

Crop Prediction Based on Artificial Intelligence and Machine Learning

Jigyasa Yadav¹, Etishree Thakre¹, Rani Mohadikar¹, Md. Mumtaj¹, Kapil K. Wankhade¹

¹Department of Computer Science and Engineering, Jhulelal Institute of Technology,
Nagpur, MS, India

Abstract:

Agriculture is one of the major revenues producing sectors of India and a source of survival. The kind of yield desire is the hardest task for the cultivating space. Artificial Intelligence (AI) and Machine Learning (ML) are the recent technologies from the last few years. Machine Learning can acquire a blast in the farming field by changing the pay situation through developing the ideal yield. This paper focuses on predicting the yield of the crop by applying Deep neural network technique. The outcome of these techniques is compared on the basis of mean absolute error. The prediction made by machine learning algorithms will help the farmers to decide which crop to grow to get the maximum yield by considering factors like soil temperature, moisture, area groundwater level, etc.

Keywords:

Agriculture, Artificial Intelligence, Machine Learning, Crop prediction, Deep Neural Network.

I. Introduction:

Farmer's in India play a major role to feed the growing population of the country. This makes crop examination and expectation as significant as yield creation. Nourishment crops in India are for the most part occasional and require observing of soil parameters like the measure of Moisture, Temperature, Humidity, Type of soil, Area groundwater level. Rainfall and Temperature are the two environmental factors that impact the crop yield.

Estimating agriculture yield before harvest is an important issue in agriculture, as the changes in crop yield from year to year influence international business, food supply, and global market prices. Also, the early prediction of crop yield provides useful information to policy planners. Appropriate prediction of crop productivity is required efficient planning of land usage at economic. In recent times, the forecasted of productivity at the within-field level has increased. The most influencing factor for crop productivity is whether conditions. If whether based on predictions is more precise, then farmers can be alerted well in advance so that the major loss can be mitigated and would be helpful for economic growth. The prediction will also aid the farmers to make the decision such as the choice of alternative crops or to discard a crop at an early stage in case of a critical situation. Further, predicting crop yield can facilitate the farmers to have a better vision of the cultivation of seasonal crops and their scheduling. Thus, it is necessary to simulate and predict the crop yield before cultivation for efficient crop management and expected outcomes as there exists a non-linear relationship between crop yield at the factors influencing crop, machine learning techniques might be efficient for yield production.

Crop yield prediction is an important agricultural problem. Each and Every farmer is always trying to know, how much yield will get from his expectation. In the past, yield prediction was calculated by analyzing the farmer's previous experience on a particular crop. The Agricultural yield fundamentally relies upon climate conditions, bugs, and arranging of gather activity. Accurate information about the history of crop yield is an important thing for making decisions related to agricultural risk management. This research focuses on the evolution of a prediction model that may be used to predict crop yield

production. The proposed method uses Artificial Intelligence and Machine Learning techniques to predict crop.

a) Artificial Intelligence:

Artificial intelligence is a technology that connected physical objects in the world that are accessible through the internet. Applications are developed based on Artificial Intelligence devices for monitoring and controlling in various domains including home appliances, smart cities, smart homes, and so on. In the agriculture domain, few kinds of research have proposed architectures based on AI with machine learning to predict the type of crops and small plants. AI with machine learning is a mature technology and a lot of work has been done for the agriculture domain. AI with machine learning architecture was proposed for predicting the types of crops in the agriculture domain [1].

b) Machine Learning:

Machine learning is a new technology and has been used for analyzing the soil moisture, temperature, and humidity of the environment, pH value to classify them. It is useful for the type of crops and small plant predictions based on certain parameters like environmental conditions, the moisture of soil, pH level in farmland. Artificial Intelligence fundamentally utilized in farming space for kinds of harvests forecast utilizing altered help vector machine calculation [2].

c) Deep Neural Network:

A Deep Neural Network (DNN) is an Artificial Neural Network (ANN) with numerous layers between the information and yield layers. The DNN finds the right numerical control to transform the contribution to the yield, regardless of whether it be a straight relationship or a non-direct relationship. The system travels through the layers ascertaining the likelihood of each yield. For instance, a DNN that is prepared to perceive hound breeds will go over the given picture and compute the likelihood that the pooch in the picture is a sure variety. The client can survey the outcomes and select which probabilities the system should show (over a specific edge, and so forth.) and return the proposed mark. Each mathematical manipulation as such is considered a layer, and complex DNN has many layers, hence the name "deep" networks [3].

II. Literature Review:

Machine Learning (ML) deals with problems where the relationship between input and output variables is not known or hard to obtain. The "learning" term here denotes the automatic acquisition of structural descriptions from examples of what is being described. Unlike traditional statistical methods, ML does not make assumptions about the correct structure of the data model, which describes the data [4]. This trademark is helpful to display complex non-straight practices, for example, a capacity for crop yield expectation. ML methods most effectively applied to Crop Yield Prediction (CYP). Directed Learning calculation comprises an objective/result variable (or ward variable) which is to be anticipated from a given arrangement of indicators (autonomous factors). Using these sets of variables, we generate a function that map inputs to desired outputs. The training process continues until the model achieves a desired level of accuracy on the training data [5].

K-MEANS: K-means is a measurable, unsupervised, non-deterministic, iterative technique for gathering diverse articles into clusters. The k-means calculation has ended up being to a great degree intense in making clusters in various handy applications in creating regions, for instance, in Bioinformatics, advertising division, PC vision, geostatistics, stargazing, and cultivation. It is the least complex unsupervised learning calculations known for its speed, ease, and convenience [6].

KNN: K-Nearest Neighbours (KNN) algorithm is one of the simplest, easy to understand, versatile and one of the topmost machine learning algorithms. KNN is a non-parametric supervised learning algorithm. Additionally, it is an instance-based learning or a lazy algorithm. At the point when an inquiry to the database is made, the calculation utilizes the preparation examples to let out an answer. That is why for KNN the training phase is very fast compared to other classifier algorithms. However, the testing phase becomes slower and costlier, which is in terms of time and memory. KNN's motivation is to utilize a database where the information focuses are isolated into a few classes to foresee the order of another example point. Informally, this means that for a labelled dataset consisting of training observations (x, y) and the algorithm is used to capture the relationship between x and y . KNN Algorithm is based on feature similarity: How closely out-of-sample features resemble our training set determines how we classify a given data point [7].

APPRIORI ALGORITHM: Defines the parameter K Apriori is a calculation for incessant thing set mining and affiliation guideline learning over value-based databases. It continues by recognizing the continuous individual things in the database and extending them to bigger and bigger things sets as long as those things sets show up adequately regularly in the database [8]. The continuous thing sets dictated by Apriori can be utilized to decide affiliation rules which feature general patterns in the database: this has applications in area s, for example, showcase bushel examination.

LINEAR REGRESSION: Linear Regression model for crop yield prediction is to develop the Linear Regression models for crop yield prediction, Linear Regression analysis is majorly used for prediction functions because it provides predicted entity as a function of depended entities [9].

SVM: SVM creates a hyperplane or set of hyperplanes in a high- or infinite-dimensional space, which is utilized for regression, classification or other tasks. SVM uses linear functions for learning. In the case of nonlinear cases, SVM uses a kernel technique to plot the data into a higher dimensional feature space, in which linear functions can be applied. This strategy finds wide applications in signal preparation, time arrangement investigation, climate expectation, crop efficiency forecast, and so forth. SVM is less vulnerable to the overfitting problem as it is capable of mapping the functions to a hyperplane. A study on the impact of climate factors such as temperature, rainfall, relative humidity, sunshine hours, daily temperature range, and rainy days are considered to predict paddy yield is proposed. Further SVM Classifier is used to recognize the plant as either harvest or weed. The method is implemented to detect maize and weed plants. SVM-based Open Crop Model (SBOCM) for rice improvement stage and yield forecast is endeavoured Soil order to recognize reasonable harvest for the dirt utilizing SVM is proposed in the key issue in SVM displaying is the assurance of the part capacities, hyperparameters, and punishment coefficient. In [10], the application of SVM in analysing aerial hyperspectral observations taken over a cornfield to predict crop yield, biomass, plant height, and leaf greenness is proposed. The results were compared with a stepwise regression method and found to be better.

PSO: Particle Swarm Optimization (PSO) is an optimization method that is included in the family of Swarm Intelligence. Particle Swarm Optimization was a technique introduced by Eberhart and Kennedy that flocking birds. The PSO algorithm 28 groups each particle (point) with a certain speed of flight but without quality and size in the multidimensional search space. It is assumed that there is a group (swarm) consisting of m particles flying with a certain speed in a Dimensional search space, and each particle changes its location (i.e., state or solution) based on both its own best search point in the history of the search process and that of other particles in the swarm (or neighbourhood).

III. Proposed System:

This system proposes a method using of deep learning technique which analyses the dataset from the past data fed and classifies or predict the plants or crops. Deep learning concepts use in agriculture and the efforts that apply to execute deep learning techniques, in various agricultural sectors. Deep learning application is required in this field as it provides a major impact on the modern techniques, it extends the Machine learning by adding more depth into the model. The main feature of deep learning is the raw data process to increase accuracy and classification. Plant recognition, predicting future crop yield is the main target where deep learning can be implemented. A large dataset is required to train the model, while some techniques use text data to train the model. Deep learning in agriculture has many environments and it can proliferate the agriculture sector. We have worked with convolution neural network models and used deep learning in the system by training the model with data sets.

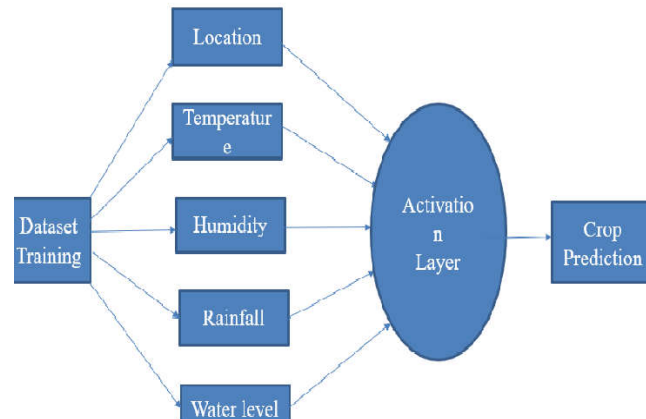


Fig. 1. Parameters for Crop Prediction

The proposed system consists of the 5 modules:

- 1) The creation of UI (User Interface) for taking an input.
- 2) The prediction based on KNN (K-Nearest Neighbour).
- 3) The fine prediction using CNN (Convolution Neural Network).
- 4) The real time input and output.
- 5) To optimization and result analysis.

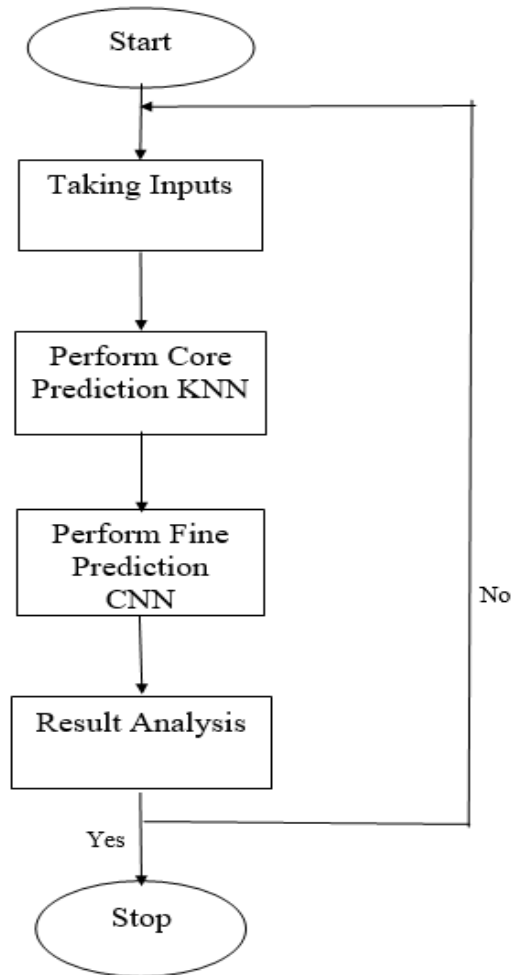


Fig. 2. Flow Chart

The prediction of the crop is dependent on numerous factors such as Soil Nutrients, weather, and past crop production to predict the crop accurately. The creation of UI (user interface) is used for taking input from the user according to the parameters/factors.

The k-nearest neighbors (KNN) calculation is a basic, simple to-actualize regulated AI calculation that can be utilized to tackle both characterization and relapse issues. The KNN classification method can be applied over the small datasets, it is analyzed that when the size of the dataset gets increased accuracy to get reduced. This classifier can be divided the data into two parts training data and testing data.



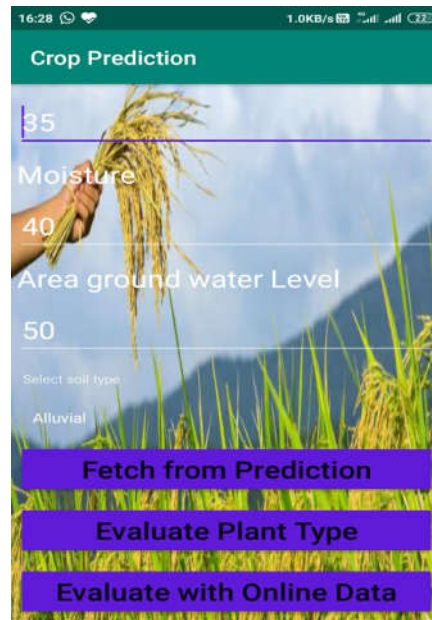
Fig. 3. User Interface

The Convolutional Neural Network (CNN) the developed agricultural management needs simple and accurate estimation techniques to predict yields in the planning process. The necessity of the present study is to identify whether convolutional neural network (CNN) models could effectively predict yield for typical climatic conditions of the mountainous region, evaluate CNN model performance relative to variations of developmental parameters and compare the effectiveness of multiple linear regression models with CNN models.

The core aim of our research is to establish a model that will efficiently predict a particular crop.

Following steps needs to be followed for opening and performing different operations:

- i. Open the mobile application.
- ii. Depending on the current user location, the system mines the soil properties in the respective area from the soil repository. In a comparative methodology, climate parameters are extricated from the climate informational collection.
- iii. A crop can be cultivable only if apropos conditions are met. These incorporate broad parameters that partnered with soil and climate. These constraints are compared and the apt crops are ascertained. Various Linear Regression is utilized by the framework to anticipate the harvest. The forecast depends on past creation information of harvests, for example, distinguishing the substantial climate and soil parameters and contrasting it and current conditions that will foresee the harvest all the more precisely and in a down to earth way.
- iv. Enter the values in various parameters like Soil Temperature, Moisture, Area groundwater level and also fetch the data from online according to the parameter.



v. Fig. 4. Taking Inputs

- vi. Snap-on the referenced tab to assess the plant type.
- vii. The most profitable crop is predicted by the system using the Multiple Linear Regression algorithm and the user is provided with multiple suggestions of crop conferring to the duration of the crop.

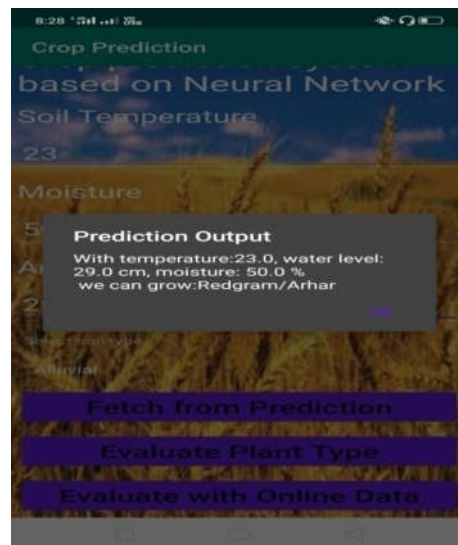


Fig. 5. Final Output

IV. Result:

As discussed in the related work number of algorithms have been implemented using K-means, Linear Regression, SVM, etc. In this paper, we have discussed the Deep neural networks and Convolutional neural network.

Table 1: Comparative Accuracy

Sr no.	Algorithms	Accuracy
1	K-means Algorithm	84%
2	Linear Regression Method	89%
3	Support Vector Machine (SVM)	80%
4	Convolution Neural Network (CNN)	93%

V. Conclusion:

Agriculture is the backbone and important sector of developing country's economies. It contributes to the development and is considered as an important source. Crop yield prediction is remaining as a challenging issue for farmers. The proposed system takes into consideration the data related to soil, weather and past year production and suggests which are the best profitable crops that can be cultivated in the appropriate environmental condition. Also, this RF system takes into consideration the past production of data which will help the farmer get insight into the demand and the cost of various crops in the market. As maximum types of crops will be covered under this system, a farmer may get to know about the crop which will help the farmer get insight into the demand and the cost of various crops in the market. As maximum types of crops will be covered under this system, a farmer may get to know about the crop which may never have been cultivated. This project aims to improve accuracy by using deep nets. We presented a machine learning and artificial intelligence approach for crop prediction, which demonstrated performance in various crops. In this way, we concluded Deep nets for crop prediction. Deep nets will be used to search an acquire data sets for crop prediction. Thus, processing all those attributes will increase the accuracy of the normal prediction.

References:

1. Teresa Priyanka, Pratishtha Soni, C. Malathy, Eurasian Journal of Analytical Chemistry, Issue 02 December 2018, ISSN:1306-3057 OPENACCESS 2018 13 (SP): 6-12.
2. Radhika, Narendiran, International Journal of Computer & Mathematical Sciences IJCMS ISSN 2347 – 8527 Volume 7, Issue 4 April 2018.
3. Dumitru Erhan, Christian Szegedy, Alexander Toshev, Dragomir Anguelov; The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 2147-2154.
4. G Ruß, "Data Mining of Agricultural Yield Data: A Comparison of Regression Models", Conference Proceedings, Advances in Data Mining – Applications and Theoretical Aspects, P Perner (Ed.), Lecture Notes in Artificial Intelligence 6171, Berlin, Heidelberg, Springer, 2009, pages : 24-37.
5. Tripathi S, Srinivas V V, Nanjundiah R S, "Downscaling of Precipitation for Climate Change Scenarios: A Support Vector Machine Approach", J Hydrol, 2006, pages : 621-640.
6. Camps-Valls G, Gomez-Chova L, Calpe-Maravilla J, Soria-Olivas E, MartinGuerrero J D, Moreno J, "Support Vector Machines for Crop Classification using Hyper Spectral Data", Lect Notes Comp Sci 2652, 2003, pages : 134-141.

7. B. Devika, B. Ananthi, International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 12 | Dec 2018, e-ISSN: 2395-0056, p-ISSN:2395-0072, www.irjet.net.
8. Rub, G., “Data Mining of Agricultural Yield Data:A Comparison of Regression Models”, 9th Industrial Confrence, Vol.5633, pp.24-37, 2009.
9. Sellam,V., Poovammal, E., “Prediction of Crop Yield using Regression Analysis”, Indian Journal of Science and Technology, Vol. 9(38), pp.15, 2016.
10. Perpetua Noronha, Divya. J, Shruthi. B.S, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 5, Special Issue 2, October 2016.