JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue

JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

AGRICULTURAL PRICES FORECASTING USING MACHINE LEARNING

[1]Dinesh, Dr. P. V. Lakshmi

[1]M.Tech Scholar, [2]Professor

[1] Dept of Computer Science and Engineering,

[1] GITAM University, Visakhapatnam, India

Abstract : Agriculture and agricultural sectors account for the majority of the country's income and economy in India. Farmers have had a difficult time anticipating agricultural produce prices in the past. Farmers are currently losing a significant amount of money as a result of price volatility induced by climatic change and other price-influencing variables. Farmers are unable to receive the price for their produce that they seek. The purpose of this research is to create a decision-making support model for predicting agricultural commodity prices. This method can be used as a reference for determining what crops a farmer should plant, taking into consideration parameters such as annual rainfall, WPI, and other variables. The system produces a detailed 12-month projection. The decision tree regression technique is a machine learning regression technique.

Keywords - Price Prediction, Machine Learning, WPI, Decision Tree Regression.

INTRODUCTION

Agriculture is the most significant economic pillar in our country. For the vast majority of families, agriculture is their primary source of income. Agriculture takes up 60% of the land in the country to support the demands of the inhabitants. Our research aims to find a practical solution to the problem of crop value prediction so that farmers can be assured of a return. A vast variety of agricultural products are available at the market. A multitude of factors determine agricultural product prices, and even the same product can be priced differently in different markets. Agricultural commodity prices are naturally volatile, meaning they can rise or fall at any time, causing economic chaos. This system forecasts crop value using the Decision Tree Regression approach.

A good crop price forecasting system can present clients with more options that will meet their needs. Finally, the results are given in the form of a web application that farmers can use. Farmers will profit from the work done here to forecast horticultural products costs since they will be able to sow crops based on their expected expenses in the future. Farming items are subject to regular rates, which are spread out over the course of the year. These rates will be guaranteed on Rate on Investments if they are revealed to farmers ahead of time (ROI). Horticulturists can keep track of these graphs and estimate advertising rates to pass on to farmers.

RELATED WORK

Supervised Machine learning Approach for Crop Yield Prediction in Agriculture Sector. The Random Forest Algorithm and Decision Tree are used to forecast annual agricultural productivity. It uses Kaggle data for training reasons. The system requires fewer computer resources and can produce results rapidly due to the algorithm's relative simplicity. Only structured data could be used with it. As a result, it does not apply uniformly to different types of data.

Crop yield prediction through proximal sensing and machine learning algorithms.

It uses IoT-based devices to collect data from New Brunswick and Prince Island. Following the data gathering, machine learning techniques like Linear Regression, KNN, Elastic-Net, Support Vector Regression, and Linear Regression are employed. The SVR approach offers the highest accuracy when compared to the other methods. It was able to explain how internal and external factors like climate, environment, and other variables affect agricultural productivity. It was put to the test using a small dataset. If it's tested on a much larger dataset, it might be more dependable.

The performance of the autoregressive integrated moving average (ARIMA), partial least square (PLS), artificial neural network (ANN), and PLS combined with response surface methodology (RSM), dubbed RSMPLS, was tested. RSMPLS can be used to create a non-linear relationship between prior prices. The developed service is also linked to the smart agro management platform, allowing users to retrieve previous pricing and estimate future prices. In the model, only one parameter is employed (prices). More variables, such as climate, market location, and planting space, will result in more precise results.

Demand based crop recommender system for farmers.

The Decision Tree algorithm is used to classify crops and make recommendations to farmers. The dataset only included data from the prior year to make the results more relevant. They also designed a user-friendly GUI for text-to-speech and vice versa conversion using nlp algorithms. The GUI with nlp algorithms will be quite beneficial to farmers. Because the dataset will be updated to the most recent year's data, the results will be more relevant. When determining the crop, there are insufficient aspects to consider. To acquire more accurate results, more features are required.

Decision Making Support System for Prediction of Prices in Agricultural Commodity

To aid farmers in projecting prices, a decision-making help model has been developed. This design incorporates a gateway via which farmers can access their accounts by logging in with their credentials. Farmers must provide the name of the commodity as well as the previous selling price of the crop. This application will be able to provide normal crop pricing based on previous prices, which will aid farmers in making better decisions and forecasting prices. They're predicting crop prices with inefficient data since they're using previously reported in prices from consumers and effectively taking an average. This model will not always generate accurate and exact results because no extra parameters are taken into account.

An efficient analysis of crop yield prediction using Hadoop framework based on random forest approach.

The Random Forest Algorithm is integrated with MapReduce programming on the Hadoop system. Will increase the scalability of the process by allowing for more efficient data management. There may not be a modular and platform-independent solution for crop yield forecast difficulties.

METHODOLOGY

Tools used: The two most crucial requirements are Python and Flask. Python is a high-level programming language that is used to define system models. As a result, you'll require a Python 3.5+ environment that includes the sklearn, numpy, and pandas libraries, as well as a fully functional Python 3.5+ environment. The flask html-python framework is also required. Flask is a web framework written in the Python programming language. It doesn't have its own set of tools or dependence on libraries.

Users can obtain and view the algorithm's results through the website. The approaches listed below are being used to develop the most interactive website possible, where users may learn about the amount of land available, the amount of land used for cultivation, the number of farmers who use the land, and the various crops that have been cultivated. Using the social media links at the bottom of the website, farmers can also communicate with one another. They can reach out to agricultural specialists by dialling the hotline number or visiting the website; there is also a mobile app they can use to receive answers to their questions. This is a mechanism that helps a farmer keep track of what's going on in the present.

Process: The current research examines supervised learning models, which are a type of machine learning. The system architecture defines the conceptual model in terms of diverse structures and multiple viewpoints of the system. The proposed framework can be used to calculate the crop price. This model shows a vast amount of data being captured and preprocessed to remove unwanted data like NULL and absurd numbers. During the preprocessing stage, we separated the dataset into training and testing datasets. Train the dataset to detect the crop price in the dataset using appropriate supervised learning algorithms. Find crop prices using machine learning methods for each new data point added to the dataset. After gathering data, a suitable machine learning technique must be employed to compute the results. We applied a number of machine learning algorithms, including random forest, polynomial regression, and decision trees, to improve the model's efficiency and capabilities. Accuracy and precision will be calculated for the proposed model. This system design includes flow data, machine learning techniques, and modules for detecting crop pricing and feature selection.

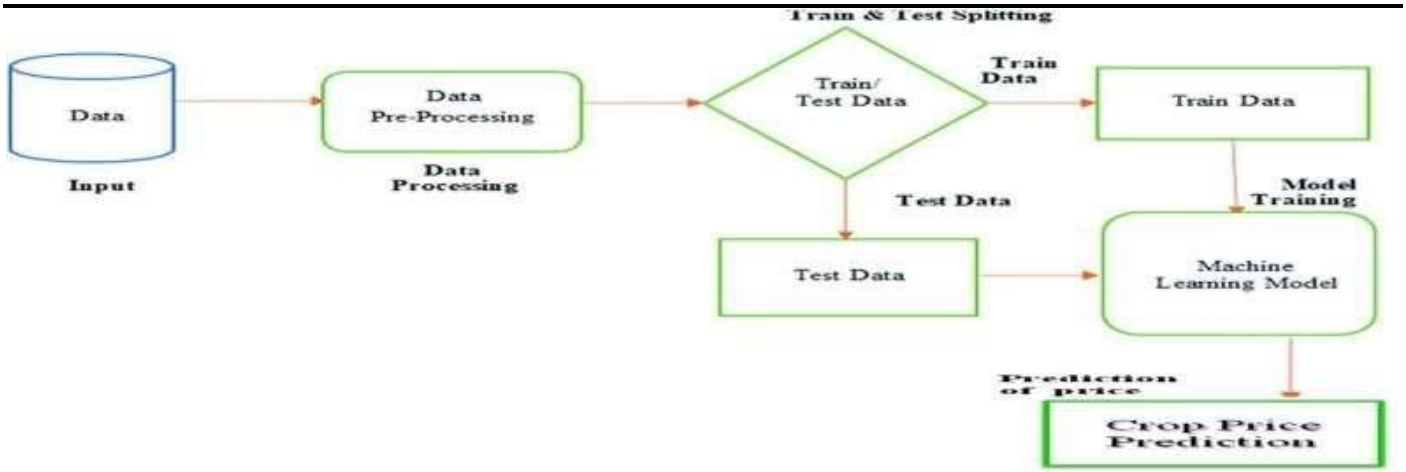

Fig 1. System Architecture

The following implementation modules are a big part of the research work.

1. Data Gathering
2. Exploration of Data
3. Prediction using machine learning
4. Web-based software

Data Gathering

The dataset was created using agricultural data taken from a public repository. There are a number datasets that can be used to gather data. We were able to obtain information on the characteristics of each crop's rainfall.

Exploration of Data

Exploratory data analysis is a sort of data exploration. When a component gathers and analyses data, data analysis (EDA) is a key phase that must be done before any showcasing. This is because it is vital for a data research associate in nursing to virtually comprehend the data's notion while avoiding raising concerns. Knowledge inquiry is frequently quite useful in acquiring a hold on the structure of data, the appropriation of qualities, and therefore the distance of special features and interrelationships within the informative index.

Prediction using machine learning

The Decision Tree is one of the most extensively used and useful models for supervised learning. It can be used to solve regression and classification problems, albeit the latter is more common. This tree-structured classifier has three types of nodes. The Root Node is the graph's first node, and it represents the entire sample. It can be subdivided further into nodes. The internal nodes reflect the features of a data collection, while the branches represent the decision rules. Finally, the Leaf Nodes represent the outcome. This method is quite helpful in resolving decision-making problems.

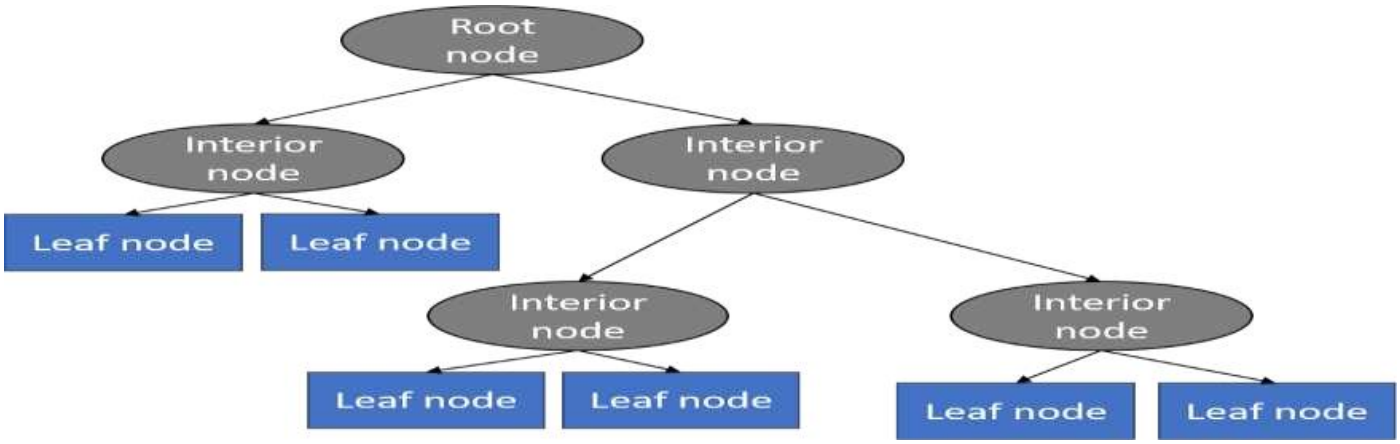

Fig 2. Decision tree model

By answering True/False questions, a specific data point is travelled through the entire tree until it reaches the leaf node. The final forecast is the average of the dependent variable's value in that leaf node. After many cycles, the Tree is able to predict an acceptable value for the data point.

STEPS TO IMPLEMENT THE ALGORITHM

Step 1:- Produce a dataset with training data for rainfall and the wholesale pricing index

. **Step 2:-**Select all the rows and column 1from dataset to "x" Which is independent variable

Step 3:-Select all of the rows and column 2 from dataset to "y" Which is dependent variable

Step 4:- Fit decision tree regress or to the dataset

S**tep 5:-**Predict the new value **Step 6:-**Visualize the result and check the accuracy

RESULT AND DISCUSSION

The decision tree is one of the most powerful complete regression algorithms accessible, capable of successfully and efficiently working with very small datasets. The criterion function is used to create the tree. One of the criteria used is the entropy, often known as the impurity measure. In classification, entropy is the most widely used impurity measure or splitting condition. It is defined as follows:

For a given node t, p(i|t) denotes the proportion of samples that belong to class c. If all samples at a node belong to the same class, the entropy is 0; otherwise, it is maximal.

$$E(t) = -\sum_{i=1}^{c} p(i|t) \log_2 p(i|t)$$

we have a uniform class distribution.

Three equations are used to evaluate the performance of decision tree regression:

1. Mean absolute error

2. The squared error of the error

3. R2 score

The accuracy of regression methods is calculated using the three equations above. Different algorithms were developed and tested, with the results compared to see which one was the most effective. The decision forest approach generated the best results for the majority of crops, despite the fact that both the choice tree regression and the random forest algorithms performed well. The random forest algorithm, on the other hand, only worked for a few algorithms. As a result, decision tree regression is utilised to anticipate prices, which aids farmers in making decisions with a 95.4 percent accuracy.

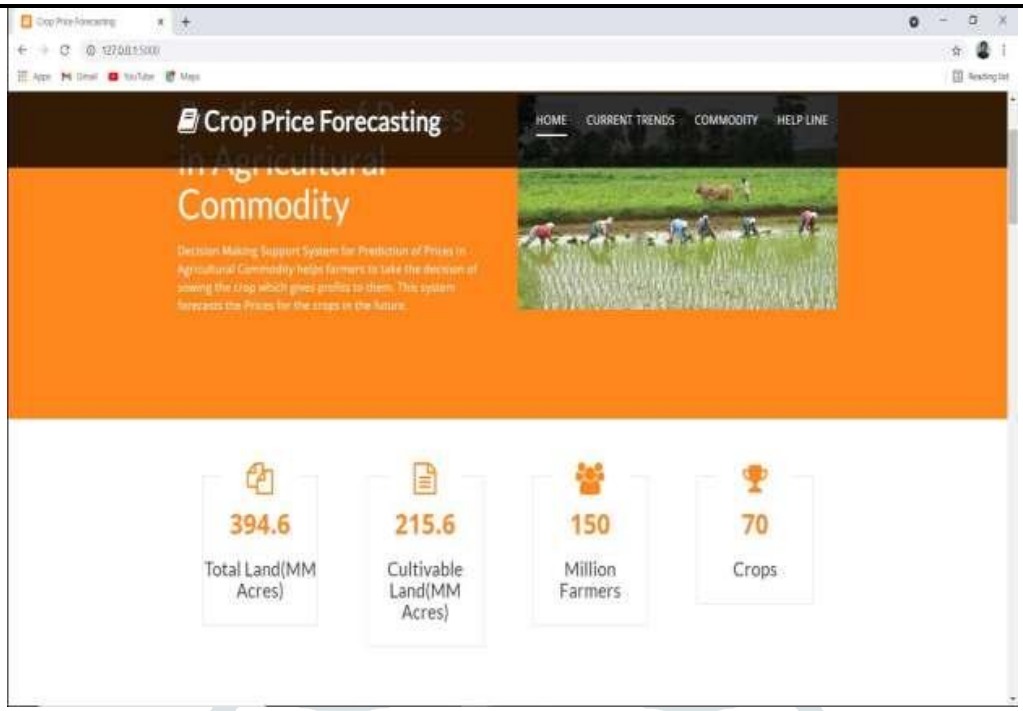

Fig 3. Home page of the web site with the amount land and crops information.

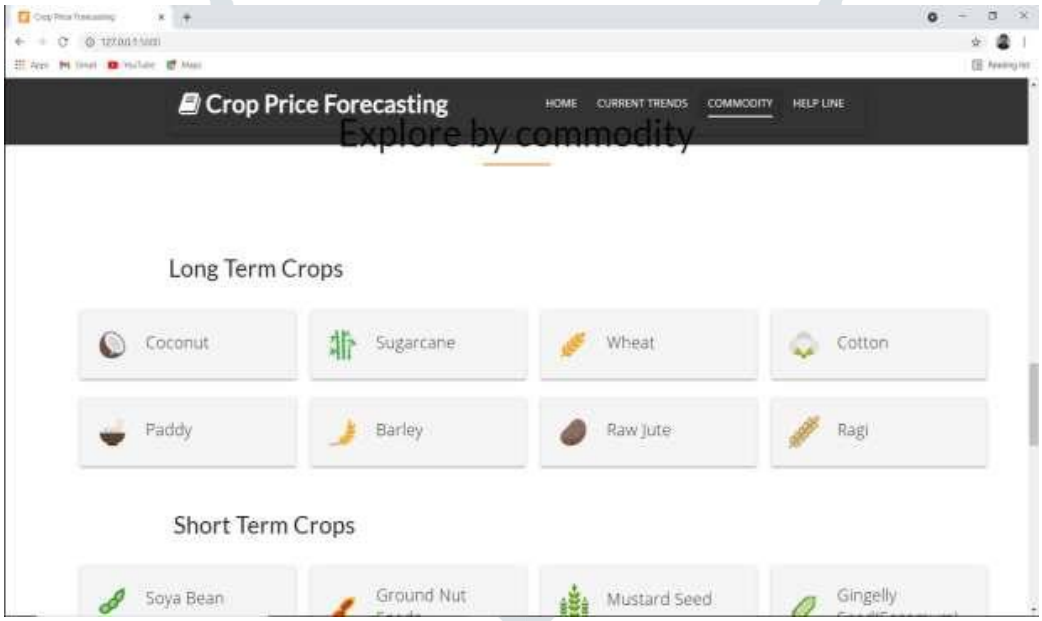

Fig 4. Long Term and Shot term crops list are displayed from which users can select a crop.

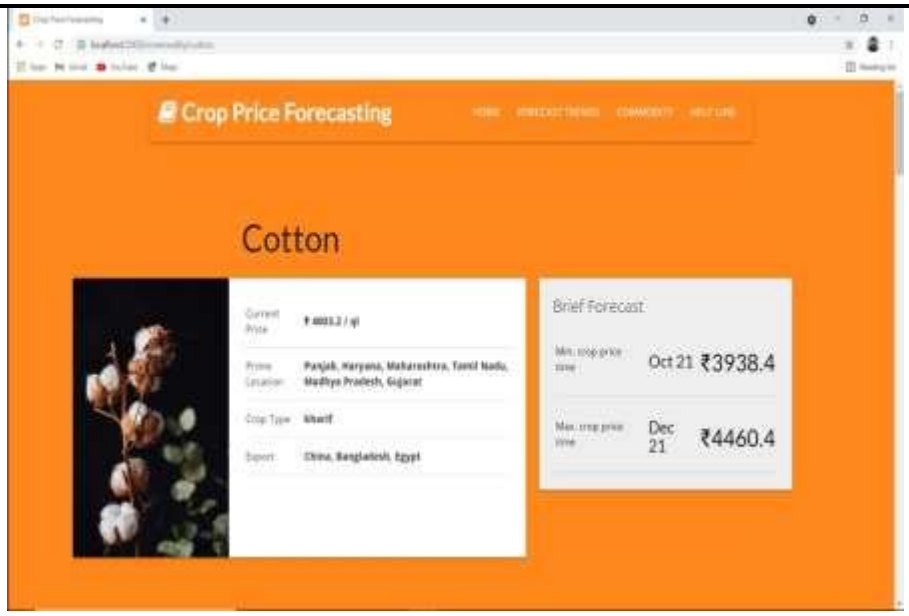

Fig 5. The crops details like minimum and maximum price and location were displayed.

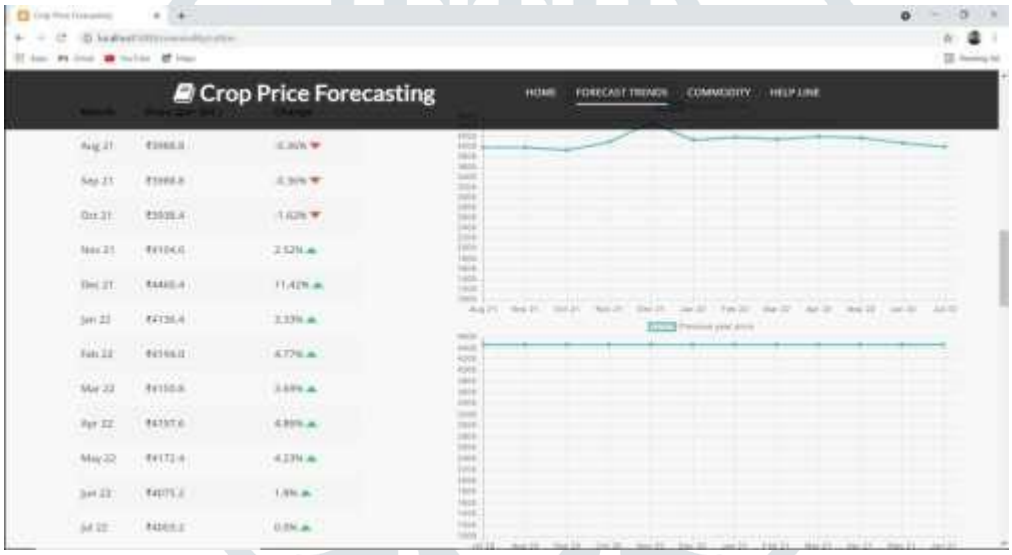

Fig 6. The forecasting trends of the crop like its next 12 months price and the graphs that shows the statstics of the price for present and next 12 months.

Fig 7. The help line details and the social media links were provided here.

The model receives no input from the user. It simply provides information such as current trends, as well as the most profitable and least profitable crops. It also includes details about the most significant commodity crops. The crops can be looked at individually, with a list of long-term and short-term crops available. Farmers can choose a crop, then navigate to the commodity to see information

such as the maximum and lowest price per quinta, the area of growth, and the destinations for the products. They can see the 12-month forecast alongside graphs that display statistics for the current year and the following year.

CONCLUSION & FUTURE SCOPE

The proposed model was developed to help farmers make better decisions about which crops to plant, when to plant them, and where to plant them. Even before the growing process begins, our system anticipates the price of the crop of choice, delivering crucial information to the farmer. A variety of methods, such as decision trees, neural networks, and SVM, can be used to anticipate prices. In our model, we use a decision tree. It has a high level of accuracy and has been trained on a variety of kharif and ragi crops (such as paddy, arhar, bhajra, barley, and others).

The project makes use of a web application to estimate crop prices and forecasts, which is based on effective machine learning techniques and technologies and has a user-friendly interface. The training datasets obtained so far provide enough data to forecast market pricing and demand. To correctly measure the accuracy of each system employed, the root mean square error is obtained for each technique, and the most accurate system is then picked. Farmers will be able to make better decisions when bidding for better market prices for their crops using this strategy. As a result, the strategy aids farmers in reducing their difficulties and avoiding suicide attempts.

REFERENCES

1. https://www.educba.com/machine-learning-algorithms/

2. Kaur, Manpreet, Heena Gulati, and Harish Kundra. "Data mining in Agriculture on crop price prediction: Techniques and Applications." International Journal of Computer Applications 99.12 (2014): 1-3.

3. Ramesh, D., & Vardhan, B. V. (2013). Data mining techniques and applications to agricultural yield data. International journal of advanced research in computer and communication engineering, 2(9), 3477-3480.

4. S. Kulkarni, S. N. Mandal, G. S. Sharma, M. R. Mundada and Meeradevi, "Predictive Analysis to Improve Crop Yield using a Neural Network Model," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Bangalore, 2018, pp. 74-79.

5. Sadiq A Mulla, S. A. Quadri. "Crop-yield and Price Forecasting using Machine Learning." International journal of analytical and experimental modal analysis. August 2020.

6. S. Kulkarni, S. N. Mandal, G. S. Sharma, M. R. Mundada and Meeradevi, "Predictive Analysis to Improve Crop Yield using a Neural Network Model," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Bangalore, 2018, pp. 74-79.

7. S. V. Bhosale, R. A. Thombare, P. G. Dhemey and A. N. Chaudhari, "Crop Yield Prediction Using Data Analytics and Hybrid Approach," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, 2018, pp. 1-5.

8. T. Siddique, D. Barua, Z. Ferdous and A. Chakrabarty, "Automated farming prediction," 2017 Intelligent Systems Conference (IntelliSys), London, 2017, pp. 757-763.

9. S. Bhanumathi, M. Vineeth and N. Rohit, "Crop Yield Prediction and Efficient use of Fertilizers," 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2019, pp. 0769-0773.

10. S. V. Bhosale, R. A. Thombare, P. G. Dhemey and A. N. Chaudhari, "Crop Yield Prediction Using Data Analytics and Hybrid Approach," 2018 Fourth International Conference on Computing