JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue

JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Enhancing Agriculture Crop Prediction and Yield Optimization through Advanced Time Series Analysis Techniques

[1]Ramkumar Devendiran, [2]Devalapalle Nikhil Reddy, [3]Madipalli Sri Ram Gowtham,

[4]Gannabathula Sai Kanth Pushkar, [5]Gopuram Karthik,

[1]Assistant Professor Senior Grade, [2,3,4,5]UG Student,
[1,2,3,4,5] School of Computer Science and Engineering,
[1,2,3,4,5]VIT-AP University, Amaravati, Andra Pradesh, India-522237.

Abstract : With population expansion and environmental issues, crop production is a crucial component of ensuring global food security, and it must be optimized sustainably. Machine learning methods have become an effective tool in agriculture, providing creative ideas to improve crop yield. This abstract examines the application of deep learning techniques in crop agriculture, with a focus on how these techniques have the potential to completely transform this industry. A feature of machine learning, which is a subset of machine learning, is its capacity to automatically recognize and extract patterns from huge datasets. In terms of crop production, this entails utilizing a variety of data sources, such as satellite imaging, weather records, soil information, and sensor data, to guide decision-making processes at all stages of crop cultivation. Deep learning has numerous uses, and crop monitoring is one of the main uses of deep learning in agriculture. In order to provide real-time insights into crop health, neural networks can evaluate satellite and drone footage. This allows them to spot problems like pest infestations, illnesses, and nutritional deficits. This makes it possible for farmers to take prompt remedial action, limiting crop losses and the need for chemical interventions. The ability to predict yield is another crucial component of crop production. To predict future agricultural yields, machine learning models can use past data on crop performance, including environmental factors, cultivation techniques, and yield records. The Time Series Analysis Techniques used to optimize resource management, especially by farmers, allow them to make decisions about planting, harvesting, and resource allocation that directly increase efficiency.

IndexTerms **- Agriculture, Crop Prediction, Yield Optimization, Time Series Analysis, Advanced Techniques.**

I. INTRODUCTION

Agriculture is the foundation of our civilization, sustaining billions of people globally with food and a means of subsistence. The uncertainty surrounding crop output, however, is one of several difficulties the agriculture sector must overcome. Crop yields can be considerably impacted by variables including weather, soil quality, and pest infestations, making it challenging for farmers to plan properly. In recent years, new opportunities for resolving these difficulties have emerged as a result of the integration of cutting-edge technology into agriculture. The application of deep learning algorithms for crop prediction is one such technical achievement. A type of artificial intelligence called machine learning has proven to have amazing ability in a number of areas, including picture identification and natural language processing. Deep learning models can examine huge volumes of data when used in agriculture. Our data set contains 57k rows in which it contains attributes like Domain, Area, Element, Item, Year, Units, Value. we can use this huge database with the help of this database we can use deep learning techniques like linear regression, long short-term memory for training the model. These both techniques are type of recurrent neural network. For our database model we have chosen linear regression because it predicts the value of unknown data by using another related and non-data value. In this technique we can analyze the accuracy through the evaluation metrics include mean absolute error and mean square error. In this linear regression technique, it trains by itself with the help of epochs i.e., iterations and shows us the errors. This is the best technique for predicting the huge database.

Here we used Streamlit, it is a python framework readily available for the creation of web applications dedicated to Machine Learning and Data Science. With Streamlit, the development and deployment of web applications is made convenient. This framework enables users to write applications in a manner similar to that of Python code.

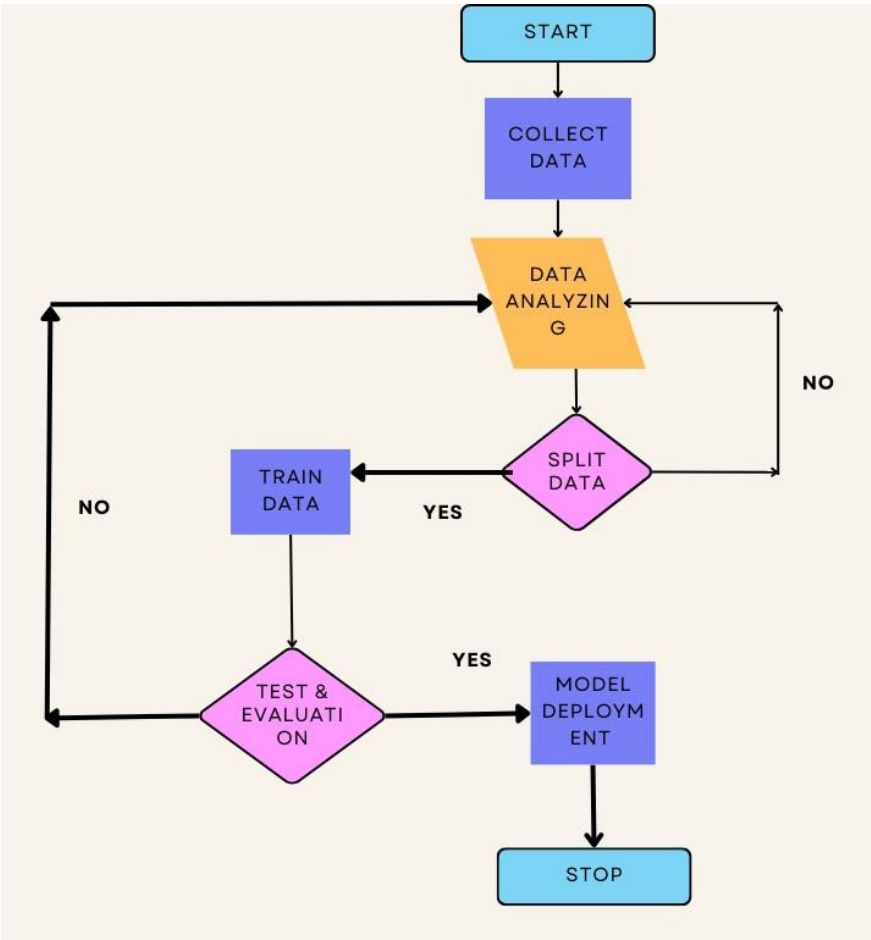

Fig 1: Flowchart of agricultural crop prediction model

Key features of our Crop Prediction Platform:

- Data Accessibility: Our user-friendly website provides easy access to our comprehensive database, ensuring that users can quickly retrieve the information they need to make informed decisions about their crops.
- Machine Learning Models: Our platform employs state-of-the-art deep learning models that have been trained on vast datasets to deliver highly accurate crop yield predictions.
- Customization: Users can tailor predictions to their specific regions and crop types, ensuring that the insights provided are relevant to their unique farming conditions.
- Historical Analysis: The platform allows users to delve into historical data trends, helping them identify patterns and correlations that can inform future planting and harvesting decisions.
- Real-time Updates: Our platform constantly updates its predictions based on real-time weather and environmental data, ensuring that users have the most current information at their fingertips.
- Collaboration and Knowledge Sharing: Users can collaborate with experts, share insights, and contribute to a growing community of agricultural data enthusiasts.

We want to provide farmers and stakeholders in the agriculture industry with the tools they need to maximize crop output, decrease risks, and ultimately contribute to food security by fusing cutting-edge deep learning algorithms with a user-friendly web interface. Our platform makes agriculture more accessible to everyone and more productive in the future.

II. Motivation

Researching and developing methods, for predicting rice crop growth using time series analysis is a task in the field of agriculture. Time series analysis involves studying data collected over a period of time to identify patterns, trends and make predictions for the future. In relation to rice cultivation this approach can offer insights. Enhance decision making for farmers, agricultural authorities and other stakeholders. Rice serves as a food for billions of people globally. Its production is highly influenced by various factors such as weather conditions, pest outbreaks and the availability of resources like water and fertilizers. By analyzing data on rice production, climate patterns, soil quality and other relevant parameters through time series analysis researchers can develop models. These models can forecast rice yields identify irregularities or anomalies in crop growth patterns and optimize the allocation of resources to ultimately improve food security.

Recent technological advancements like satellite imagery, IOT sensors and machine learning algorithms have significantly enhanced the accuracy of predicting rice crop growth. These innovations enable real time data collection and analysis which allows for responses to threats to crops as well as opportunities, for better crop management. In the face of growing population and climate change it is crucial to conduct research and development in predicting rice crop yields through time series analysis. This area of study offers potential in improving efficiency minimizing waste and supporting farmer's livelihoods while ensuring global food security. The outcomes will benefit people, around the world by providing a sustainable supply of rice while addressing needs.

The study of rice crop prediction, through time series analysis, which utilizes data to forecast rice yields is an interdisciplinary field. It combines agronomy, data science and technology to improve practices and ensure food security. By examining time series

data that includes factors like weather conditions, soil quality and crop health researchers can develop models for estimating future rice production. These predictions play a role in managing crops allocating resources efficiently and planning policies accordingly. To effectively share these insights with others it would be beneficial to create a website that serves as a central hub for sharing research findings, data sets and useful tools. This website would offer a to use interface designed for farmers, policymakers and researchers alike to access both time and historical data. Such accessibility would greatly assist in making informed decisions. Additionally, this platform could foster collaboration. Facilitate the exchange of knowledge, among stakeholders. Ultimately it will contribute towards promoting productive rice farming practices on a scale.

III. Contribution

The main contribution of this paper is that we have integrated our model with our website to provide a user-friendly interface to the farmers. These are some keys contributions in our project:

3.1Website interface

In our research paper we introduce a website interface that predicts crop yields using Linear Regression and Time Series Analysis. Users can input information, about their state and crops and our model will generate forecasts. This innovation connects the worlds of data science and agriculture providing a tool for farmers and policymakers. We have combined methods with modern web technology in our approach making it a valuable addition, to the field of crop yield prediction.

3.2Real Time data

The proposed method utilizes Linear Regression to forecast crop yields, which's an aspect of promoting sustainability, in agriculture. Additionally, we have created a user website that offers real time access, to these forecasts empowering farmers with information to make informed decisions. By combining data science and web technology we have made our model more accessible and practical. Our efforts not contribute to the field of precision agriculture. Also showcase the tangible benefits of applying machine learning to address real world challenges in the agricultural industry.

3.3Scalability

Our project showcases the power of Linear Regression in predicting crop yield. Provides a user website where users can input country and crop data to receive accurate predictions. This demonstrates the potential, for research. By expanding our dataset incorporating machine learning models and collaborating with field experts we aim to improve the accuracy and usefulness of our predictions. This research has the potential to revolutionize food production strategies by introducing precise crop yield forecasting methods and offering valuable insights, to the agricultural community.

IV. Proposed Model

In the field of agriculture and predicting crop yields, the decision of which machine learning algorithm to use is crucial. Choosing the right model can have a significant impact on the accuracy of predictions and, as a result, the effectiveness of farming practices. In our crop yield prediction project, we have selected Linear Regression as the main predictive model. We believe this selection is justified for several reasons. First, Linear Regression offers simplicity and interpretability. Second, it is suitable for predicting continuous outputs. Third, it assumes a linear relationship between variables. Fourth, it provides information about feature importance and coefficients. Additionally, its ease of integration, scalability, and efficiency make it a favorable choice.

4.1Data Collection: Gather historical data on crops, including factors like weather, soil type, crop variety, and past yields. Collaborate with relevant agricultural agencies or use publicly available datasets.

4.2Data Pre-processing

Clean and prepossess the data by handling missing values, outliers, and ensuring uniform data formats. Normalize or standardize numerical variables and encode categorical ones.

4.3Model Selection: we choose Linear Regression for time-series data. Linear Regression is a simple yet effective model for yield prediction. It assumes a linear relationship between the selected features and the target variable (yield). The model's equation can be represented as:

$$\text{Yield} = \beta 0 + \beta 1 * \text{Feature1} + \beta 2 * \text{Feature2} + ... + \beta n * \text{Feature} \quad \text{(1)}$$

Where,

Yield is the predicted crop yield.
β0 is the intercept term.
β1, β2, …, βn are the coefficients associated with each feature.
Feature1, Feature2, …, Feature are the selected input features.

4.4Data Splitting: We split the dataset into training, validation, and test sets to train, tune, and evaluate the model's performance, respectively.

4.5Evaluation: We assessed the model's performance using the test dataset. Metrics such as Mean Absolute Error (MAE), Mean Square Error (MSE), and R-squared can quantify prediction accuracy.

- We got these results for our model:
- Mean Absolute Error: 1388.9472566440481
- Mean Squared Error: 5112432.468132844

4.6 Deployment: We Developed a user-friendly website interface that allows users to access the crop prediction model. Once the model is trained and evaluated, it is put to work on a web page. The website will give the user a user-friendly interface where he can input preferences – in this case, country, crop and year that he is interested in. The website will then use the trained model to generate predictions based on the input preferences. Integrate the model into the website using Streamlit through Localhost, ensuring scalability and responsiveness. We have used render web services to deploy our website in which user can access it from his own location and get the results of his interested crop. Here we pushed our project repository to GitHub and from there we uploaded that in cloud web hosting platform for deployment by using Streamlit framework and cloud we hosted our website successfully. For initializing we used a command $Streamlit run app.py

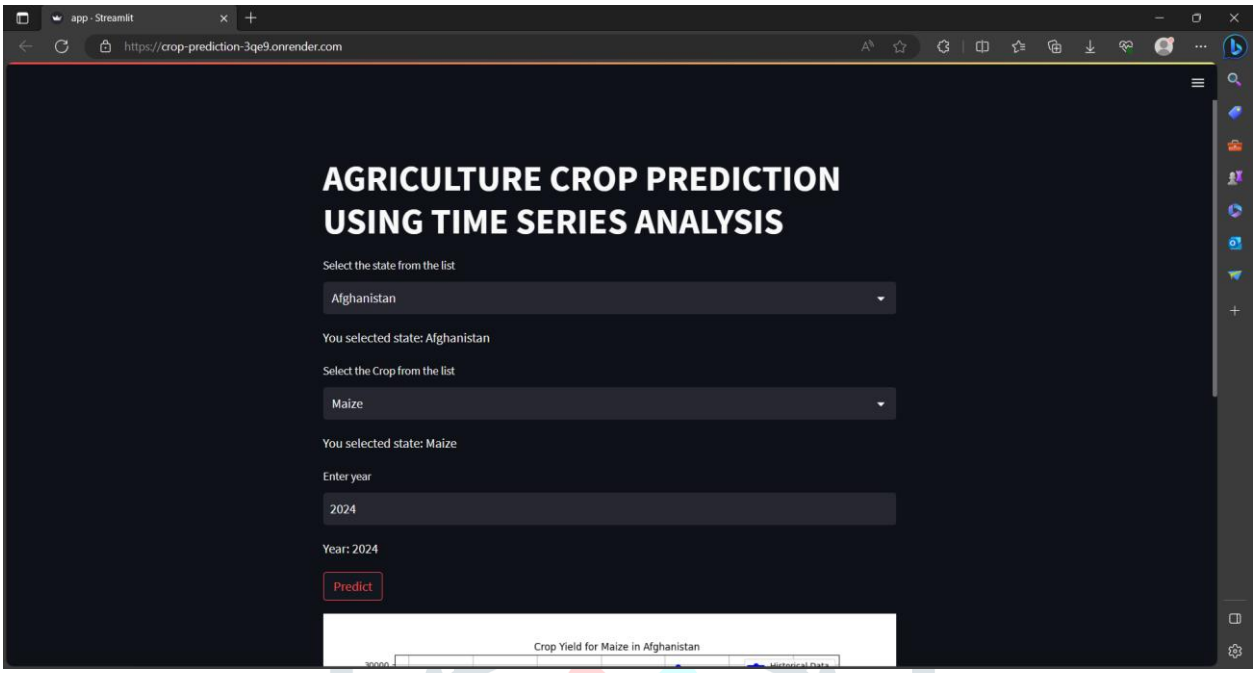

Fig 2: User interface of the Agricultural crop prediction model

4.7 User Interaction and Process of predicting crop yield: Using the website, we allowed users to enter data on the present country, particular crops and prediction year. Then these inputs given are read by the trained Linear Regression model. The trained linear regression Model at this stage which have been trained on a data set through which it was also learned and this data set has been trained on the Yield dataset which was collected from the "Kaggle website". The trained Linear Regression model read by the inputs i.e., The Country, the Crop and year and derived the same dataset we can say "Yield dataset" and it will read the file from the "Kaggle website" for the prediction of the future crop value.

By reading the file the main reading is done by using pandas. The user enters the Crop, the Country and prediction year into the website. And then the User enters these details, and then we check our Model in this way, by checking with the two inputs like Crop, the Country and year. But this is not considered, the user will be entering the information and then asking about a prediction from our Linear regression model. The Crop and the Country inputs are taken by the Linear Regression model and the model looked up the exact data that is needed in it by given its inputs on its dataset through the Linear Regression Model. With the Linear Regression model and with the given value according to the inputs the simple calculation can be performed. The Linear regression as we know Linear Regression is used to finding the best linear relationship that is possible between two sets of the data.

4.8 Detailing: Introducing the expectation results to clients in an easy-to-understand design, for example, Pie diagrams or Structured presentations, permitting them to go with informed choices in regards to planting, reaping, and crop the board. By following this methodology, users can easily access and utilize our crop prediction system through a user-friendly website, making informed decisions about agricultural practices and maximizing crop yield.

V. RESULTS & DESCRIPTION

The figure 3 shows the results of the crop rice, paddy, wheat, potatoes, maize in Afghanistan and also compared with the maize and wheat production in the form of bar graph where x-axis stands for year and y – axis stands for production.

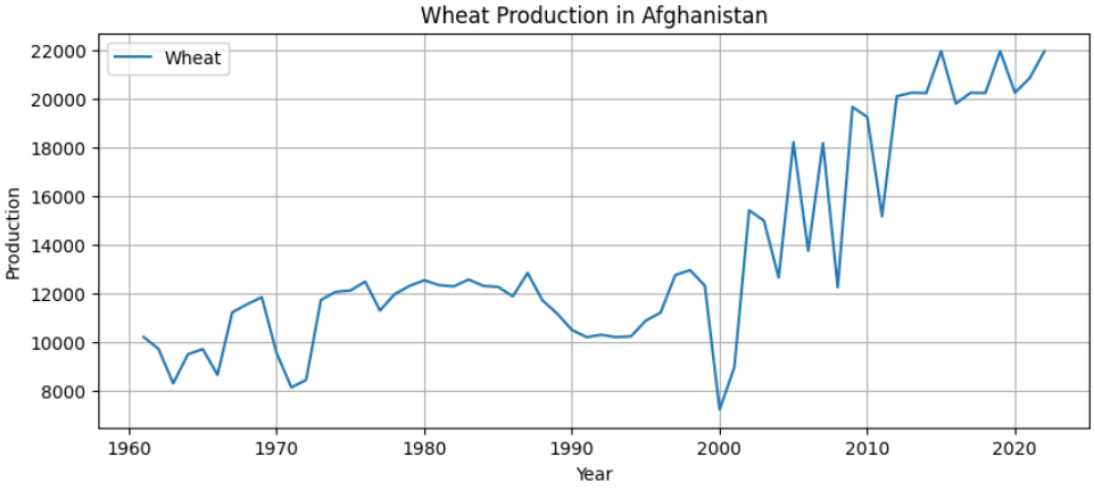

Fig 3: Wheat Production in Afghanistan

From the graph wheat production in Afghanistan seems to be volatile up to the year of 2010-12 from there it started rising. Here the production is mentioned in the form of Hg/Ha (hectogram per hectare). It achieved highest value in the year 2020 i.e., 22000 Hg/Ha.

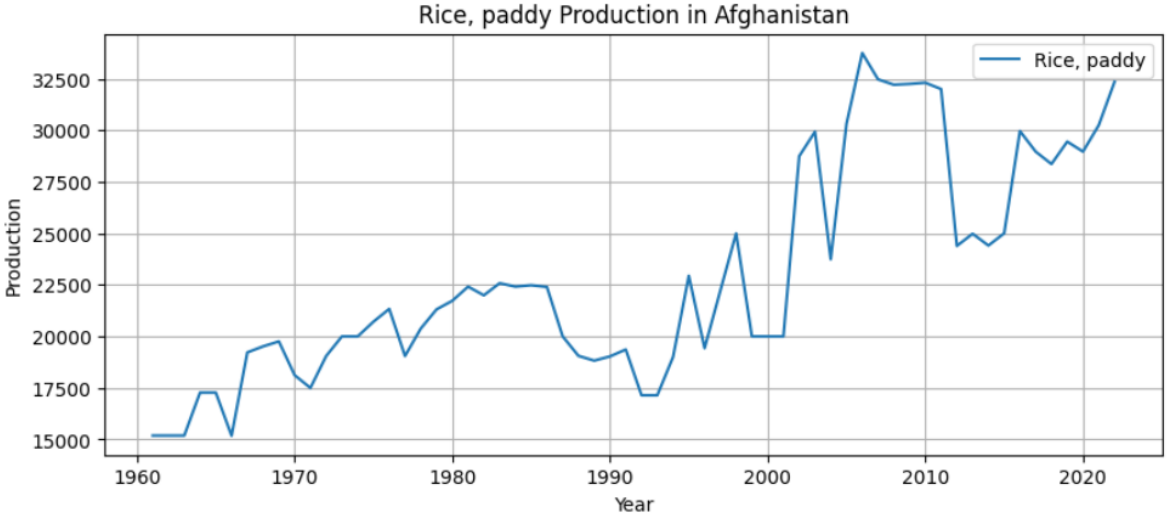

Fig 4: Rice and Paddy Production in Afghanistan

At a glance we can see from the above graph rice, paddy production in Afghanistan it experienced ups and downs throughout the research. Here the production is mentioned in the form of Hg/Ha (hectogram per hectare). It achieved highest value in the year 2022 i.e., 32500 Hg/Ha. Here the x- axis is having difference of 10 years of data and y- axis having a production value correspondingly.

Fig 5: Potatoes Production in Afghanistan

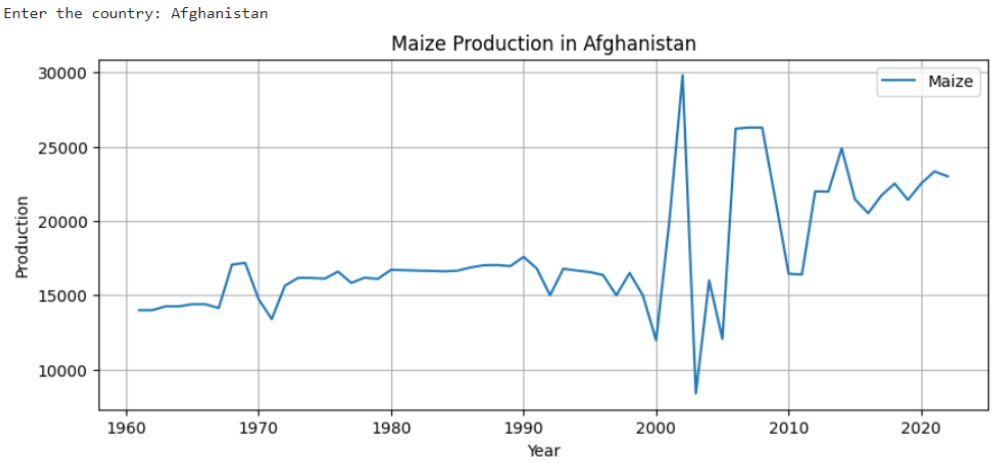

Fig 6: production of Maize in Afghanistan.

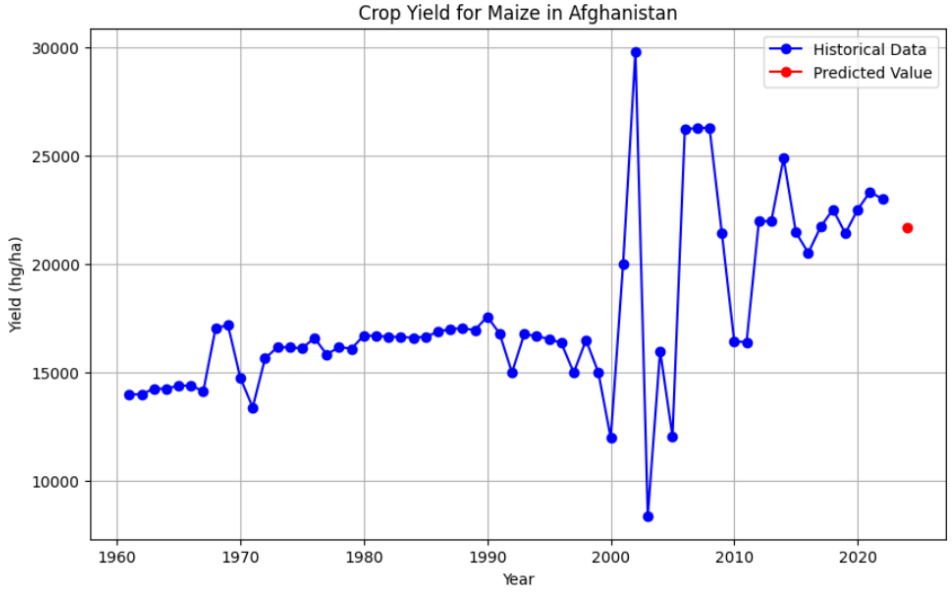

Fig 7: Prediction of crop yield for Maize in Afghanistan

This above graph represents prediction of crop yield of "Maize" in Afghanistan for the year 2024. In the above graph the blue dots represent the production of maize in that year in Afghanistan. The red dot represents the prediction for the crop yield of maize in that year using our linear regression model. The crop yield for maize in the year 2024 of Afghanistan is 21700.746055869182 hg/ha.

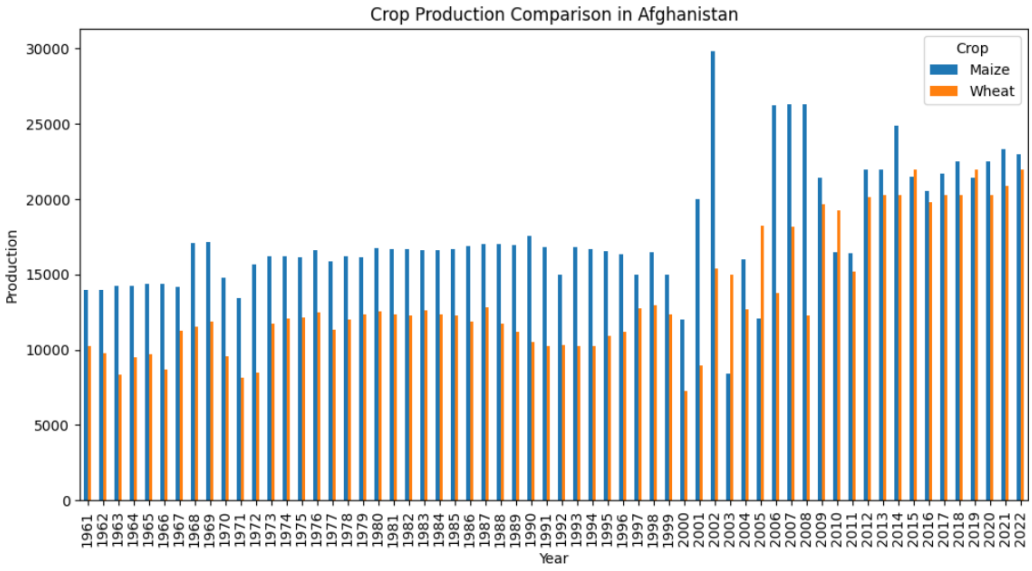

Fig 8: Comparison of Wheat & Maize crops production in Afghanistan

The Above graph represents the comparison of "Maize" and "Wheat" crops production in Afghanistan overall the years (1961 - 2022).

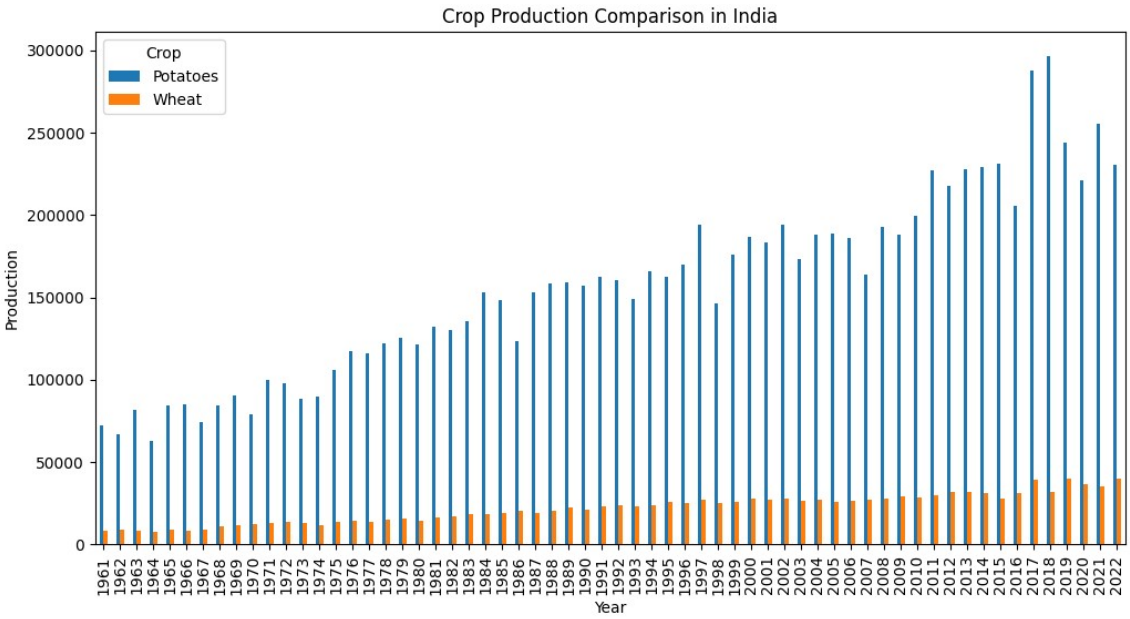

Fig 9: Comparison of Potatoes & Wheat crop Production in India

The Above graph represents the comparison of "Potatoes" and "wheat" crop production in India. Over these years' wheat is showing a minimal growth in production in India, where as potatoes are showing relatively raise and drops in its production overall these years. Potatoes has achieved its highest value in the year 2018.

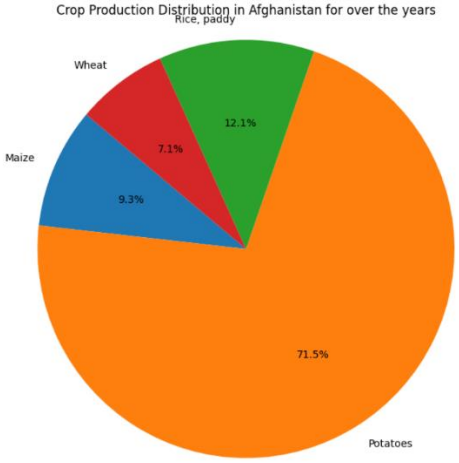

Fig 10: Pie chart of Crop Production Distribution in Afghanistan

The following pie chart gives information about the percentage of crops produced on average throughout the research. Here the potatoes hold highest distribution in Afghanistan where remaining crops together holds 1/3rd of the total distribution.

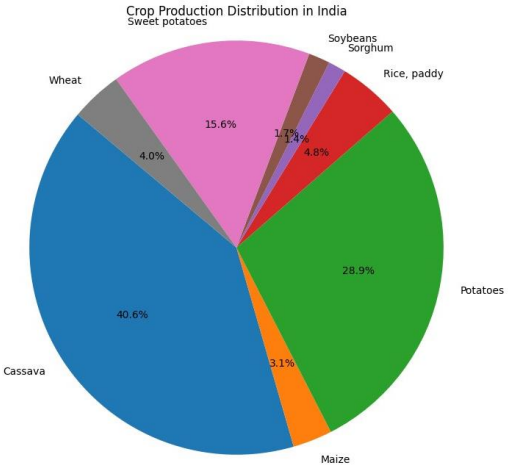

Fig 11: Pie chart of Crop Production Distribution in India

The pie chart shows the crop production distribution of different crops in India. The largest slice of the pie is for Cassava, at 40.6%. This is followed by Potatoes at 28.9%, sweet-potatoes at 15.6% and other crops in that are cultivated in India.

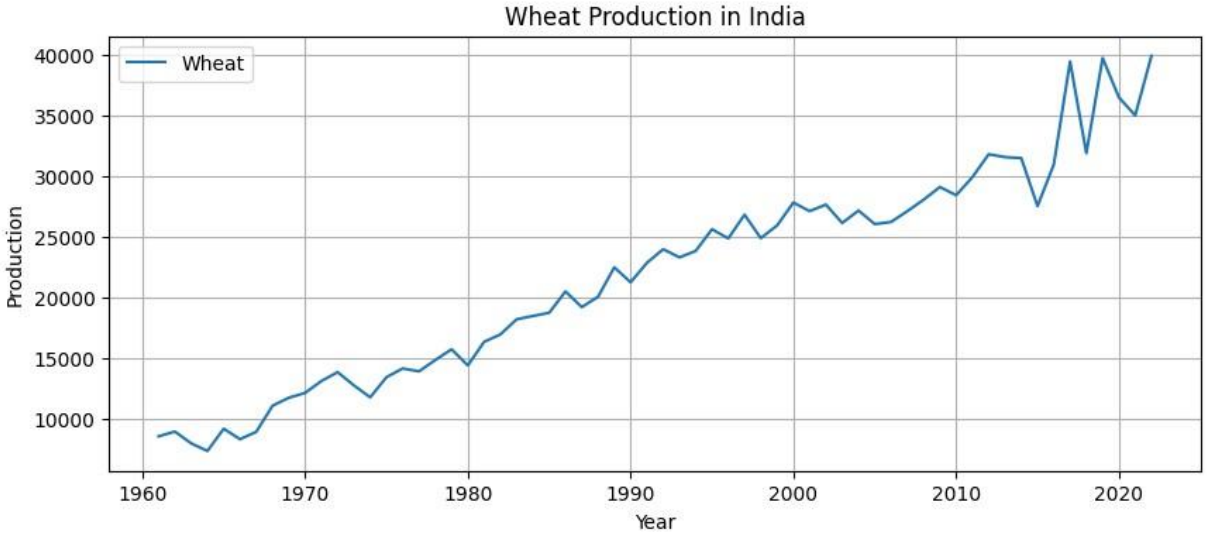

Fig 12: Wheat Crop Production in India

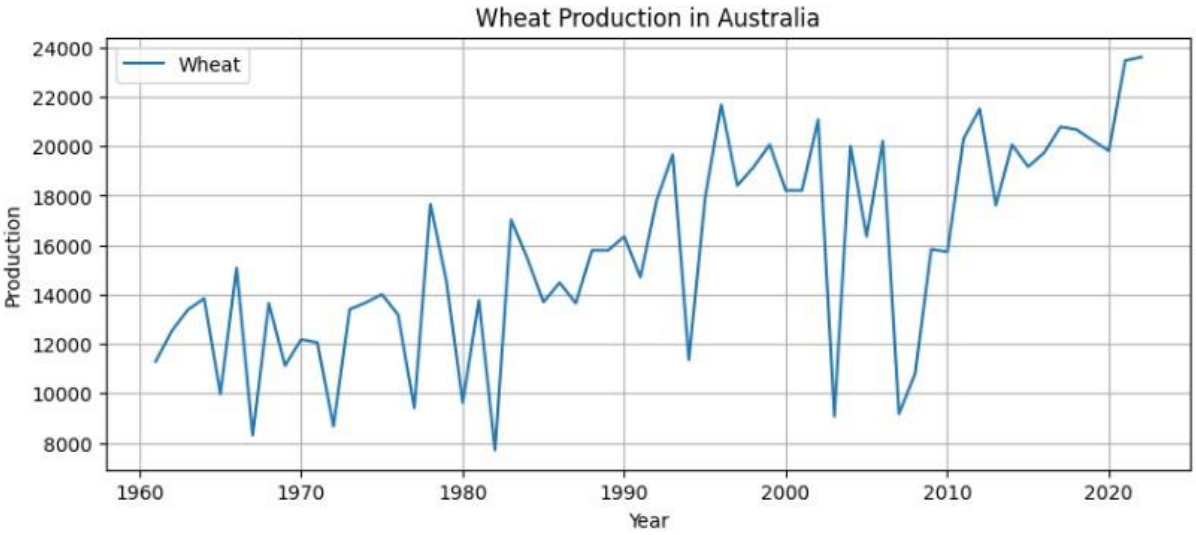

Fig 13: Wheat Crop Production in Australia

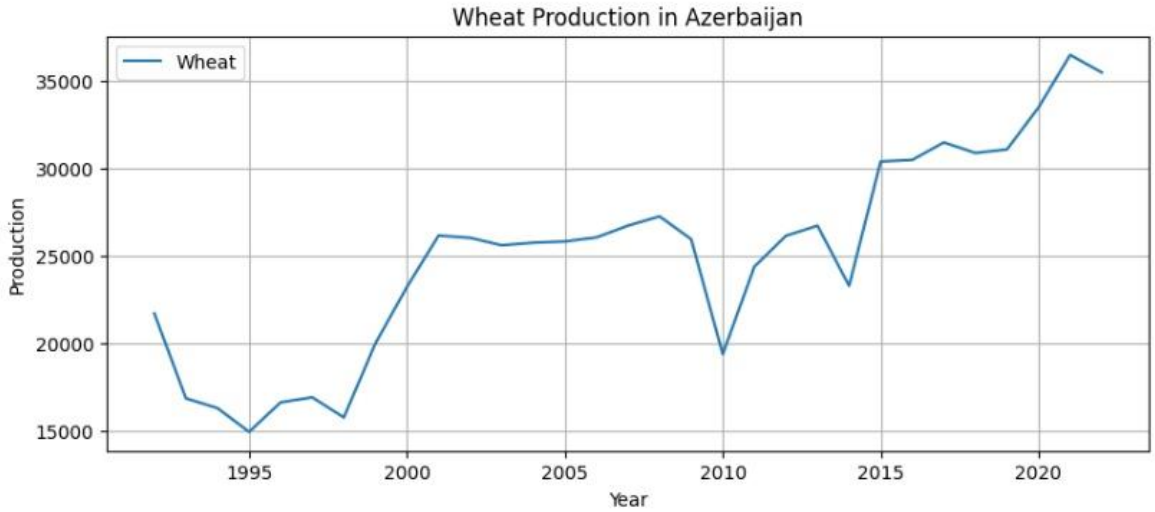

Fig 14: Wheat Crop Production in Azerbaijan.

Wheat production seems to be quite impressive in India over most of the period. On the one hand Australia not had been a notable country as it experienced drastic falls and not a significant growth for constant. Other hand Azerbaijan has experienced both downfalls, grown up, as it produced more than Australia, on average around 22000 crops been produced.

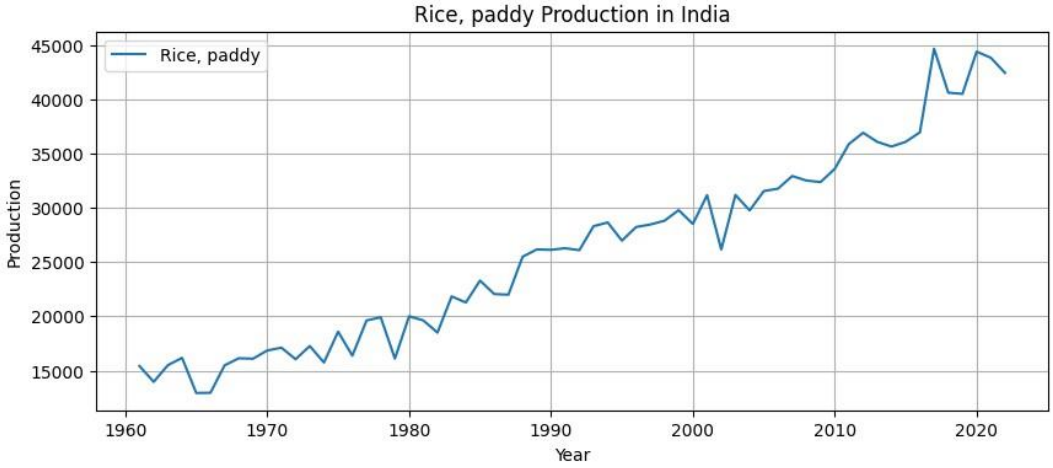

Fig 15: Crop Production of Rice and Paddy in India

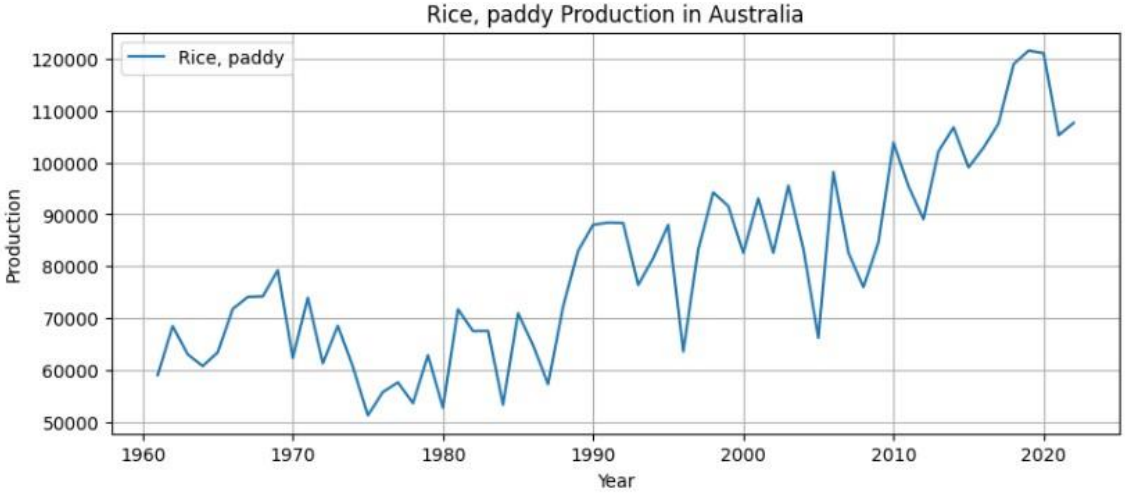

Fig 16: Crop Production of Rice and Paddy in Australia

Fig 17: Crop Production of Rice and Paddy in Azerbaijan

Rice and paddy have been the most talk about crop for India as it seems an enormous growth of production. Where Australia has experienced some growth, though some falls has been noticed constantly. Even though at the end of the scale, Australia has recorded the highest production touching 120000. On the other hand, Azerbaijan has drastic downfall for thought as it peaked in the year 2000 and from there it didn't reach there once again.

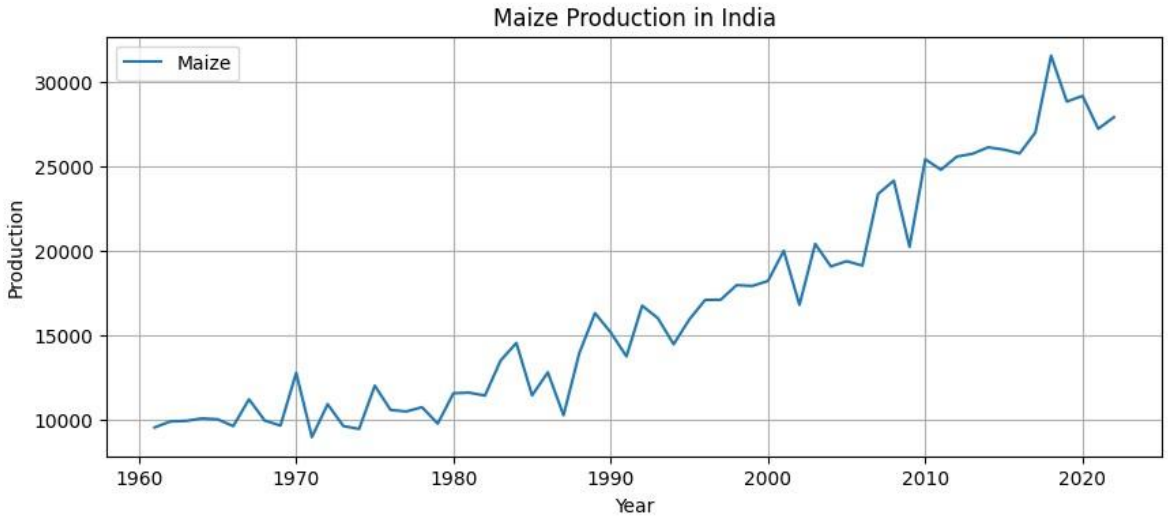

Fig 18: Crop Production of Maize in India

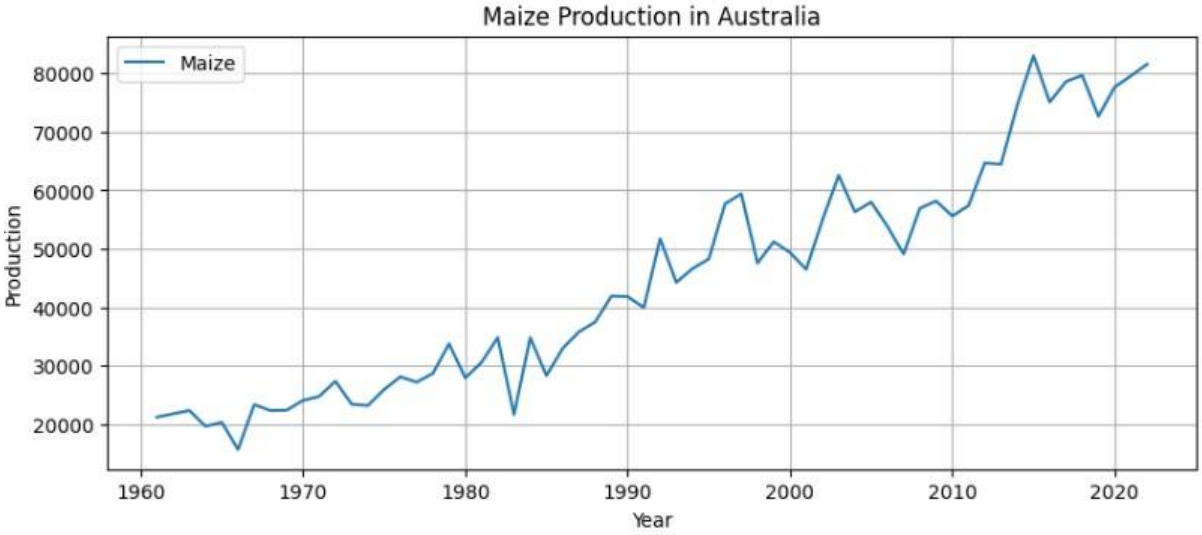

Fig 19: Crop Production of Maize in Australia

Fig 20: Crop Production of Maize in Azerbaijan

As, we Compared Maize in 3 Countries, we can conclude that Australia holds the top position compared to other's. We can see the significant grow on producing the Maize right from the year 1991.On the other hand India too have the upper hand but not able to reach too highest. Azerbaijan having the comfortable grown but stands at the low compared to Australia, India.

REFERENCES

[1] C. A. Reynolds,M. Yitayew,D. C. Slack,C. F. Hutchinson,A. Huete &M. S. Petersen. Estimating crop yields and production by integrating the FAO Crop Specific Water Balance model with real-time satellite data and ground-based ancillary data. https://doi.org/10.1080/014311600750037516

[2] Mamunur Rashid; Bifta Sama Bari; Yusri Yusup; Mohamad Anuar Kamaruddin; Nuzhat Khan. A Comprehensive Review of Crop Yield Prediction Using Machine Learning Approaches With Special Emphasis on Palm Oil Yield Prediction. https://ieeexplore.ieee.org/document/9410627
[3] Anna Mateo-Sanchis; Jose E. Adsuara; Maria Piles; Jordi Munoz-Marí; Adrian Perez-Suay; Gustau Camps-Val. Interpretable Long Short-Term Memory Networks for Crop Yield Estimation.https://ieeexplore.ieee.org/document/10041987
[4] Dhivya Elavarasan; P. M. Durairaj Vincent. Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications. https://ieeexplore.ieee.org/document/9086620
[5] Mengjia Qiao; Xiaohui He; Xijie Cheng; Panle Li; Haotian Luo; Zhihui Tian; Hengliang Guo. Exploiting Hierarchical Features for Crop Yield Prediction Based on 3-D Convolutional Neural Networks and Multikernel Gaussian Process. https://ieeexplore.ieee.org/document/9404810
[6] Mr.Sandeep Kumar Hegde, Rajalaxmi Hegde,Thanmaya.S, Kiran.N, Vivek Kumar. CROP ANALYSIS AND PREDICTION USING DATA MINING TECHNIQUES. https://www.ijrar.org/papers/IJRAR19H1149.pdf
[7] Ekaansh Khosla, Ramesh Dharavath & Rashmi Priya. Crop yield prediction using aggregated rainfall-based modular artificial neural networks and support vector regression. https://link.springer.com/article/10.1007/s10668-019-00445-x
[8] Usharani Bhimavarapu ,Gopi Battineni and Nalini Chintalapudi. Improved Optimization Algorithm in LSTM to Predict Crop Yield. https://www.mdpi.com/2073-431X/12/1/10
[9] Khaki, S.; Wang, L. Crop Yield Prediction Using Deep Neural Networks. https://www.frontiersin.org/articles/10.3389/fpls.2019.00621/full
[10] Gholipoor, M.; Nadali, F. Fruit yield prediction of pepper using artificial neural network. https://www.sciencedirect.com/science/article/abs/pii/S0304423819301177
[11] Uferah Shafi, Rafia Mumtaz, Zahid Anwar, Muhammad Muzyyab Ajmal, Muhammad Ajmal Khan. Tackling Food Insecurity Using Remote Sensing and Machine Learning-Based Crop Yield Prediction. https://ieeexplore.ieee.org/document/10267965
[12] Yuchi Ma, Zhou Zhang. A Bayesian Domain Adversarial Neural Network for Corn Yield Prediction. https://ieeexplore.ieee.org/document/9906947
[13] Seyed Mahdi Mirhoseini Nejad, Dariush Abbasi-Moghadam, Alireza Sharifi, Nizom Farmonov, Khilola Amankulov. Multi spectral Crop Yield Prediction Using 3D-Convolutional Neural Networks and Attention Convolutional LSTM Approaches. https://ieeexplore.ieee.org/document/9956008
[14] Mengjia Qiao, Xiaohui He, Xijie Cheng, Panle Li, Haotian Luo, Zhihui Tian, Hengliang Guo. Exploiting Hierarchical Features for Crop Yield Prediction Based on 3-D Convolutional Neural Networks and Multi kernel Gaussian Process. https://ieeexplore.ieee.org/document/9404810
[15] Niketa Gandhi; Leisa J. Armstrong; Owaiz Petkar; Amiya Kumar Tripathy. Rice crop yield prediction in India using support vector machines. https://ieeexplore.ieee.org/document/7748856
[16] Niketa Gandhi; Owaiz Petkar; Leisa J. Armstrong. Rice crop yield prediction using artificial neural networks.https://ieeexplore.ieee.org/document/7801222
[17] Ratchaphum Jaikla; Sansanee Auephanwiriyakul; Attachai Jintrawet. Rice yield prediction using a Support Vector Regression method. https://ieeexplore.ieee.org/document/4600365
[18] Ruchi Verma. Crop Analysis and Prediction. https://ieeexplore.ieee.org/document/10029218
[19] S. Thirumal; R. Latha. Automated Rice Crop Yield Prediction using Sine Cosine Algorithm with Weighted Regularized Extreme Learning Machine. https://ieeexplore.ieee.org/document/10142403