



PREDICTING YIELD OF CROP AND DETECTING FERTILIZER EFFICIENCY USING MACHINE LEARNING

¹Dr.Mahantesh Mathapati,²Dr.Jyoti Metan,³Kavita Baragani

¹Professor,²Associate Professor,³Assistant Professor

¹Dept. of Computer Science & Engineering, ²Dept. of Information Science & Engineering, ³Dept. of Chemistry

¹Amruta Institute of Engineering & Mangement Sciences, Bengaluru,India

²Atria Institute of Technology, Bengaluru,India

³RajaRajeswari College of Engineering, Bengaluru,India

Abstract : One of the sectors that contributes most to our nation's GDP is agriculture. But still, the farmers don't get the worth price of the crops. It mostly happens due to improper irrigation or inappropriate crop selection, or also sometimes the crop yield is less than that of expected. The population of the world is always expanding, hence adequate crop production is required. India is a country based mostly on agriculture, and the increase of agricultural profits and agro-industrial goods support its budget. Monitoring crop development and yield estimates is crucial for a nation's commercial expansion. In agriculture, estimating yield is a crucial issue. Crop harvest estimates have a continuous impact on national and global economy and play a crucial role in the nutrition administration and nutrition security. Every farmer cares about how much output he can reasonably be expected to produce. With the aid of this estimate, farmers will be able to select crops that are suitable for their farm in terms of temperature, humidity, soil pH, season, and fertiliser. Sections of nutrients, such as nitrogen (N), phosphorus (P), and potassium (K), as well as the crop's location and zone, run parallel to it. In order to create a traditional JK model, all of these record features will be taken into consideration when training the records using numerous suitable machine-learning methods. The organization's goal is to be precise and accurate in estimating crop harvest and to provide the final user with the relevant, essential manure fraction based on the plot's full of atmospheric and soil elements, which will increase crop harvest and farmer income.

IndexTerms - yield prediction, fertilizer, Support Vector Machine (SVM)algorithm, regression algorithm.

I. INTRODUCTION

Cultivating is the principle control of India. Around 70 percent of an essential auxiliary business depends on cultivating. So for the improvement of cultivating, numerous farmers have begun utilizing the new innovations and techniques. In case individuals don't require mindfulness, nearly the development of the yields in an opportune period and at an ideal spot. For this situation, a plan to recognize the appropriateness of harvests and yield dependent on different components that influence the creation can build the quality and the yield of yields, along these lines increment the financial development and accomplish gainfulness. Farming is a business with a hazard, and a dependable harvest yield forecast is essential for choices identified with agribusiness chance administration. The vision of satisfying the world's nourishment needs for the expanding populace all through the world is getting progressively significant in these ongoing years. In the long run, it helps in accomplishing ZERO earnings. Forecasts could be utilized by crop supervisors to limit misfortunes when negative conditions may happen.

The Data Analysis is a procedure of cleaning investigation, information displaying with the target of finding valuable data and ends. So as to extricate some examples, it is a procedure of dissecting, removing, and anticipating important data from immense information. Ranchers utilize this strategy to gather their client's crude information for helpful data. This examination can likewise be utilized in the field of Agriculture. Most ranchers[1] were relied upon on their long-terms encounters in the field on specific yields to expect a better return in the following collecting time frame, But still, they don't get worth the cost of the harvests. It is, for the most part, occurs because of the ill-advised water system or unseemly harvests determination, or likewise, some of the time, the harvest yield is not as much as that of anticipated. Agrarian specialists[2] demand the requirement for a productive system to foresee and improve the harvest development, and the Majority of research works in the farming spotlight on natural instruments to recognize crop development and improve its yield. The result of harvest yield fundamentally relies upon parameters, for example, an assortment of yield, seed type, and ecological parameters, for example, state, district, year, season, crop, temperature, humidity, area, and soil moisture. By examining the dirt and climate at specific area best harvest so as to have

more yield and the net yield can be anticipated. This forecast will support ranchers[3]. To pick fitting harvests for their homestead as indicated by the dirt kind, temperature, stickiness, water level, dispersing profundity, soil PH, season, manure, and months. Harvest yield valuation is a troublesome undertaking; meanwhile, it is influenced by different aspects, for example, the hereditary capability of harvest cultivar, soil, climate, development rehearses (date of planting, a measure of the water system and manure, and so forth.) and biotic pressure. A few strategies[4] for crop yield estimation have been grown, for example, factual, agrometeorological, experimental, biophysical, and robotic.

India is a heavily populated nation, and a haphazard shift in climate situations is necessary to verify the nutritional resources of the ecosphere. Designers face significant problems in conditions of dry spells. Sort of soil takes on considerable harvest yield work. Proposing[5] the use of manures will allow the ranchers to make the best choice for their editing. The amount of study Brilliant agribusiness is the method for passing on data from conventional ranchers to the informed ranchers. To get appraisals of total physical creation capacities for the yields of different harvests in indicated states, considering different mechanical variables and a recently created climate file as data sources. Relapse and coefficient of assurance examination alongside the Average Fault rate were done to mark a not too bad correlation between our real outcome, which is the so-called target and forecast model, which is a well-disposed crossing point for farmers, which contributes the investigation so that generation dependent on accessible information. Various Data collection strategies[6] were utilized to foresee the harvest yield for augmenting the harvest profitability Accurate and opportune checking of horticultural harvest conditions and assessing potential yield yields are fundamental procedures for operational projects[7]. Because of the significance of anticipating crop yield, the motivation behind this examination is to apply a few gauging techniques for assessing crop yield evaluates in Ghana[8]. Harvest yield estimating, which gives data to chiefs. The system provides easier and faster access to all the basic information regarding the district, rainfall, the area under irrigation, crop, season yield, fertilizers used through which user can analysis the crop and also select the option of prediction where he can select the crop production parameters to get the suitable crop for his farm. This system provides simple visualization so that users can understand and analyze things in an easy way[9].

There is a main machine learning algorithm that is used to determine whether farmers get their yield or not. In the present system, the SVM algorithm is used as a classifier, which proves to give the highest accuracy among the other algorithms. [10] Instances can be classified by more than one output. There is a broad survey conducted on the type of machine learning algorithm used, based on which the yield of a crop is predicted.

II. LITERATURE REVIEW

Support vector regression (SVR) classic model is used as the leading technique to forecast rice harvest. From 1995 to 2015, we combine the climate change data and farm production data collection in Taiwan. Results of the experiment show that even the performance of intermediate SVR systems in terms of root mean square error (RMSE) and that the correlation coefficient (CC) is higher than the standard SVR model. The hybrid SVR models avg RMSE and CC is 60 and 0.996, overall. Hence, hybrid SVR simulations are proper to predict cultivated output since it offers high consistency and high strength

prediction outcomes suggested by Jheng, T[11].

Agriculture is practiced in India as a conventional job, although agriculture is still not practicing as the technologically-driven or technologically oriented profession. As a consequence, India's farming activity is just not producing significant economic results. Data mining to cultivation is one of this kind ongoing trend which focuses on improving agricultural practices by taking into account data on crop yields. In order to evaluate the independent or Combination impact of weather parameters and soil type on crop yield, this crop production data can only be applied to the atmosphere and soil data suggested by Manjunatha, M[12].

Farmers typically prepare the cultivation system on the basis of past experience. They end up cultivating undesirable crops because of the lack of detailed information about cultivation. In order to help farming make decisions that can make their farmland more productive and competitive, the research is attempting to develop an intelligent knowledge prediction framework

on Bangladesh agricultural development. The way of cultivating here, however, is already at the initial level. The studies suggest the rank of profitable crop based on the area prior to cultivation method. It shows the crops that seem to be cost-effective for

cultivating in a given piece of land suggested by Shakoor, M[13].

Since growers have had many effects on their production cycle, they also recognized the importance of encouraging future behavior on accurate climate models, science research has also increased in recent decades, and scheduling offices are searching

for good practices, today's researchers and strategy offices are attempting to connect this whole actuality into a new approach called "Climate Smart Agriculture". 'Crop planning' is a software designed and manufactured at the CCAFS hackathon, constructed as a global forum for the exchange and seeding of crop timelines and agro-climate database. It analyses pen datasets

such as historical development, land spread, and overall climate conditions and incorporates a user experience for a crowdsourced set of cultivation activity dates; it is calculated to offer easy contact to related data and modernized crop datebooks from

farmers and moreover share organization performance from limited specialists suggested by Grajales [14].

In farming, documentation is available on the internet of reports on climate and soil type, GPS monitoring, water supplies, use of fertilizer pesticides, field features, and economic conditions for commodities. Big data software has tremendous potential for referencing these details and creating detailed knowledge through geospatial analysis, remote sensing, advanced statistical algorithms, cloud tools, and progress storage schemes. This paper recommends an inspiration-based facts supervision system for

cultivation and expects to condense the technical gap stuck between agro consumers and statistics suggested by Shah P.[15]. The main knowledge of this research focuses on Indian producers as it discusses the key issues of having different productconsumer status, weather warnings and also provides a variety of development support. This will help farmers successfully sell their goods on the world market and gain considerable profit. At the very same time, obtaining details about weather conditions is a major challenge; Method of fertilization includes the use of specific crop method. Since it was not good to communicate to the farm owners' production, the government also approved various strategies, benefits to the agricultural areas and to the farmers for their resources. All this information can be stored and properly analyzed and shared with growers suggested by Sekhar C.[16]. A regression-based, complicated time series guide that tackles the quietness, obstruction, truthfulness, and spuriousness of the various aspects of the messy context and offers a stronger and even more realistic outcome than the legally sanctioned and unquestionable approaches. The frequency-based partition table has been used as a discourse partition and as the universe of discourse, real output. Eventually, to effect the fuzzification mechanism, Fuzzy Practical Connections of varying levels were performed. Additionally, Another Regression Analysis Method was applied to perform the defuzzification process suggested by Garg A [17].

No program is also in place to recommend farmers what plants they want to grow. Throughout this journal, by analyzing trends in historical data, we introduce an effort to predict agricultural productivity and quality that a producer can acquire through his field. We use a nonlinear regression shifting window strategy to predict the variety of factors involving crop yields. These include rainfall, temperature, market values, land area, and past crop yield. The study is carried out within a few districts of Tamilnadu State, India. Our program aims to recommend the perfect crop options for a farm worker to respond to their requirements of today's prevailing social crisis suggested by Raja S[18].

III. DESIGN AND IMPLEMENTATION

The functional block figure of the crop yield prediction scheme is shown in the figure 1.

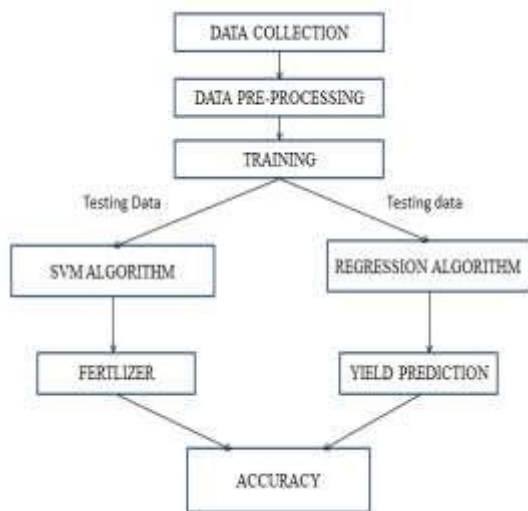


Fig- 1 Functional block of crop yield prediction

The above block diagram works in four modules. To construct a positive machine learning perfect, it is authoritative that an association has the capability to train, test, and certify them previous to installing into production..

A.Data Pre-processing

Data is gathered from different sources; it is connected in row format. Information preparing .is a method that is developed to transform over the unfinished information into a clean dataset. It alludes to change made useful for our information previously nourishing used in the calculation. Hear the crude facts in the yield information are cleaned, and the metadata is annexing to it by evacuating the things which are changed over to the number.

Table- 1 Dataset For Crop Yield Prediction

	A	B	C	D	E	F	G	H	I	J
1	State Name	District Name	Crop Year Season	Crop	Temperature	Humidity	soil moisture	area	Production	
2	Andaman and Nicobar Islands	NICOBARS	2000 kharif	Almond	36	31	43	1254	3000	
3	Andaman and Nicobar Islands	NICOBARS	2000 kharif	Other kharif pulses	37	40	46	1	1	
4	Andaman and Nicobar Islands	NICOBARS	2000 kharif	Box	36	41	50	101	310	
5	Andaman and Nicobar Islands	NICOBARS	2000 Whole Year	Beans	37	41	51	176	641	
6	Andaman and Nicobar Islands	NICOBARS	2000 Whole Year	Cashewnut	39	40	54	738	165	

Table-2 Dataset For Fertilizer Detection

	A	B	C	D	E	F
1	n	p	k	amt of n	amt of p	amt of k
2	3	5	6	64	50	60
3	1	4	2	40	46	30
4	5	1	5	93	16	32
5	3	1	3	63	20	39
6	4	6	6	87	37	39

B. Training the dataset

In Training the dataset module, it includes training and testing phase. Training: The idea of using training data in a machine learning program is a simple concept, but the way in which these innovations work seems to be very foundational. Training data as initially collected to data used to help the system understand how to learn and generate complex results by

adding concepts along with better results. Testing: A test dataset is a dataset independent of the training dataset but, which follows the same distribution of probability as the training dataset. If a model benefits well with the training dataset, then minimal overfitting has occurred.

C. Estimating Algorithm

We use a vector help and regression algorithm. SVM can be used for classification challenges and for regression challenges. This is one aspect of preparation. Training data and named data are fed here as inputs for training the classifier. SVM gains knowledge of current data and classifies unlabelled data. To improve crop productivity, various data mining techniques were used to predict crop yield. We use the regression algorithm for yield prediction and use the SVM

algorithm for fertilizer detection. Precise and timely monitoring of conditions for agricultural crops and estimating future crop yields are important processes for operational programs. Due to the importance of predicting crop yield the purpose of this study is to apply multiple forecasting methods to assess crop yield.

Support Vector Machine (SVM) Algorithm

The Support Vector Machine (SVM) is an obscure machine learning classic model. Which uses a clarification algorithm to classify the 2 group problems? Support Vector Machine is a distinct classifier which is generally defined by hyperplane separation. Support Vector Machine is also defined as labeled training data. This outputs an absolute hyperplane, which groups the new example. In the 2nd system, this hyperplane acts as a line dividing the two parts of the class on each side. Support Vector machine is said to be fast and depending on the classification algorithm, which acts on limited data. The support vector machine working can be seen here, let x and y be the classifier named coordinate support vector machine to analyze the data and give hyperplane as output, which groups the data tags. The hyperplane, which is nearest to each tag, is said to be largest. Machine Learning is usually used to make decisions based on the predictions; machine learning finds the best accuracy in prediction based on data that we give to the system.

$$ACCURACY = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Where TP represents the true positive, TN represents the true negative, FP represents false positive, and FN represents false negative.

Regression Algorithm

Regression algorithms will come under the category of the supervised machine learning algorithm, which is also a subpart of the machine learning algorithm; supervised machine learning mainly depends upon target output and input relations which uses to predict new data regression algorithm produces output values based on input data fed into the system. The output here always depends upon the input we give to the system.

$$J = \frac{1}{n} \sum_{i=0}^n (\text{pred}_i - y_i)^2 \quad (2)$$

Where J represents the cost function of regression is the root mean squared error (RMSE) between the predicted value of y and the true value of y.

D. Prediction

Machine Learning is about Prediction. The predictive model is a testing phase. It maintains extremely around scalability and consuming the estimates to create conclusions. Considerably of Machine Learning is encouraged by complications that must have responses.

IV. RESULTS AND DISCUSSION

In this paper, an effort is made in order to know the yield of crop and detecting fertilizer efficiency it is processed by implementing both the regression algorithm and support vector machine(SVM) algorithm. These models have experimented with different types of crops in various regions across the country to predict the output. Also, fertilizer data was trained using the Support vector machine algorithm and evaluated to get the result of how much amount of nitrogen, phosphorus, and potassium is required for the agricultural land.

The model for the crop yield was determined by using various parameters like type of name of the State, name of the district year of the crop is grown, type of season, crop type, temperature, humidity, and soil moisture with respect to the previous dataset. Comparing with the previous paper, errors can be easily identified by the SVM.SVM is a secure algorithm for understanding and designing when compared with others.

For prediction of the yield, the user will enter the data. The user should enter the details one after Another. The output of the yield prediction is shown in Fig. 2. The user enters the input for the fertilizer data and the out of the fertilizer data is shown in Fig. 3.

Crop Yield and fertilizer Prediction System

Crop Yield and fertilizer Prediction System

state: 1
 district: 3
 year: 5
 season: 7
 crop: 8
 Temperature: 5
 humidity: 10
 soilmoisture: 16
 area: 14
 predict: [153.31004]

Fig- 2 Output for yield prediction

```

enter nitrogen:3
enter posporus:5
enter potasium:6
Amount of nitrogen Fertilizer: [68.72145981]
Amount of posporus Fertilizer: [42.56602245]
Amount of potasium Fertilizer: [45.44868007]
  
```

Fig- 3 Output for fertilizer detecting

IV. CONCLUSION

Agriculture is the foundation of states such as India. The use of technology in agriculture, however, is to be given paramount importance to precision farming. This study indicates a method that will benefit farmers get an understanding of harvest projections established on weather restrictions and areas being cultivated. Using this farmer will be able to determine whether to grow the specific crop or go for alternative crops in the event, the yield forecasts are unfavorable. This research effort can be expanded on to the subsequent level. We comprise create a recommend arrangement for farmers to produce and distribute agricultural products. In which farmers will determine in which season which crop should be shown so that they can gain more. For an organized dataset, this method works. We can incorporate independent data structure in the future too. It means data format whatever, our system should be running with the same performance.

ACKNOWLEDGMENT

The result obtained will help the agriculturalists to predict the harvest of the crop so that the farmers can opt for the healthier plant, which offers high yield and likewise detects the manure efficiency for that particular region considering various parameters. This technique can support farmers to get a better return and achieve zero hunger.

REFERENCES

- [1] Feilong Wang, Fumin Wang*, Yao Zhang, Jinghui Hu, "Rice yield estimation at pixel scale using relative vegetation indices from unmanned aerial systems," Open Fund of Laboratory of Target background conditions in field scale, this Microwave Properties, 2023
- [2] Anill SuatTerksiz, D. TurgayAltıralar, "Use Deep Neural Networks For Crop Yield Prediction: A Case Study Of Soybean Yield in Lauderdale County, Alabama, USA," IEEE, 2023
- [3] Mr. A Suresha, Dr. P. Ganesh Kumar, Dr.M.Ramalatha, "Prediction of major crop yields of Tamilnadu using K-means& Modified KNN," Proceedings of the International Conference on Communication and Electronics System (ICCES) IEEE Xplore, 2023.
- [4] NikietaGandhi, Leisa J Armstrong, OwaizPeatkar, Amiya KumariTripathy, "Rice crop yield prediction in India using support vector machines, Thirteenth International Joints Conference on Computer Science & Software Engineering (JCSSE), 2022.
- [5] XinranGao; Jianxi Huang; Hongyuan Ma; Wen Zhou; Dehai Zhu, "Regional Winter Wheat Maturity Date Prediction Using Remote Sensing-Crop Model Data Assimilation and Numerical Weather Prediction," 7th International Conference on Agro-geoinformatics, 2022.
- [6] TishampatiDha; Doug Gray; Carl Menges, "Multiple Crop Yield prediction using dual-polarimetricTerraSAR-X strip map imagery," IEEE International Geoscience and Remote Sensing Symposium, 2021.
- [7] ZongnanLi; Zhongxin Chen, "Remote sensing indicators for crop growth monitoring at different scales," IEEE International Geoscience and Remote Sensing Symposium, 2020
- [8] J.Ramirez-Villegas and A. Challinor, "Assessing relevant climate data for agricultural applications," Agricultural Forest Meteorology, 2019.
- [9] Claudio O. Stockles, Steves A. Martin & Gayalon S. Campbells, "Crop System, a Cropping System Simulation Model: Water/N Budgets & Crop Yield", Agricultural System, 2018
- [10] PonceGuevara, K. L., PalaciosEcheverria, J. A., Maya Olalla, E.,DominguezaLimaico, H. M., Suarezl Zambrano, L. E., Roseroo Montalvo, P.D., Alvarado-Perez, J. C. (2017), "GreenFarm-DM: A tool for analyzing vegetable crop data from a greenhouse using data mining technique (1st trial)", IEEE 2nd Ecuador Technicals Chapters Meetings (ETCM), 2018.
- [11] Jenga, T.Z., Li, T.H., Lee, C.P., "Using hybrid support vector regression to predict agricultural output", Twenty-seventh Wireless and Optical Communication Conference (WOCC), 2018.
- [12] Manjunath, M., Parkvi, A., "Estimation Arecanut Yield in Various Climate Zone Karnataka using Data Mining Techniques: A Survey," International Conference on Current Trends Towards Converging Technology (ICCTCT), 2018.
- [13] Shakor, M. T., Rahaman, K., Raysta, S. N., Chakrabarty, A., "Agricultural production output prediction using a Supervised Machine Learning technique," First International Conference in Future Generations Computing Applications (NextComp), 2017.
- [14] Grajales, D. F. P., Mehjia, F., Mosquera, G. J. A., Piedrahita, L. C.,Basurto, C., "Crop planning, making smarter agriculture with climatic data," 4th International Conference on Agro Geoinformatics (Agro-Geoinformatics), 2015.
- [15] Sha, P., Hirematha, D., Chaudary, S., "Towards development spark based agriculture information system includes geospatial data," IEEE International Conference on Big Data (Big Data), 2017.
- [16] Sekar, C. C., Sekar, C., "Productivity improvements in agricultural sector using bigs data tools," International Conference on Big Data Analytics and Computational Intelligence (ICBDAC), 2017.
- [17] Garg, A., Garg, B., "A robust and novel regression-based fuzzy time series algorithm for prediction of rice yield," International Conference on Intelligent Communication and Computational Techniques (ICCT), 2017.
- [18] Raja, S. K. S., Rihi, R., Sundarsan, E., Srijit, V., "Demand based on crop recommendation systems for farmers", IEEE Technological Innovation in ICTs for Agricultural and Rural Developments (TIAR), 2017
- [19] Afrina, S., Khan, A. T., Mahia, M., Ahsan, R., Mishala, M. R., Ahmed, W.,Reahman, R. M., "Analysing of Soil Property & Climate Data to Prediction Crop Yields & Clusters Differentiate Agriculture Region of Bangladesh," IEEE/ACIS Seventeenth International Conference on Computer & Informations Sciences (ICIS), 2018.
- [20] Sahau, S., Chawala, M., Khasre, N., "Ancient analyze of crops yields predictions using Hadoop frameworks based on a random forest approach," International Conference on Computing, Communication and Automations (ICCCA), 2017.