

Market Price, Crop and Crop yield Prediction Using Machine Learning

Sumanth Kumar B¹, SuryaKant Kumar², Yashraj Singh³, Suraj M⁴Dr. D V Ashoka⁵

(Information Science and Engineering, JSS Academy of Technical Education, India)

Abstract: This Project is an Attempt to minimize the losses occurs in Agricultural field. The majority of experimenter's work on agribusiness focuses on biological mechanisms to identify crop growth, improve yields and price prediction hence presentation of agribusiness is influenced by several weather aspects. Metrological parameters include wind speed, temperature; moisture, and Humidity, precipitation parameters include rainfall, region-wise rainfall, irrigation, etc., and soil parameters which include PH, organic carbon, phosphorus, fiber, etc. And due to continuous change in climate conditions, everything is messed. In India farmers still follow the traditional technology which they adopted from their ancestors. But the problem is that in the earliest time climate was very healthy everything happened on time. But now most of the things have been changed due to global warming and many other factors. The leading annoyance with agribusiness in India is the shortage of rainfall in seasonal periods. Humidity is also required for production, though it has been unreasonable, it also transforms as a weakness. This Project is an attempt of predicting the outcome of harvest supported the current data by using RFA Random Forest Algorithm and Back Propagation.

Keywords: RFA Random Forest Algorithm, Back Propagation, Price prediction, Crop Growth, Improve Yield

I. Introduction

Agribusiness is a considerable significant sector of the Indian Economizing. Indian Farming The sector occupies 18 percent of India's GDP and employs 50% of the country's workers. But most notable experimentations have documented a successive reduction in the offering made by agribusiness to the Indian economy and may also be a conception of the money needed for industrial growth to the dimensions although it is demographically the most comprehensive financial sector and plays a momentous function in the overall socio- economic material of India. Correct and convenient monitoring of farming crop requirements and evaluating possible harvest products are fundamental techniques for functioning programs Crop outcome forecasting, which supplies knowledge for decision Makers. Because of the effectiveness of predicting crop, crop price and yield is the definition of this study and to apply several Algorithm procedures for contemplating crop yield computations in various locations. For estimating the yields of various crops in specific states using aggregate physical production functions. A newly developed weather index and various technological factors are used as inputs. To make a respectable assessment among our real result, that is known as target, and prediction model, that is a pleasant interface for farmers and offers the evaluation of rice manufacturing primarily based totally on to be had data, regression and coefficient of dedication evaluation, in addition to the Average Error rate, have been carried out.

II. Problem Statement

The production of agriculture is affected by several climate factors. Like as crop price, metrological parameters, precipitation parameters, and soil parameters. And due to continuous change in climate conditions, everything is messed. In India agriculturalists still follow the classic technology which they assumed from their predecessors. The detailed and accurate information about the crop and price across various states helps us to transform our ability to manage and implement. Here we investigate the crop and yield data collected by various states and average price of the crop and its location using machine learning technique.

III. Objective

Advancements in Artificial Intelligence and implementing the way people use the technology with crop and price prediction in agriculture the learning of the connotations of the concepts, access to the latest approaches, methods, and theory. Realizing the investigation topics based on the current research Concentrate on your domain of expertise even if one field uses the same words, they usually represent completely. And furthermore enhances the difference of the publication's survey to restrict sidetracks and recognize to illustrate what is excluded.

1. **Forecasting of Crop Yield:** production levels. Any farmer is interested in knowing how much harvest is expected. In a one-time harvest, speculation is made by examining the knowledge of farmers in a particular field and crop.
2. **Prediction of Crop:** In addition to crop prediction and crop price being defined by multiple factors, crop selection also depends on its price and plays a major role.
3. **Prediction of Crop Price:** Crop Price and Testing you should make a wise decision before planting a particular type of crop. About the cost of the plant will help to bring about a better conclusion that leads to the correct judgment of losses and control of the risk of price change.

IV. Related Work

[1] **Lstm Based Crop Price Prediction System (2021):** The procedure is used to predict the expense of rare crops using elements such as the place and tentative sowing date. Two approaches were compared one is using machine learning algorithms and the second one is a hybrid sequential modeling approach. In this analysis aspects like Area gathered, the result of a crop, and time-series data of one-time cost of produce and crop season are taken into consideration. As the price of crops varies from location to location, a particular location of Pune has been selected for this study. Following datasets, one is having time-series data of crop, area, and crop production taken from data.gov.in and another time-series data of crop and past prices from agmarknet.gov.in are used. Both datasets are merged based on crop names, year, and month. The Data was checked for any kind of correlation and the parameters that had less correlation were removed. Parameters viz. 'Crop Year', 'Crop', 'Area', 'Yield', 'month', 'Modal Price (Rs. /Quintal)' were considered for correlation. Library called Lazy Predict is used to fit the dataset with several algorithms like Random Forest Regressor, Decision Tree Regressor, KNN, MLP Regressor.

[2] **Cluster Prediction and Forecasting System using Supervised Machine Learning Algorithms (2021) :** This method used in the system is the deceleration of the decision tree which is a way of minimizing machine learning. The parameters considered in the forecast are: - rainfall, wholesale price indicators (minimum support [MSP], farming costs, planting costs, etc.). Accurate forecasting of crop prices; is important in crop production management. Such forecasting will also help affiliated industries that rely on agriculture to find raw materials to strategize for their business. With the benefits of this app, farmers get an early forecast that allows them to maximize their profits and prevent major losses. Which will also increase the country's economy.

[3] **Estimates in Agriculture: Predictable Prices for Arecanuts in Kerala (2020):** Kerala's monthly defense costs are expected using a series of time and engine learning ideas. Models such as SARIMA, the Holt Winter's Seasonal strategy, and the LSTM neural network were used, and use was tested based on RMSE value in the areca nut database and costs from 2007 to 2017. The LSTM neural grid model was the most appropriate data model. Information obtained from the Department of Economics and Statistics; Kerala included a monthly price for areca nuts from 14 states in Kerala during the period 2007-2017. Views from the years 2007-2016 were used for the training set, and prices from 2017 were used for testing. Since available data was limited, it was not a good idea to allocate data for verification set data as it would interfere with learning. The acquired database contains values that are not available in most regions. The line translation method was used in regions with less than 50% deficiencies. Wayanad had very low rates, which led to data fraud. These figures are completed using the multi-line retrospective method. The prediction variables were determined using the correlogram, which included the correlation between prediction and predictive variability. Neighbors with a significant value or equal to 0.7 were considered predictor variables. The database has been simplified by considering the monthly data rate for the regional framework.

[4] **Emerging Economic Predicting Framework for Analyzing the Timeline Data Quality Analysis(2020):** Controls for machine experience are met to illustrate the cost of cutting. Strategies for predicting harvest costs are required to be realistic in life because they encourage the organizers of the collection chain and leadership to take the right stance by measuring the exact need and supply. In India, crop prices on the market are handwritten on a daily basis, which is often the result of human error such as improper data entry or data failure for days. In addition to such errors, price fluctuations in themselves make making accurate and robust forecasting decisions a challenge. All of these details in crop cost estimates are considered in this construction, different processes are used to create robust crop cost models to look at a few factors such as, (i) historical prices and market value of crops, (ii) climate history data and their impact on crop production and transportation, (iii) related factors and the quality of data from statistical analysis.

[5] **The Crop Price Prediction System uses a machine-learning algorithm (2020):**The main purpose is to show the yield payment for the following method. The task to be done is to find relevant data models

aimed at achieving high and standard accuracy in estimating crop values. To solve this problem, different Data Mining strategies are tested with different data sets. This paper proposes a system that uses data analysis techniques to predict crop prices. The proposed method will anticipate the cost of the plant established on factors such as Harvested Area, Planted Area, etc. using machine learning algorithms. This will give farmers an idea of what the future price of the crop will be. The design promotes the creation of configurations by combining data from mixed assumptions, data analysis, and advanced analysis that can help report the cost of crop marks and increase revenue for agricultural exports. Complete research concludes that XGBoost is the right approach to the project. Linear Regression is used to obtain important information from a given agricultural database. Neural Network is designed to predict the price that increases percentage accuracy. The root means that a square error is calculated across all strategies to determine the accuracy of each system used and to select the most suitable system. Strong interdependence was found between the yield/production of soybeans and exports/imports, similarly between US soybean exports and imports.

[6] **The smart Core Price Predict uses a valid machine learning algorithm (2021)** :A site model is proposed for farmers to be able to support their finances, needs, and possibilities as well as other metrics that select their preferred crop. The dashboard will show the best yields as well as the most damage and distribution that they approach. Predictability will be made in the next 12 months. Details of 20+ plants will be taken from management and will be represented in a systematic way that represents the rise and fall of crop prices every month and reflects the details of the crop such as its type, location, and export characteristics. which will help farmers to plan and manage their finances and sow/harvest properly. The interface is an easy-to-use interface and predicts the resolution of the decision tree. An in-depth statistical analysis of the above-mentioned data was performed to compile an advanced exchange method that could be reached by agricultural experts so that they could use the forecast for the next 12 months. Firstly modern data sets will be adopted that include rainfall costs and supermarkets per month per plant. After preliminary processing of the data, the model will be trained and judged accordingly. If found suitable, the website will be designed and the ML model will be used in the background.

[7] **Estimating Prices Using Machine Learning in Madhya-Pradesh: A Pilot Study (2020)** :Certain crops were selected for pricing taking into account factors such as sub-crop, crop yields, and price volatility, Wheat and paddy crops were avoided due to low price volatility and MSP which determines their prices. A common approach to machine learning concepts followed in this project is the following steps: Identify conclusions and obtain data: Discussion of domain experts, identifying influential factors, and obtaining historical data from key stakeholders. Prepare data: Clear in advance, clean and convert the data into a machine-readable format. Train model: Using trained data to develop multiple value guessing algorithms. Test Model: Testing and output analysis of various algorithms and short listing of highly efficient algorithms. Use and live to test: Use the live test algorithm over a period of time and further development using essential learning. In pricing forecasts the following steps are followed, pricing decisions, data collection, data purification, and formatting algorithm training, live prediction follow-up tests.

[8] **Focus on Choosing a Model Based on Conclusive Conversion Price Index(2022)**: Many of the most widely used methods for forecasting timelines are affected by factors such as data quality, as well as general instabilities during times of high volatility, which are typical of the Indian economy. A response to these shortcomings was to introduce model-specific modeling and demonstrate its functionality in two separate ways: feature concatenation (FC) and time-series dependent features (TD), which take into account weather, pre-market arrivals, and data quality Predict crop prices accurately. Furthermore, a context-based model is proposed for selection and retraining using trend analysis in order to improve crop pricing estimates, especially during periods of high volatility in prices. According to the results of the experiment, the model-based strategy selection strategy improves the accuracy of the Tomato and Maize crops in both types of feature submissions. Two factors can be monitored in the near future - effective data quality improvement and context-based model development. To reduce problems with data quality, strategies such as obtaining non-existent values from a nearby market place can be employed.

[9] **Crop-yield and Price Forecasting using Machine Learning(2020)**: The suggested model has been materialized to support agriculturalists create more profitable decisions concerning which harvest is most appropriate during his expected duration of sowing and the surroundings. This procedure indicates the product and expense of the produce of preference, delivering the agriculturalist valuable knowledge satisfactorily before beginning the process of cultivation. Many guessing algorithms can be used to generate yield and price estimates such as SVM, decision trees, Neural networks, etc.This representative uses the Decision Tree. It is trained in a few ragi plants (such as paddy, arhar, bhajan, barley, etc.), and Kharif brings logical precision. In order to provide a sustainable livelihood to growing populations, agronomists need to increase the supply of

food in the current farming environment to avoid deforestation. The longer-term assignment accomplished ought to optimize agriculture techniques to boost profits, harvest rate, and payments in an exceeding effects technique. A State-of-the-art deal presentiment design may be generated wherever a dashboard can foreknow the demand matter directions exploitation statistical procedure for a minimum of duration and therefore the presentation convention of miscellaneous produce. A Medium for Agriculture ought to use enormous knowledge, Machine Learning, AI, satellite metaphorical approach, and weather information to evaluate the land place and monitor crop healthiness on a true-time basis. so it'll glimpse cuss and illness infestations, evaluate the product result and earnings, and conjointly forecast expenses.

[10] **Guess the price of a stock using algorithms for machine-readable learning algorithms (2021):** Pre-product costs and reflect costs for the next 12 months are recommended. The flash-based website is used to display data and the website operates with an adequatemachine to understand algorithms and communication technologies that are easy to use for users. The job data sets used provide information showing the cost and demand of the market plants. This program allows farmers to solve their problems and increase their income. Various algorithms such as decision trees, vector support machines, neural networks, in-depth learning, etc. are used to predict crop prices. The proposed model uses a machine learning algorithm called Decision tree Regressor. Trained with some of the Kharif and Ragi plants (eg iPaddy, Wheat, Cotton, Barley, etc.) it provides better accuracy. The model can continue to be trained in climate-sensitive farming methods, provide fertilizer proposals, and identify crop monitoring systems, pest warnings, disease outbreaks supported by advanced AI models.

[11] **Harvest Prediction Using Machine Learning (2020):** A set of information is collected from a variety of sources and the data provided is used for analysis (descriptive and diagnostic). For the past ten years annual crop summaries have been used. The datasets accept the behavior of anarchic time series. Combining the primary and necessary abstracts the Random Forests approach is used for Global and Regional Crop Yield Predictions. The entire database is divided into 2 parts: 75% of the model training database and the remaining 25% of the data will be used for model testing. To anticipate prospective possibilities Commanded device learning algorithms can be devoted to what has existed comprehended in history utilizing labeled instances. After a sufficient amount of practicum, the method can yield marks for any untested practical information. To change the model, the information algorithm can differentiate its products with accurate, organized products and can detect errors.

[12] **Crop Yield Prediction using Machine Learning Algorithms (2021):** Primary emphasis is the forecast of produce and product analysis with the usefulness of machine learning procedures. Several machine learning procedures are used for authentic computations. Random Forest classifier was employed for the harvest outcome foretelling for the designated section. This technique is executed for crop foretelling from the accumulation of one-time data. The proposed procedure enables farmers to select which crop to cultivate. This appointment is operated to explore release and acquire an understanding regarding the products that can be conceived to make efficient and reasonable harvesting. The unvarnished and fortunate prognoses of different crops across different districts will help the farmers. This will improve Indian economizing by maximizing the yield rate of cultivated crops. In the coming years, a data-independent system can be applied to this system. In all formats, this system should work with the same accuracy. Integrating the details of soil into the system is an advantage, as for selecting crops soil knowledge is a good parameter. Proper irrigation systems are also needed in feature crop cultivation. About rainfall can be depicted whether extra water availability is needed or not. This work can be enhanced to a higher level by availing it to different parts of India.

[13] **An interaction regression model for crop yield prediction (2021) :** An exchange regression representative for harvest yield foretelling is presented, which produced three significant assistance. Rather, it transcended machine learning algorithms concerning foretelling accurateness in a separate encyclopedic case study, which employed the aforementioned data of three Midwest circumstances from 1990 to 2018. Second, it remembered a dozen $E \times M$ interchanges for soybean and grain products, which are spatially concentrated and can be utilized to form counter-intuitive and insightful assumptions. Third, it explained the subsidies of climate, soil, administration, and their interchanges to crop harvest. Acquiring these subsidies accompanying is significantly significant since most of the crop yield projection algorithms are not capable of satisfactorily addressing both prediction accuracy and explain ability. The presented representative and computational experimentations had some restrictions. For instance, the feature and interaction appointment algorithms were heuristic in nature, which can find high-quality solutions efficiently but global optimality is not guaranteed. By increasing the number of features (genetic information), the proposed algorithm may lose its efficiency in terms of running time in finding robust features and interactions. This model is seeking self -or two-way interactions. The forthcoming investigation should focus on exploring the

possibility of including additional data (such as high dimensional genotype data, plant traits, straightforward administration strategies, and satellite images) to improve prophecy accuracy and construct additional biologically and agronomic ally insightful findings. **[14]Impact of El Niño on Oil Palm Yield in Malaysia (2021):** The investigation suggests a refreshed oil palm creation index, anointed Fresh Fruit Bunch Index (FFBI) emanated from the monthly oil palm FFB outcomes. It has been demonstrated sufficient in the improvement of the movement and oil palm yield modeling in Malaysia regarding El Niño events. The FFBI establishes a comparatively additional heightened correlation with ONI than FFB results established on Spearman’s rho correlation examinations at 0.01 alpha levels. The statistic benchmark features a remarkably more increased predictive accurateness compared to the FFB model. The FFBI representative has more miniature misconceptions founded on residual calculations. The proposed FFBI provides an improved strategy to model the consequence of El Niño occurrences within the Malaysian feather palm industry, as financial losses caused by El Niño events cannot be neglected. An exceedingly powerful El Niño event is figured to generate 3.07 billion USD of conceivable trouncing when launched in December 2021 operating a 6% discount rate. The accommodation combined possibility losses of 13 past El Niño events, ranging from 1986 (banning 2018/19 weak El Niño) is approximately 9.55 billion USD, which is roughly 2.84% of Malaysia’s 2020 Gross domestic product.

[15]Stability and Adaptability of Maize Hybrids for Precision Crop Production in a Long-Term Field Experiment in Hungary (2021): this deconstruction is concentrated on Bearable agribusiness performs in the attractions of human beings. It is additionally efficient in utilizing help possessing counterbalance with the circumstances. In other words, sufferable agribusiness is ecologically applicable economically achievable, and socially desirable. This investigation demonstrated the ultimate oil range in FAO340, starch in FAO330 on NPK, and nitrogen fertilizer protein in FAO350. Grain product is estimated which demonstrated ultimate equilibrium in FAO420 Agronomy 2021, 11, 2167 12 of 14 as a result of NPK fertilization and FAO430 as a result of nitrogen fertilization. Depending on the given fertilizer, farmers can use the various type of hybrids to obtain stable and high yields. It was uncovered that nitrogen fertilizer resulted in the most significant grain yield in maize, and potassium and phosphorus with nitrogen stabilized the kernel yield along with other parameters. It instructs operating the entire fertilizer NPK amount with (150 Kg/ha nitrogen, 115 Kg/ha potassium, 135 Kg/ha phosphorus) on maize combinations.

V. Methodology

The dataset collected is the input of our project. This dataset undergo pre-processing technique where the unwanted data or the empty data are removed from the datasets. After preprocessing the data is loaded to the system and different machine learning algorithm is executed to bring out the best one algorithm. The scope of this project is to bring out the predictions on Crop, yield, and price. Finally, according to the user input in the front end the model processes the task in the backend the displays the output of the predictions.

VI. General Procedure

A web app can be developed to display the prediction of the output and for providing a smooth use. The overall procedure for identifying and classifying the crop and price is depicted in fig.1

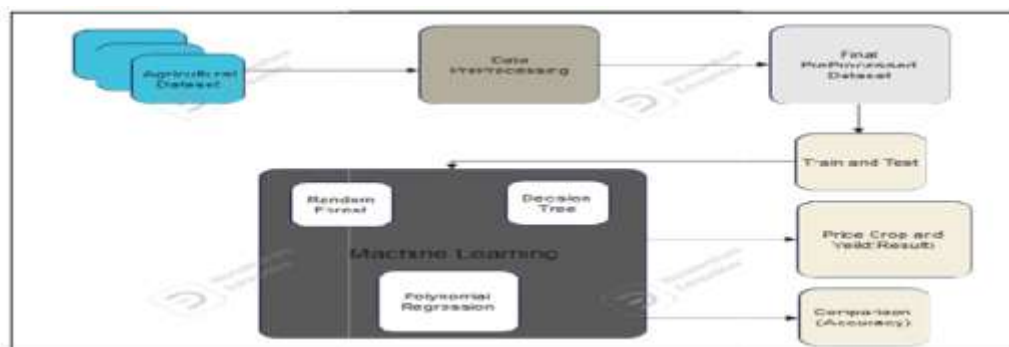


Fig.1: General Procedure

A model trained with a suitable datasets should be identifying the crops and prices respectively. After identification, the model should classify the crop, yield and price based on different states. Fig. 2 represents an expected output.

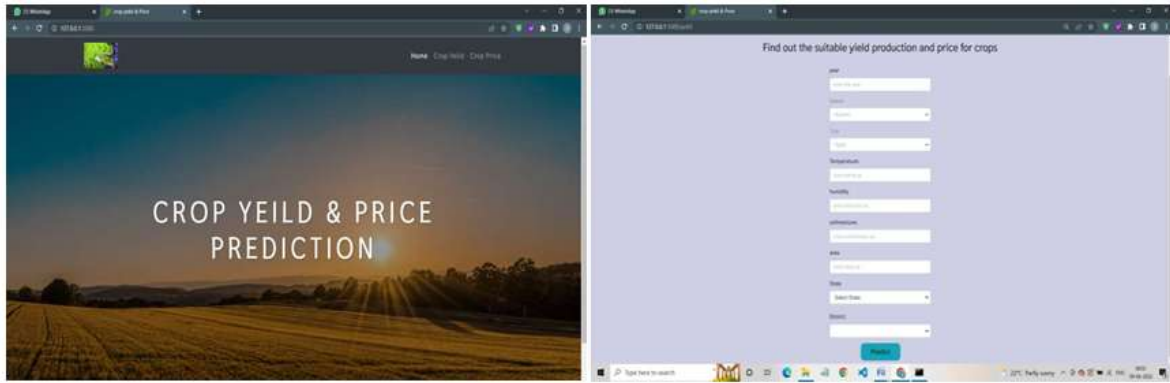


Fig 2: Expected Output

VII. Conclusion

We have proposed an approach of Predicting the crop, crop-yield and crop-price using RFA Random Forest and Back Propagation Algorithm. This approach helps reduce the loss faced by the farmers and improve yield to get better Agribusiness. In this phase of the project, we reviewed the literature on the price prediction, crop-yield prediction. This literature helps us understand the challenges that we face in the price dataset to identify the price and yield. According to the literature experimental results we came to know that random forest and back propagation helps in crop-price Identification and helps to eradicate the challenges that we have mentioned in the dataset that comes in the dataset and also the accuracy rate of this Algorithm is far better than other Algorithm. This helps the farmer to study about Agricultural ecosystem and can be improved by integrating this with other departments like sericulture, and other towards to development of villages.

VIII. Acknowledgment

It is our pleasant duty to place on record our deepest sense of gratitude to our respected guide Dr. D V Ashoka, who is a professor and Dean at JSS Academy of Technical Education, Bangalore for the constant encouragement, valuable help, and assistance in every possible way.

References

- [1]. Kiran M. Sabu, T.K. Manoj Kumar, Predictive analytics in Agriculture: Forecasting prices of Arecanuts in Kerala, *Procedia Computer Science*, [2]. Applied different Algorithm comparison <https://www.analyticsvidhya.com/blog/2021/05/5-regression-algorithms-you-should-know-introductory-guide>
- [2]. F. F. Haque, A. Abdelgawad, V. P. Yanambaka and K. Yelamarthi, "Crop Yield Analysis Using Machine Learning Algorithms," 2020 IEEE 6th World Forum on Internet of Things (WF-IoT), 2020, pp. 1-2, doi: 10.1109/WF-IoT48130.2020.9221459.
- [3]. M. Kavita and P. Mathur, "Crop Yield Estimation in India Using Machine Learning," 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), 2020, pp. 220-224, doi: 10.1109/ICCCA49541.2020.9250915.
- [4]. P. S. Nishant, P. Sai Venkat, B. L. Avinash and B. Jabber, "Crop Yield Prediction based on Indian Agriculture using Machine Learning," 2020 International Conference for Emerging Technology (INCET), 2020, pp. 1-4, doi: 10.1109/INCET49848.2020.9154036.
- [5]. R. Kumar, M. P. Singh, P. Kumar and J. P. Singh, "Crop Selection Method to maximize crop yield rate using machine learning technique," 2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), 2015, pp. 138-145, doi: 10.1109/ICSTM.2015.7225403.
- [6]. Analysis Of Crop Yield Prediction Using Data Mining Techniques (2015).
- [7]. Effects of NP Fertilizer Placement Depth by Year Interaction on the Numeral of Maize (*Zea mays* L.) Plants after Emergence Using the Additive Main Effects and Multiplicative Interaction Model (July 2021).
- [8]. Bojtor, C.; Mousavi, S.M.N.; Illés, Á.; Széles, A.; Nagy, J.; Marton, C.L. Stability and Adaptability of Maize Hybrids for Precision Crop Production in a Long-Term Field Experiment in Hungary. *Agronomy* **2021**, *11*, 2167. <https://doi.org/10.3390/agronomy11112167>
- [9]. Khor, Jen Feng & Ling, Lloyd & Yusop, Zulkifli & Tan, Weilun & Ling, Joan & Soo, Eugene. (2021). Impact of El Niño on Oil Palm Yield in Malaysia. *Agronomy*. *11*. 2189. [10.3390/agronomy11112189](https://doi.org/10.3390/agronomy11112189). (2021).
- [10]. Ansarifar, J., Wang, L. & Archontoulis, S.V. An interaction regression model for crop yield prediction. *Sci Rep* **11**, 17754 (2021). <https://doi.org/10.1038/s41598-021-97221-7> [12]. Crop Yield Prediction using Machine Learning Algorithms (2021).
- [11]. Ranjani Dhanapal et al 2021 *J. Phys.: Conf. Ser.* **1916** 012042 Crop price prediction using supervised machine learning algorithms
- [12]. Jain, Ayush & Marvaniya, Smit & Godbole, Shantanu & Munigala, Vitobha. (2022). Towards Context-based Model Selection for Improved Crop Price Forecasting. [10.1145/3493700.3493725](https://doi.org/10.1145/3493700.3493725). (January 2022).
- [13]. AIGGPA; Crop price prediction using machine learning in Madhya Pradesh; A pilot study, Atal Bihari Vajpayee Institute of Good Governance and Policy Analysis, March 2020.
- [14]. Ghutake, I., Verma, R.K., Chaudhari, R., & Amarsinh, V. (2021). An intelligent Crop Price Prediction using suitable Machine

- Learning Algorithm. ITM Web of Conferences.
- [15]. Ashoka, D. V., Abhilash, C. B., & Bharath, M. B. (2019, March). A Framework for Enhancing the Agriculture yield using Cloud Clusters. In 2019 1st International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE) (pp. 162-166). IEEE.
- [16]. Chetan, R., Ashoka, D.V., Ajay Prakash, B.V. (2021). Smart Agro-Ecological Zoning for Crop Suggestion and Prediction Using Machine Learning: An Comprehensive Review. In: Chiplunkar, N., Fukao, T. (eds) Advances in Artificial Intelligence and Data Engineering. Advances in Intelligent Systems and Computing, vol 1133. Springer, Singapore. https://doi.org/10.1007/978-981-15-3514-7_94