International Journal of Multidisciplinary Research and Growth Evaluation
ISSN: 2582-7138
Received: 01-06-2021; Accepted: 15-06-2021
www.allmultidisciplinaryjournal.com
Volume 2; Issue 4; July-August 2021; Page No. 230-233

Crop yield prediction based on Indian agriculture using machine learning

Ashwin Kowshik [1], Kishor Gowda HK [2], Rithik Somesh BR [3], Yashas S [4], Dr. Ramesh B [5], Nithyashree R [6]
[1-6] Department of Computer Science and Engineering, Malnad College of Engineering, Hassan, India

Corresponding Author: **Ashwin Kowshik**

Abstract
India being an agriculture country, its economy predominantly depends on agriculture yield growth and agro industry products. Data Mining is an emerging research field in crop yield analysis. Yield prediction is a very important issue in agriculture. Any farmer is interested in knowing how much yield he is about to expect. Analyze the various related attributes like location, pH value from which alkalinity of the soil is determined. Along with it, percentage of nutrients like Nitrogen (N), Phosphorous (P), and Potassium (K) Location is used along with the use of third-party applications like APIs for weather and temperature, type of soil, nutrient value of the soil in that region, amount of rainfall in the region, soil composition can be determined. All these attributes of data will be analyzed, train the data with various suitable machine learning algorithms for creating a model.
The system comes with a model to be precise and accurate in predicting crop yield and deliver the end user with proper recommendations about required fertilizer ratio based on atmospheric and soil parameters of the land which enhance to increase the crop yield and increase farmer revenue.

Keywords: Indian Agriculture, Machine Learning, Rice Crop, Vegetation Indices

1. Introduction
Agriculture is the most important sector of Indian Economy. Indian agriculture sector accounts for 18 percent of India's GDP and provides employment to 50% of the country's workforce. But latest studies have shown a steady decline in the contribution made by agriculture to the Indian economy although it is demographically the broadest economic sector and plays a significant role in the overall socio-economic fabric of India.
The Data Analysis is process of inspecting cleansing, modelling data with the goal of discovering useful information and conclusions. It is a process of analyzing, extracting and predicting the meaningful information from huge data to extract some pattern. This process is used by companies to turn the raw data of their customer to useful information. This analysis can also be used in the field of Agriculture. Most farmers were relied on their long-terms experiences in the field on particular crops to expect a higher yield in the next harvesting period but still they don't get worth price of the crops. It is mostly happening due to improper irrigation or inappropriate crops selection or also sometimes the crop yield is less than that of expected. Agricultural researchers insist on the need for an efficient mechanism to predict and improve the crop growth and Majority of research works in agriculture focus on biological mechanisms to identify crop growth and improve its yield. The outcome of crop yield primarily depends on parameters such as variety of crop, seed type and environmental parameters such as sunlight (Temperature), soil (ph.), water (ph.), rainfall and humidity. By analyzing the soil and atmosphere at particular region best crop in order to have more crop yield and the net crop yield can be predicted. This prediction will help the farmers. To choose appropriate crops for their farm according to the soil type, temperature, humidity, water level, spacing depth, soil PH, season, fertilizer and months.
Agriculture is the most important sector of Indian Economy. Indian agriculture sector accounts for 18 percent of India's GDP and provides employment to 50% of the country's workforce. But latest studies have shown a steady decline in the contribution made by agriculture to the Indian economy although it is demographically the broadest economic sector and plays a significant role in the overall socio-economic fabric of India. Smart agriculture is the way of conveying information from traditional farmers to the educated farmers. To obtain estimates of aggregate physical production functions for the yields of various crops in specified states, considering various technological factors and a newly developed weather index as inputs. Regression and coefficient of determination analysis along with Average Error rate were carried out to make a decent comparison between our actual result which is called target and prediction model that is friendly interface for farmers, which gives the analysis of rice production based on available data.

Different Data mining techniques were used to predict the crop yield for maximizing the crop productivity Accurate and timely monitoring of agricultural crop conditions and estimating potential crop yields are essential processes for operational programs Because of the importance of predicting crop yield, the purpose of this study is to apply several forecasting methods for evaluating crop yield estimates in Ghana. Crop yield forecasting, which provides information for decision Makers.

2. Literature review

Literature survey describes about the existing work on the given project. It deals with the problem associated with the existing system and also gives user a clear knowledge on how to deal with the existing problems and how to provide solution to the existing problems.

2.1 Rice yield estimation at pixel scale using relative vegetation indices from unmanned aerial systems.

Timely and accurate prediction of rice yield information is closely related to the people's livelihood, which has been attached great importance by all levels of government. Satellite remote sensing provides the possibility for large-scale crop yield estimation, but they are usually limited by spatial and spectral resolution. Unmanned Aerial Vehicles (UAV) remote sensing with hyperspectral sensors can obtain high spatial temporal resolution and hyperspectral images on demand. Generally, time-series Vegetation Indices (VIs) are used for estimating grain yield. But multi-day vegetation indices may be affected by different background and illumination condition, so the differences between vegetation indices may include the effects induced from external condition, which will pose a negative effect on the accuracy of crop yield estimation. Therefore, in this study, the relative vegetation index and relative yield were proposed and used to estimate rice yield at pixel scale. And the optimal growth stages for crop yield estimation would also be determined. Hyperspectral images of critical rice growth stages at tillering stage, jointing stage, booting stage, heading stage, filling stage, ripening stage were obtained from July 28 to November 24 in 2017.

2.2 Use of deep neural networks for crop yield prediction: A case study of soybean yield in Lauderdale County, Alabama, USA

World population is constantly increasing, and it is necessary to have sufficient crop production. Monitoring crop growth and yield estimation are very important for the economic development of a nation. The prediction of crop yield has direct impact on national and international economies and play important role in the food management and food security. Deep learning gains importance on crop monitoring, crop type classification and crop yield estimation applications with the recent advances in image classification using deep Convolutional Neural Networks. Traditional crop yield prediction approaches based on remote sensing consist of classical Machine Learning methods such as Support Vector Machines and Decision Trees. Convolutional Neural Network (CNN] and Long-Short Term Memory Network (LSTM] are deep neural network models that are proposed for crop yield prediction recently. This study focused on soybean yield prediction of Lauderdale County, Alabama, USA using 3D CNN model that leverages the spatiotemporal features. The yield is provided from USDA NASS Quick Stat tool for years 2003-2016

2.3 Prediction of major crop yields of Tamilnadu using K-means and Modified KNN

Agriculture is the principal source of livelihood for more than 40 percent of the population of this state. According to Food and Agricultural Organization (FAO) researchers, between 2010 and 2050 the world population will increase by one third. The demand for crop production will increase by 60% higher than the current production. Hence prediction plays a major role to find out the demand of crop production for maximizing the yield. For that in this paper we propose a prediction method for the major crops of Tamil Nadu using K-means and Modified K Nearest Neighbor (KNN). MATLAB and WEKA are used as the tool for clustering and classification respectively. The number result shows that our method is better than traditional data mining approach.

2.4 Rice crop yield prediction in India using support vector machines

The sustainability and productivity of rice growing areas is dependent on suitable climatic conditions. Developing better techniques to predict crop productivity in different climatic conditions can assist farmer and other stakeholders in better decision making in terms of agronomy and crop choice. This paper presents the review on use of such machine learning technique for Indian rice cropping areasThis paper discusses the experimental results obtained by applying SMO classifier using the WEKA tool on the dataset of 27 districts of Maharashtra state, India. The dataset considered for the rice crop yield prediction was sourced from publicly available Indian Government records. The parameters considered for the study were precipitation, minimum temperature, average temperature, maximum temperature and reference crop evapotranspiration, area, production and yield for the Kharif season (June to November) for the years 1998 to 2002. For the present study the mean absolute error (MAE), root mean squared error (RMSE), relative absolute error (RAE) and root relative squared error (RRSE) were calculated. The experimental results showed that the performance of other techniques on the same dataset was much better compared to SMO.

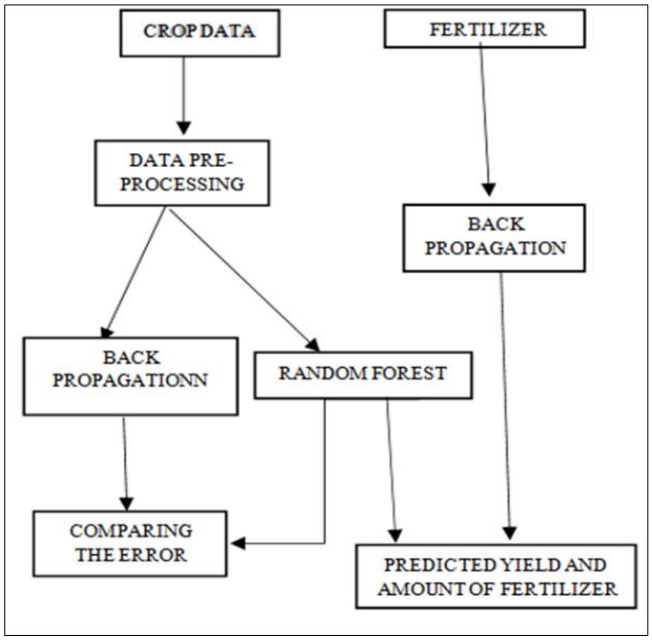
3. Problem statement

The production of agriculture is affected by several climate factors. Like as metrological parameters (Humidity, wind speed, temperature, and moisture), precipitation parameters (rainfall, region wise rainfall, irrigation etc.), and soil parameters (PH, organic carbon, phosphorus, fiber etc.). And due to continuously change in climate condition everything is messed.

In India farmers still follow the traditional technology which they adopted from their ancestor. But the problem is that in earliest time climate was very healthy everything was happened on time. But now most of the things have been changed due to global warming and many other factors. The main problem with agriculture in India is lack of rainfall in seasonal time. Humidity is also necessary for crops but it has been excessive, it also converts as drawback. Winter season is been affected so Rabi crops are widely affected. Since few years the rainfall in winter season was high as expected.

To overcome these above issues we need to develop a system which will able to find the hidden facts or results, patterns and insights. The farmer can predict which crop should sow so that he/she can get more benefit. In proposed system we are applying data analytics techniques on agriculture production based datasets and find the insights so that it can help to the farmers and their decision making.

4. Methodology



4.1 Crop yield prediction

The outcome of crop yield primarily depends on parameters such as variety of crop, seed type and environmental parameters such as sunlight (Temperature), soil (ph), water (ph), rainfall and humidity. By analyzing the soil and atmosphere at particular region best crop in order to have more crop yield and the net crop yield can be predicted. This prediction will help the farmers. To choose appropriate crops for their farm according to the soil type, temperature, humidity, water level, spacing depth, soil PH, season, fertilizer and months.

4.2 Fertilizer prediction

India is a highly populated country and randomly change in the climatic conditions need to secure the world food resources. Framers face serious problems in drought conditions. Type of soil plays a major role in the crop yield. Suggesting the use of fertilizers may help the farmers to make the best decision for their cropping situation.

Based on soil type and soil PH we suggest what kind of fertilizer should be used for particular crop.

4.3 Random forest algorithm

Random Forest algorithm is a supervised classification algorithm. We can see it from its name, which is to create a forest by some way and make it random. There is a direct relationship between the number of trees in the forest and the results it can get: the larger the number of trees, the more accurate the result. But one thing to note is that creating the forest is not the same as constructing the decision with information gain or gain index approach.

5. Abbreviations

UAV - Unmanned Aerial Vehicles

CNN - Convolutional Neural Network

LSTM - Long-Short Term Memory Network

FAO - Food and Agricultural Organization

KNN - K Nearest Neighbor

MAE - Mean Absolute Error

RMSE - Root Mean Squared Error

RAE - Relative Absolute Error

RRSE - Root Relative Squared Error

6. Comparison (with result/ performance of proposed work with existing methods)

Prediction of the crop yield using the efficient algorithm and suggest how much quantity of fertilizer should be used to get the proper yield for the crop using naïve Bayesian algorithm. The data mining techniques on historical climate and crop production data several predictions are made which increase the crop productivity. The decision support system must be implemented for the farmers to take proper decisions about soil and crop to be cultivated. They have collected the dataset with attributes of the crop season, Area and production in hectares and analyzed with various algorithms in WEKA.

7. Conclusion and future enhancement

Crop Yield Prediction and efficient utilization of fertilizers is victoriously anticipated and effective algorithm is also established from both the algorithm and acquired the great efficient result of the harvest.

Provided the farmer with the yield of a crop based on land area, rainfall, temperature and district using machine learning. Predicted the future market price of crops by taking previous crop price and predicted yield data into consideration. Predicted the crop price using multiple linear regression and random forest.

8. References

1. Niketa Gandhi *et al*. Rice Crop yield forecasting of tropical wet and dry climatic zone of India using data mining techniques. IEEE International Conference on Advances in Computer Applications (ICACA), 2016.
2. Eswari KE, Vinitha L. Crop yield prediction in Tamil Nadu using bayesian network. International Journal of Intellectual Advancements and Research in Engineering Computations, 6(2). ISSN: 2348- 2079.
3. Shruti Mishra, Priyanka Paygude, Snehal Chaudhary, Sonali Idate. Use of data mining in crop yield prediction. IEEE Xplore Compliant-Part Number: CFP18J06-ART. ISBN: 978-1-5386-0807-4, DVD Part Number: CFP18J06DVD, ISBN: 978-1-5386-0806-7.
4. Anna Chlingaryana, Salah Sukkarieha, Brett Whelanb. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review, Computers and Electronics in Agriculture, Elisver. 2018; 151:61-69.
5. Dakshayini Patil *et al*. Rice Crop Yield Prediction using Data Mining Techniques: An Overview. International Journal of Advanced Research in Computer Science and Software Engineering. 2017; 7(5).
6. Shakil Ahamed ATM, Navid Tanzeem Mahmood, Nazmul Hossain, Mohammad Tanzir Kabir, Kallal Das, Faridur Rahman, *et al*. Applying data mining techniques to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh. 978-1-4799-8676-7/15/$31.00 copyright 2015 IEEE SNPD, 2015, 1-3.
7. Menaka K, Yuvaraj N. Indian Journal of Innovations and Developments, 5(12).

8. Sekhar CC, Sekhar C. Productivity improvement in agriculture sector using big data tools. 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC), 2017.
9. Shah P, Hiremath D, Chaudhary S. Towards development of spark based agricultural information system including geo-spatial data. 2017 IEEE International Conference on Big Data (Big Data), 2017.
10. Raja SKS, Rishi R, Sundaresan E, Srijit V. Demand based crop recommender system for farmers. 2017 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR), 2017.