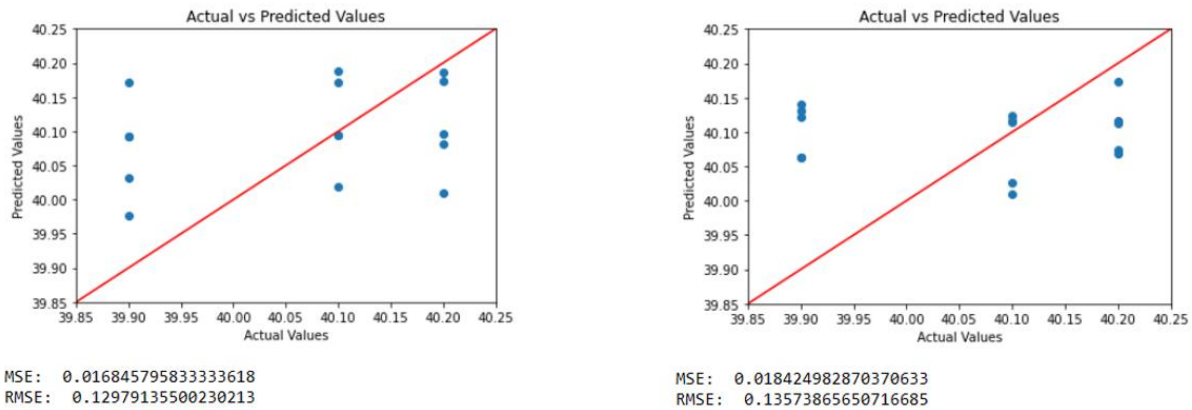
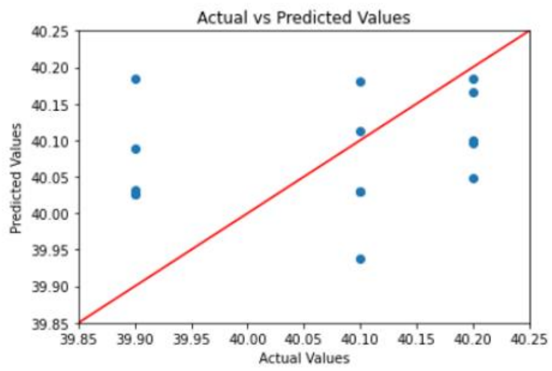


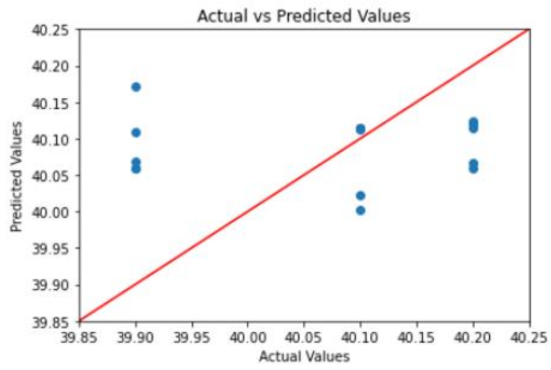
Fig 6: (A) Results for yield using SAVI
(B) Results for yield using NDVI

- RMSE value for yield using DVI by RF model: 0.1299
- RMSE value for yield using RVI by RF model: 0.1339

(A) (B)



MSE: 0.01689636215277789
RMSE: 0.12998600752687917

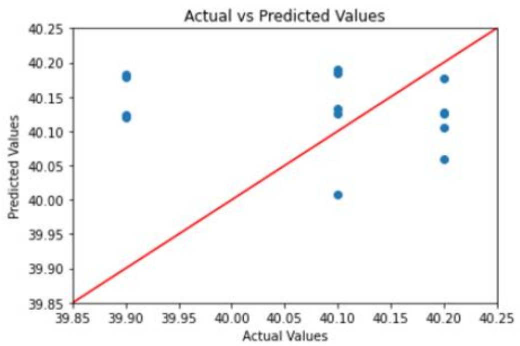


MSE: 0.017953729976852364
RMSE: 0.13399152949665274

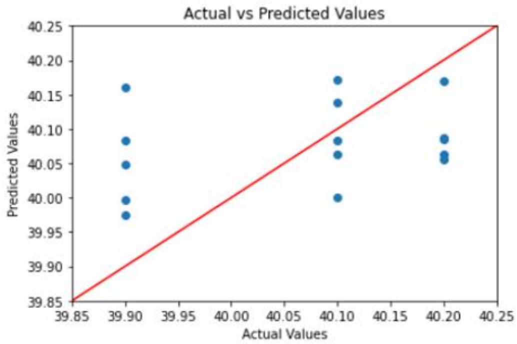
Fig 7: (A) Results for yield using DVI
(B) Results for yield using RVI

- RMSE value for yield using IPVI by RF model: 0.1632
- RMSE value for yield using MSAVI by RF model: 0.1220

(A) (B)



MSE: 0.026664895254630002
RMSE: 0.16329389227595134



MSE: 0.014897403703704381
RMSE: 0.12205492085001891

Fig 8: (A) Results for yield using IPVI
(B) Results for yield using MSAVI

Yield for each feature using LSTM:

- RMSE value for yield using NDVI by RF model: 0.1324
- RMSE value for yield using SAVI by RF model: 0.1301

(A) (B)

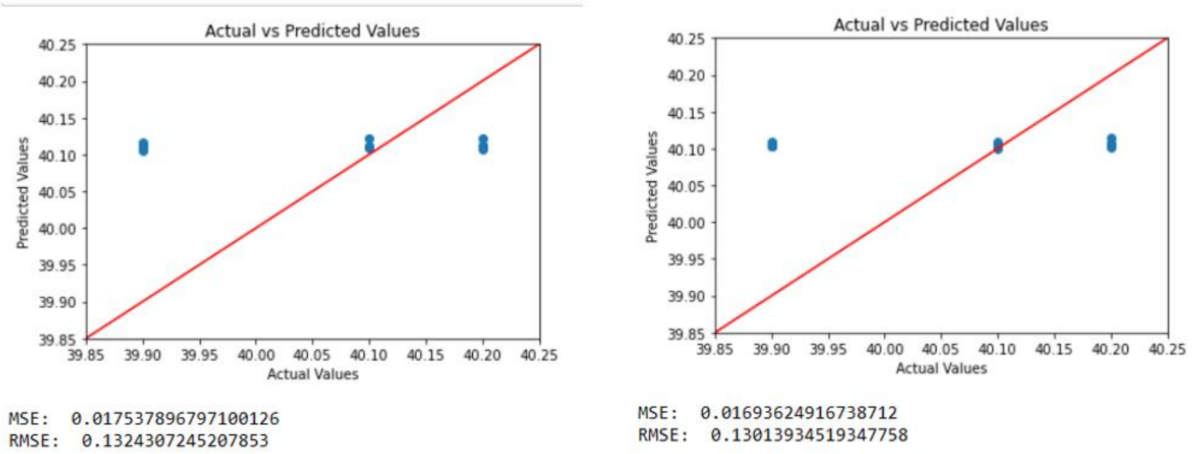


Fig 9: (A) Results for yield using NDVI
(B) Results for yield using SAVI

- RMSE value for yield using DVI by RF model: 0.1575
- RMSE value for yield using RVI by RF model: 0.1504

(A) (B)

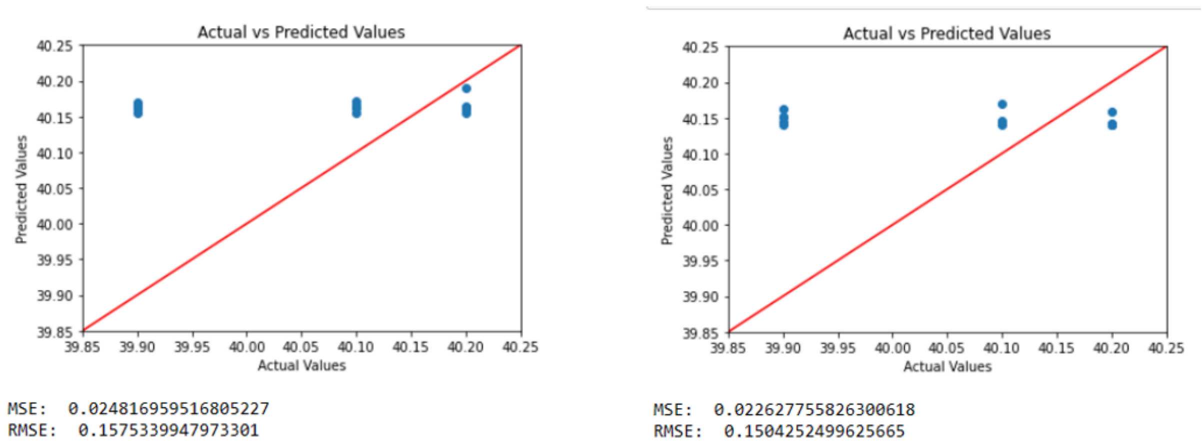


Fig 10: (A) Results for yield using DVI
(B) Results for yield using RVI

- RMSE value for yield using DVI by RF model: 0.1296
- RMSE value for yield using RVI by RF model: 0.1318

(A) (B)

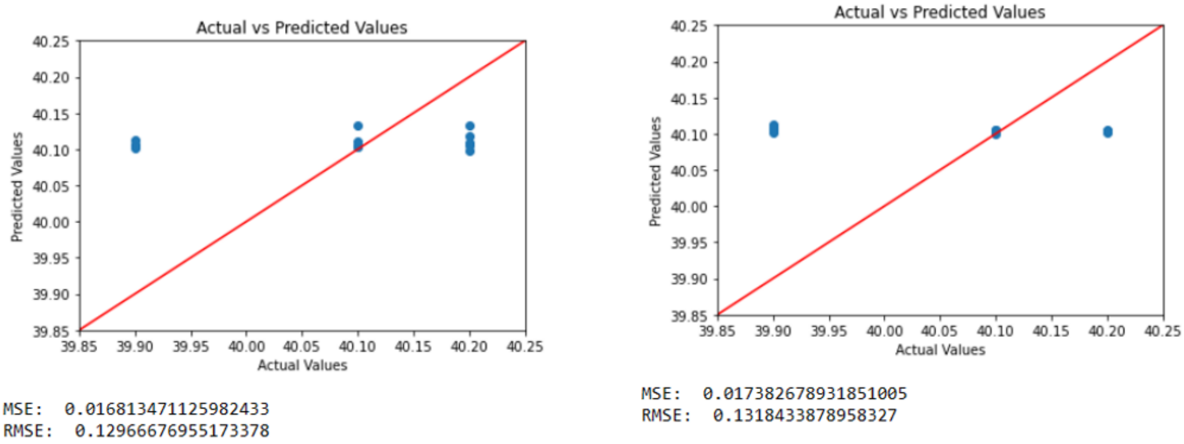


Fig 11: (A) Results for yield using IPVI
(B) Results for yield using MSAVI

- From above plots and RMSE values we can say that LSTM has predicted yield with error less than 0.15 for the given test data.
- Among all the features, IPVI gives better yield predictions with RMSE 0.129 for the LSTM model.
- Among all the features, MSAVI gives better yield predictions with RMSE 0.122 for the RF model.
  So the above two features play an important role in crop monitoring and yield prediction.

15. Conclusion

- In this study, we first processed all data including satellite and coordinates data, then used the two models, including DL model (LSTM) and one ML model (RF) to predict wheat yield.
- The LSTM model generally outperforms the random forest model in predicting the vegetation indices, as it has lower RMSE values for most of the features.
- Among the individual features, MSAVI has the lowest RMSE values for both Random forest and LSTM models,indicating the most accurately predicted feature.The combination of all features has the lowest RMSE value for the LSTM model,indicating that it is the best combination for predicting vegetation indices.
- Our findings demonstrated a new scalable, simple and inexpensive framework for estimating wheat yield on a regional scale with satellite available data which can potentially be applied to areas with sparsely observed data and worldwide for estimating crop yield.

16. Limitations

16.1. Data availability:

One Of the primary limitations of the current Ml models for crop yield prediction is the lack of the data availability. Although several datasets are available for the crop yield prediction, they are often incomplete, inconsistent and not standardized.

16.2. Complexity of the model:

Many Ml models for crop yield prediction are complex and require a considerable amount of computational resources. This makes it difficult to deploy them in resource constrained environments such as rural areas.

16.3. Lack of Interpretability:

Many ML models used for crop yield prediction are black box models which means it is difficult to interpret the model’s decision making process. This limits the ability of farmers to understand the factors that influence crop yields and the corrective measures.

17. Future Scope

- Integration of remote sensing and IoT data: The use of remote sensing and IoT data can provide more accurate and realtime information about crop growth and the other environmental factors that can influence crop yield.The integration of these data sources with ML models can enhance their accuracy and reliability.
- Developing Explainable AI models: The development of the explainable AI models that can provide insights into the decision making process can help farmers make better decisions based on the model's recommendations
- Collaborative efforts between researchers, farmers and industry can help improve the quality and availability of data for crop yield predictions. This can help develop more accurate and robust ML models for crop yield prediction. These models should be easily accessible to farmers to improve crop yields and reduce input costs.

References

1. **Centiner, H. and Burhan, K.A.R.A., 2022. Recurrent neural network-based model development for wheat yield forecasting, Adiyaman universitesi Muhendislik Bilimleri Dergisi,9(16), pp.204-218).**
2. **TY-Book,AU-Kuwata,Kentaro,Shibasaki,Ryosuke,PY-2015/07/01,SP-858,EP-861,T1-Estimating crop yields with deep learning and remotely sensed data, DO-10.1109/IGARSS.2015.7325900.**
3. **Macdonald1980GlobalCF,GlobalCropForecasting,RobertB.Macdonald and Forrest G. Hall, Science, 1980, 208, 670 – 679**
4. **Kundu, S.G., Ghosh, A., Kundu, A. and GP, G., 2022. A ML-AI ENABLED ENSEMBLE MODEL FOR PREDICTING AGRICULTURAL YIELD. *Cogent Food & Agriculture*, *8*(1), p.2085717.**
5. **TY-CHAP,AU-Sinwar,Deepak,AU-Dhaka,Vijaypal,AU-Sharma,Manoj Kumar, AU-RaniGeeta,PY-2019/10/01,SP-155,EP-180,SN-978-981-15-0662-8,TI-AI-BasedYieldPredictionandSmartIrrigation,DO 10.1007/978-981-15-0663-5_8.**
6. **N. Gandhi, O. Petkar and L. J. Armstrong, "Rice crop yield prediction using artificial neural networks," *2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR)*, Chennai, India, 2016, pp. 105-110, doi: 10.1109/TIAR.2016.7801222**
7. **R. Luciani, G. Laneve and M. JahJah, "Agricultural Monitoring, an Automatic Procedure for Crop Mapping and Yield Estimation: The Great Rift Valley of Kenya Case," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12, no. 7, pp. 2196-2208, July 2019, doi: 10.1109/JSTARS.2019.2921437.**
8. **Martínez-Ferrer, L., Piles, M. and Camps-Valls, G., 2020. Crop yield estimation and interpretability with Gaussian processes. *IEEE Geoscience and Remote Sensing Letters*, *18*(12), pp.2043-2047.**
9. **Kamir, E., Waldner, F. and Hochman, Z., 2020. Estimating wheat yields in Australia using climate records, satellite image time series and machine learning methods. *ISPRS Journal of Photogrammetry and Remote Sensing*, *160*, pp.124-135.**
10. **TY- JOUR,AU-Qiao, Mengjia,AU- He, Xiaohui,AU- Cheng, Xijie,AU-Li, Panle,AU- Luo, Haotian,AU-Tian, Zhihui,AU- Guo, Hengliang,PY - 2021/04/14 SP- 1,EP- 1,T1- Exploiting Hierarchical Features for Crop Yield Prediction Based on 3-D Convolutional Neural Networks and Multikernel Gaussian Process VL - PP,DO - 10.1109/JSTARS.2021.3073149.**
11. **E. Panwar, A. N. J. Kukunuri, D. Singh, A. K. Sharma and H. Kumar, "An Efficient Machine Learning Enabled Non-Destructive Technique for Remote Monitoring of Sugarcane Crop Health," in IEEE Access, vol. 10, pp. 75956-75970, 2022, doi: 10.1109/ACCESS.2022.3190716.**

12. **Vasit Sagan, Maitiniyazi Maimaitijiang, Sourav Bhadra, Matthew Maimaitiyiming, Davis R. Brown, Paheding Sidike, Felix B. Fritschi,Field-scale crop yield prediction using multi-temporal WorldView-3 and PlanetScope satellite data and deep learning,ISPRS Journal of Photogrammetry and Remote Sensing,Volume 174,2021,Pages 265-281,ISSN 0924-2716.**
13. **Shirui Hao, Dongryeol Ryu, Andrew Western, Eileen Perry, Heye Bogena, Harrie Jan Hendricks Franssen,Performance of a wheat yield prediction model and factors influencing the performance: A review and meta-analysis,Agricultural Systems,Volume 194,2021,103278,ISSN 0308-521X.**
14. **Bose, P., Kasabov, N.K., Bruzzone, L. and Hartono, R.N., 2016. Spiking neural networks for crop yield estimation based on spatiotemporal analysis of image time series. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(11), pp.6563-6573.**
15. **Juan Cao, Zhao Zhang, Yuchuan Luo, Liangliang Zhang, Jing Zhang, Ziyue Li, Fulu Tao,Wheat yields predictions at a county and field scale with deep learning, machine learning, and google earth engine,European Journal of Agronomy,Volume 123,2021,126204,ISSN 1161-0301.**
16. **Pantazi, X.E., Moshou, D., Alexandridis, T., Whetton, R.L. and Mouazen, A.M., 2016. Wheat yield prediction using machine learning and advanced sensing techniques. *Computers and electronics in agriculture*, *121*, pp.57-65.**
17. **Diego Gómez, Pablo Salvador, Julia Sanz, José Luis Casanova,Modeling wheat yield with antecedent information, satellite and climate data using machine learning methods in Mexico,Agricultural and Forest Meteorology, Volume 300,2021,108317,ISSN 0168-1923.**
18. **Rajkumar Dhakar, Vinay Kumar Sehgal, Debasish Chakraborty, Rabi Narayan Sahoo, Joydeep Mukherjee, Amor V.M. Ines, Soora Naresh Kumar, Paresh B. Shirsath, Somnath Baidya Roy,Field scale spatial wheat yield forecasting system under limited field data availability by integrating crop simulation model with weather forecast and satellite remote sensing,Agricultural Systems,Volume 195,2022,103299,ISSN 0308-521X.**
19. **Glorot, X. and Bengio, Y., 2010, March. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics* (pp. 249-256). JMLR Workshop and Conference Proceedings.**
20. **Mukesh Singh Boori, Komal Choudhary, Rustam Paringer, Alexander Kupriyanov,Machine learning for yield prediction in Fergana valley, Central Asia,Journal of the Saudi Society of Agricultural Sciences,Volume 22, Issue 2, 2023,Pages 107-120,ISSN 1658-077X.**
21. **Cedric, L.S., Adoni, W.Y.H., Aworka, R., Zoueu, J.T., Mutombo, F.K., Krichen, M. and Kimpolo, C.L.M., 2022. Crops yield prediction based on machine learning models: case of west african countries. *Smart Agricultural Technology*, p.100049.**
22. **Sathya P, Gnanasekaran P. (2023) Paddy Yield Prediction in Tamilnadu Delta Region Using MLR-LSTM Model. *Applied Artificial Intelligence* 37:1.**
23. **Rafael Battisti, Paulo C. Sentelhas, Kenneth J. Boote,Inter-comparison of performance of soybean crop simulation models and their ensemble in southern Brazil,Field Crops Research,Volume 200,2017,Pages 28-37,ISSN 0378-4290.**