IJIREEICE

ISSN (Online) 2321-2004
ISSN (Print) 2321-5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

Vol. 8, Issue 4, April 2020

Time Series Prediction of Crop Yield based on Time Alignment using Machine Learning

Manjula H.N[1], Namratha S[2], Shiva Prasad K[3]

Assistant Professor, Department of Information Science & Engineering, Atria Institute of Technology, Bangalore,India[1]

Student Department of Information Science & Engineering, Atria Institute of Technology, Bangalore, India[2,3]

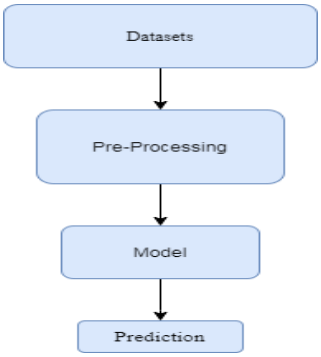
Abstract: Food production in India is largely dependent on cereal crops including rice, wheat and various pulses.The sustainability and productivity of crops growing areas is dependent on suitable climatic conditions.Variability in seasonal climate conditions can have detrimental effects, with incidents of drought reducing production. Developing better techniques to predict crop productivity in different climatic conditions can assist farmers in better decision making in terms of agronomy and crop choice. Machine learning techniques can be used to improve prediction of crop yield under different climatic scenarios. This paper presents the review on use of such machine learning techniques for cropping areas.

Keywords: Machine Learning, Linear Regression, Random Forest Regression, Prediction.

I. INTRODUCTION

With an increasing world population and changing climate, has come the necessity to secure the world food resources. Farmers are faced with having to make difficult decisions as to how to remain productive and sustainable with changing climates and market economic pressure.The provision of accurate and timely information such as meteorological, soil, use of fertilizers, use of pesticides can assist farmers to make the best decision for their cropping situations.This could benefit them to attain greater crop productivity if the conditions are suitable or help them to decrease the loss due to unsuitable conditions for the crop yield.The current study used a dataset from various places of India.Various climatic factors which are known to affect the crop yield, such as precipitation, minimum temperature, average temperature, maximum temperature, reference crop evapotranspiration, were considered with the crop yield production for the Kharif season for the 10 years.

II. STEPS FOLLOWED IN PREDICTION MODEL



The above figure shows the steps followed in the prediction model.The dataset fetched is pre-processed. The pre-processed data is fed into the model where various Regression algorithms are applied to obtain the predictions. The accuracy of prediction can vary based on various factors such as data fed into the model, the algorithm used as well as pre-processing techniques.

III. RELATED SURVEY

The paper[1] provides an idea of how to uncover additional perspectives from data on precision farming through a Big Data approach. It presents a scenario for using Information and Communication Technology (ICT) services to gather huge data in the agricultural big data environment. Big data analytics in farming applications offer a new perspective to advance weather decisions and increase yields.

IJIREEICE

ISSN (Online) 2321-2004
ISSN (Print) 2321-5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

Vol. 8, Issue 4, April 2020

In paper[2] data mining techniques are used to classify soil into low, medium and high categories in order to predict crop yield using the available dataset. This study will help soil analysts and farmers determine which sowing may lead to better crop production in which land.

Remote soil parameter monitoring is an emerging trend which has the potential to transform farming practices and increase productivity. The pH value, the temperature and the moisture content of the soil are the basic parameters which help to characterize the soil and thus to make appropriate decisions regarding the application of fertilizers and the selection of sown crops. In this work, an antimony electrode is used for pH measurement. For soil moisture content estimation, the inverse relation between soil resistance and soil moisture has been utilized and corresponding circuitry has been developed.[3]

The project in paper[4] includes various features such as remote controlled monitoring based on GPS, sensing of humidity & temperature, scarring of intruders, protection, wetness of the leaf and proper irrigation. It uses wireless sensor networks to continuously note soil properties and environmental factors. Different sensor nodes are installed inside the farm at various locations.

In the paper[5] certain Data Mining techniques were adopted in order to estimate crop yield analysis with existing data. The applications that use the K-Means approach , utilize only the basic algorithm, while many other improvements are available. Some Data Mining techniques have not yet been applied to agricultural problems.

One of the most beneficial aspects in [5] is that there is no cost to this particular device, and it can be mounted on any smartphone, which makes this device more user-friendly. Such systems are also open to farmers, and thus they will be properly informed about which crop to grow and which not to grow.

This proposed system uses the Artificial Neural Network[6], which is one of the most effective tools in modeling and prediction. There are various parameters that decide the crop productivity. Hence these parameters are considered as the input for the proposed system and based on the manipulation with these inputs, the desired output must be produced. The parameters include pH, phosphate, potassium, nitrogen, depth, temperature, rainfall.

IV. PROPOSED METHOD

Linear Regression is a machine learning algorithm based on supervised learning. It performs a regression task. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting. Different regression models differ based on – the kind of relationship between dependent and independent variables, they are considering and the number of independent variables being used.

Hypothesis function for Linear Regression :

$$y = \theta_1 + \theta_2 . x$$

While training the model we are given:
x: input training data (univariate – one input variable(parameter))
y: labels to data (supervised learning)

When training the model – it fits the best line to predict the value of y for a given value of x. The model gets the best regression fit line by finding the best θ_1 and θ_2 values.
θ_1: intercept
θ_2: coefficient of x
Once we find the best θ_1 and θ_2 values, we get the best fit line. So when we are finally using our model for prediction, it will predict the value of y for the input value of x.

A Random Forest is an ensemble technique capable of performing both regression and classification tasks with the use of multiple decision trees and a technique called Bootstrap Aggregation, commonly known as bagging. What is bagging you may ask? Bagging, in the Random Forest method, involves training each decision tree on a different data sample where sampling is done with replacement.

From a computational standpoint, Random Forests are appealing because they
• naturally handle both regression and classification.
• are relatively fast to train and to predict.
• depend only on one or two tuning parameters.

IJIREEICE

ISSN (Online) 2321-2004
ISSN (Print) 2321-5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

Vol. 8, Issue 4, April 2020

• have a built-in estimate of generalization error.

As the first step of the crop prediction, we show the results using only data information. In addition to crop information, we plan to reflect other related characteristics such as meteorological information, preference estimation by data analysis.

The framework is designed to reflect this information into learning in future. We focus on representative types of agricultural crops, especially some types of crops that are indispensable and popular in daily life. The selection of target agricultural products varies according to measurement periods. The most representative products are chosen and have been surveyed according to the specific periods.

One of the most popular and well-known methods to predict some success rate in the future. A new reinforcement learning based prediction algorithm has been proposed and showed good performances in prediction at a specific point in the future. Although the accuracy of a predicted value at a specific time point has been surveyed. One of our main concerns is to know how these kinds of non-linear mechanisms have an effect on the accuracy of prediction and statistical characteristics of yield.

A predictor that is trained by the past data and provides a success value at a specific time point, that agricultural products trade seems surely to have seasonal variations, i.e., weather movement behaves according to the season and it shows one-year periodical behaviour. Therefore, by setting the length of the past sequence as the value, we can expect that the basic predictor can detect these characteristics in the sequence. In order to predict crop in addition to predict a success rate at a specific point in the future, we use generation and filter methods by combination of basic predictors.

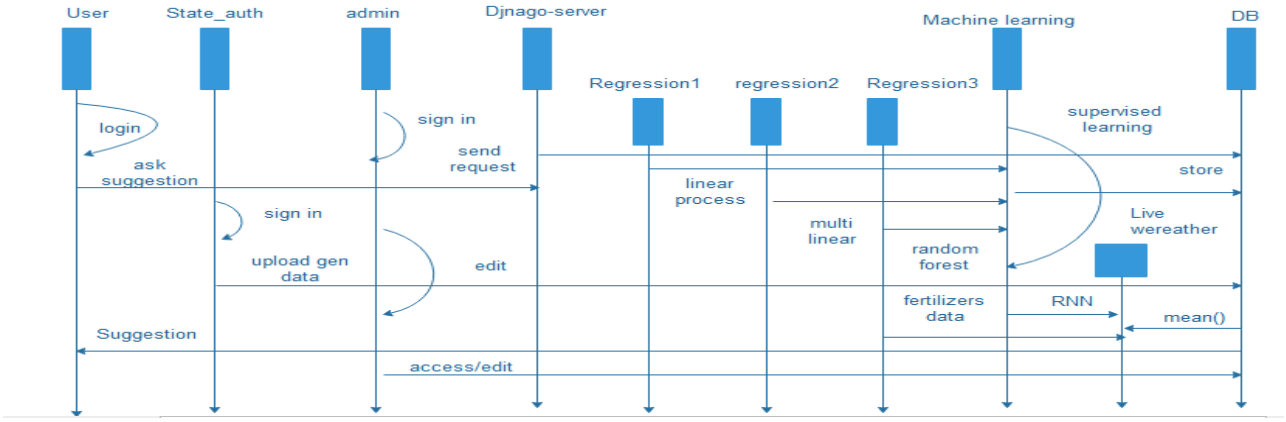


Fig.1 Sequence Diagram

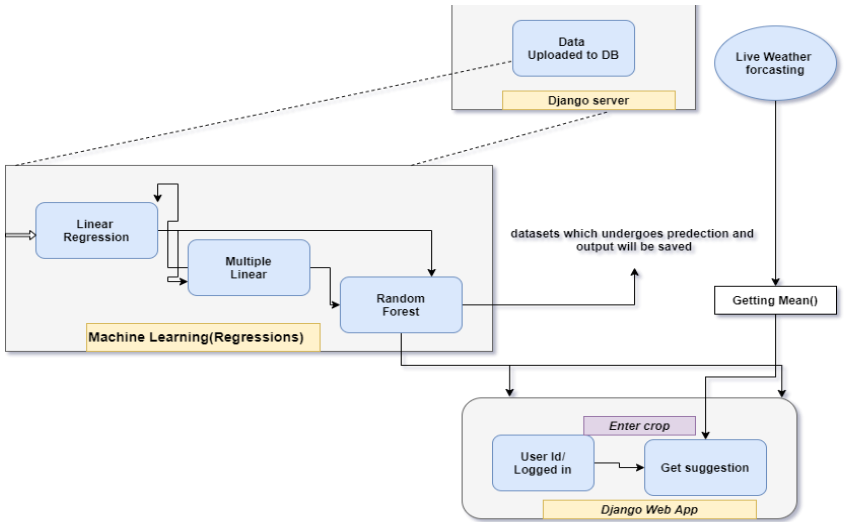


Fig. 2 Architecture Diagram

IJIREEICE

ISSN (Online) 2321-2004
ISSN (Print) 2321-5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

Vol. 8, Issue 4, April 2020

V. CONCLUSION

From the information acquired by referring to numerous project reports and survey papers belonging to the Crop yield prediction, we conclude that the concept of supervised learning methodology is better suited for training the concept when there is vast availability of unlabeled data with few labeled data. This also fits for models that are not meant to be unique to a domain.It can be concluded that other classifiers used on the current study dataset and reported earlier should be recommended for further development of a crop prediction model.

REFERENCES

[1]. M.R. Bendre, R.C. Thool, V.R.Thool,.Big data in precision agriculture: Weather forecasting for future farming,year,2015
[2]. Monali Paul, Santosh K. Vishwakarma, Ashok Verma,Analysis of Soil Behaviour and Prediction of Crop Yield Using Data Mining Approach,year,2015
[3]. Analysis of Soil Behaviour and Prediction of Crop Yield Using Data Mining Approach,An IoT Based System for Remote Monitoring of Soil Characteristics,tear 2016
[4]. Dr.N.Suma, Sandra Rhea Samson, S.Saranya, G.Shanmugapriya, R.Subhashri,IOT Based Smart Agriculture Monitoring System,year 2017
[5]. Dhivya B ,Manjula , Siva Bharathi, Madhumathi,A Survey on Crop Yield Prediction based on Agricultural Data,year 2017
[6]. GiritharanRavichandran, ,Koteeshwari R S,Agricultural Crop Predictor and Advisor using ANN for Smartphones,year 2016