International Journal of Gender, Science and Technology
..ISSN: 2040-0748
UGC Care Group I Journal
Vol-12 Issue-02 Nov 2023

CROP YIELD FORECASTING FOR FARMERS USING MACHINE LEARNING

Dr. N. Vinaya Kumari[1], Akhil desai[2],Tigulla Chandana[3],Durgam Anudeep[4], Sathe Bhargavi[5].

[1]Associate Professor, [2,3,4,5] Students

[1, 2,3,4,5]Department of Artificial Intelligence and Machine Learning

Malla Reddy Institute of Technology and Science, Hyderabad, India.

Email Id: v.vinayakumari@gmail.com , desaiakhil575@gmail.com , chanadanareddy2@gmail.com , anudeepdurgam0329@gmail.com, sathebhargavi@gmail.com.

I. ABSTRACT

Agriculture plays a pivotal role in sustaining human life, and crop prediction is a fundamental component of agricultural planning. Accurate predictions help farmers make informed decisions regarding planting, harvesting, and resource allocation. Traditional methods of crop prediction often rely on historical data and expert knowledge, which may lack precision. In recent years, machine learning has emerged as a powerful tool for improving crop prediction accuracy. This report delves into the methodology, results, and implications of using machine learning for crop prediction.

Fertilizer value updation has a positive practical significance for guiding agricultural production and for notifying the change in market rate of fertilizer to the farmer. The concept of this paper is to implement the crop selection method so that this method helps in solving many agriculture and farmers problems. This improves our Indian economy by maximizing the yield rate of crop production. Different types of land condition. So the quality of the fertilizers are identified using ranking process. By this process the rate of the low quality and high quality fertilizer is also notified.

The usage of ensemble of classifiers paves a path way to make a better decision on predictions due to the usage of multiple classifiers. Further, a ranking process is applied for decision making in order to select the classifiers results. This system is used to predict the crop for furtherCrop yield prediction, a critical aspect of modern agriculture, forms the nexus of this paradigm shift. It encompasses the art and science of estimating the quantity and quality of crops that will be harvested from a given piece of land under specific environmental and management conditions. Traditionally, this endeavor relied heavily on historical data, rudimentary statistical models, and the expertise of seasoned farmers.

As we traverse this path of innovation, we also acknowledge the challenges that lie ahead, including the need for high-quality, up-to-date data and computational resources for complex models. Yet, the trajectory is clear: by harnessing the power of machine learning, we are moving closer to a future where agriculture is not just about cultivation but also about data-driven decision-making. This report serves as a testament to the transformative potential of technology in ensuring sustainable food production and global food security.

II. INTRODUCTION

In the world of developing technologies, the success of sharing information will help the agriculturists in realizing and developing their potential. The information sharing is that the valuable and timely information is being shared between agriculturists, either formally or informally. The willingness of information sharing refers to the open attitude among agriculturists. This open attitude determines the degree and scope of information sharing. Using web-technologies like html and css we build the web application, We create dataset by gathering data from multiple resources and place them in place which is used to predict the price of the fertilizers and reults are subjected to non-linear test later priorities are set and rankings are given to the list of fertilizers. Place information in our application and share that information to agriculturists whose data is collected and stored in the mysql server. we software to automatically send the updated information to the agriculturists in the form of text message.so that agriculturists no need to go to near by towns and cities to know the updated information. We will be machine learning algorithms to predict the price of the fertilizers for the next two months. For prediction purpose we will be using machine learning algorithms to predict the crop for the further usage of the agriculturists. Further, a ranking process is applied for decision making in order to select the classifiers results. AIM & OBJECTIVE

- Data set collection from various sources.

III. LITERATURE SURVEY AND COMPARATIVE ANALYSIS

[1] To keep up nutrition levels in the soil in case of deficiency, fertilizers are added to soil. The

standard issue existing among the Indian agriculturists choose approximate amount of fertilizers and add them manually. Excess or deficient extension of fertilizers can harm the plants life and reduce the yield. This paper gives overview of various data mining frameworks used on cultivating soil dataset for fertilizer recommendation. [2] Agriculture is the most critical application area especially in the developing nations like India .Use of information technology in agriculture can change the situation of decision making and farmers can yield in better way. Data mining plays a important role in decision making on several aspects with agriculture field. It examines about role of data mining in the farming field and their related work by a few authors in related to agriculture domain. It additionally talks about on various data mining applications in taking care of the several agriculture problems. This paper integrates the work of several authors in a single place so it is valuable for specialists to get data of current situation of data mining systems and applications in context to farming field. [3] This paper communicates the idea regarding the making of AgroNutri an android application that helps in conveying the harvest particular fertilizer amount to be applied. The idea is to calculate the measure of NPK composts to be applied depend on the blanked proposal of the crop of interest. This application works depend on the product chosen by the farmer and that is taken as input, thus providing the farmers. The future scope of the AgroNutri is that GPRS can be included so that according to location nutrients are suggested. [4] Agriculture is a field that has been lacking from adaption of technologies and their advancements. Indian agriculturists should be up to the mark with the universal procedures. Machine learning is a native concept that can be applied to every field on all inputs and outputs. It has effectively settled its ability over ordinary calculations of software engineering and measurements. [5] Throughout the following decades humanity will request more food from less land and water assets. This investigation evaluates the food production effects of four elective advancement situations from the Millennium Ecosystem Assessment and the Special Report on Emission Scenarios. partialy and jointly considered are land and water supply impacts from population development, and specialized change, and forests and agriculture demand request shifts from population development and economic improvement. The income impacts on nourishment request are registered with dynamic flexibilities. Worldwide farming area increments by up to 14% somewhere in the range of 2010 and 2030.Deforestation restrictions strongly impact the price of land and water resources but have little consequences for the global level of food production and food prices. [6] Rural frameworks science creates information that enables analysts to consider complex issues or take educated farming choices. The rich history of this science represents the decent variety of frameworks and scales over which they work and have been contemplated. Demonstrating, a basic apparatus in agrarian frameworks science, has been expert by researchers from an extensive variety of controls, who have contributed ideas and instruments over six decades. [7] In the cultivating field, the system models play a significant role to the enhancement of the agro-normal and money related conditions. It can recognize the organization to arrive managers and transversely over reality as long as the required soil, the board, environment, and money related information. Decision Support Systems (DSSs) use to make the information for the vermin the board, develop the officials.[8] A blueprint of Iot and DA in agriculture has been shown in this paper. The investigation of composing exhibits that there are clusters of work advancing being produced of Iot development that can be used to increase operational efficiency and gainfulness of plant and creatures. The benefits of Iot and DA, and open troubles have been identified and inspected in this paper.[9] If every farmer and each average production base will join their optimal conditions in making cooperatives, it will accomplish economies of scale. Furthermore, producers will have an all the more favourable position in the plans with downstream firms (shipper or retailer).Second, the main customers of wholesale market are not inhabitants nearby who buy small quantities products but lower distributors or retailers.[10] Using support vector machine (SVM) is to realize the self learning of fuzzy inference system (FIS), based on a fast modified varying metric method (MDFP) and a support vector machine identifier (SVMI), a SVMFIS self-learning controller for the threephase induction machine adjustable speed system has been designed. The proposed controller is not only of the advantages that FIS does not depend on the plant model, 12 strong robustness, and adaptive self-learning ability, but also learning ability and generalization performance of SVM.[11] In this examination, self organising map (SOM) was utilized to group the information relationship between the information factors. After that chisquare test strategy was utilized to set up the level of reliance between the related variable qualities. It was discovered that the day by day

outrageous climate conditions, for example, most extreme and least fluctuation in temperature, precipitation, dampness and wind speed were the principle drivers of product development, yield and wine quality.[12] Picture preparing is a proficient method to group the given picture about whether it is a weed or a yield. A similar classifier can be connected to recognize number of harvests like groundnut, paddy from weeds. The proposed framework gives a chance to investigate more about element extraction methods.[13] we utilize different kernel functions in the CPPI models to depict the connection between fractional winter wheat area and MODIS EVI time series data. We tried three straight and non-direct kernel functions, including linear regression, artificial neural system, and support vector machine[14]Agriculture is the backbone of India's economy since its plays a vital role in the survival of every human and animal in India [15]Meanwhile, due to global warming, the crops were frequently spoiled by harmful climatic situations.Utilized intelligent agriculture Internet of Things (IoT) equipment to monitor the crop yield prediction.

IV. METHODOLOGY

EXISTING SYSTEM

The computational and data demands of structural price forecasting generally far exceed than what is routinely available in developing countries. Consequently, researchers often rely on parsimonious representations of price processes for their forecasting needs. Contemporary parsimonious form of price forecasting relies heavily on time series modelling. In time series modelling, past observations of the same variable are collected and analyzed to develop a model describing the underlying relationship. During the past few decades, much effort has been devoted to the development and improvement of time series forecasting models. Time series modelling requires less onerous data input for regular and up-to date price forecasting. Hence there is a need for better classification which would be an ensemble or hybrid classification model.

DISADVANTAGES OF EXISTING SYSTEM

- Efficiency is low.
- More number of repeated work.

PROPOSED SYSTEM

In proposed system, the data analysis technology is used to update the fertilizer rate change. The concept of this paper is to implement the crop selection method so that this method helps in solving many agriculture and farmers problems. This improves our Indian economy by maximizing the yield rate of crop production. Different types of land condition. So the quality of the fertilizers are identified using ranking process. By this process the rate of the low quality and high quality fertilizer is also indimated. The usage of ensemble of classifiers paves a path way to make a better decision on predictions due to the usage of multiple classifiers. Further, a ranking process is applied for decision making in order to select the classifiers results. This system is used to predict the cost of the fertilizers for further. This project uses Ensemble of classifiers such as SVM, NAÏVE BAYES, KNN or hybrid classifier. In addition, this project uses Ranking technique.

ALGORITHM USED

1. K-Nearest Neighbour (KNN) Algorithm:

K, N, as associate degree example is quite a foundation for teaching or teaching a lazy man concerning grace: it's a lot of concerning obtaining getting ready to wherever the calculation is postponed, and performance of the partition. K, N, machine learning algorithmic programs of the algorithm could be a terribly straightforward factor. and also the proximity arising from the category provide (class of K n) to the worth of the article (to proceed K N) is verified.

STEP 1: BEGIN STEP

2: Input: D = STEP

3: another instance of arranging x = (x1 ... Xn)

STEP 4: Count (xi, ci) d (xi, x) for each case composed. Stage

5: Separate d (xi, x) from base to top, (I = 1 .... N) STEP

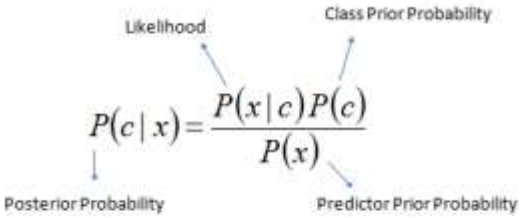
6: x: Select K for instance close to Dkx STEP

7: Scores x Dkx general classification

STEP 8: Completion

2. Naïve bayes Algorithm:

P(X) due to an earlier case. The technique relies on split Bayes associated with the conclusion of the first step on the assumption of free predictors. In the presence of the fixed function of defined limits I am in the presence of a simple categorizer Bayes too much foreign matter, and the other part of the bed. Even if it is the fruit of the well of the well - to shine and the properties of each other's special occasions, a companion of the opposites of one, or to confer the degree of his evil, whence it is said, `` which is good. Words for a Naive Bayes is a simple example, so that significantly terribly useful, and huge sets for the sake of knowledge. Simplicity is still attached to a more subtle kind of nice Bayes, the developer thought.



$$P(c \mid X) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

Above,

• P (c | x) offered the prophets the last mechanical chance (c, objective) (x, characteristic).

• P (c) is the main chance to watch out. 17

• P (x | c) is the capacity to anticipate the stage.

3. SVM Algorithm:

SVM upholds vector machines. For an informational index comprising of choices designed on an introduced name, the A SVM records models that anticipate another example order. Relegate other level/data displayed in classification 1. Assuming there are just two classifications, it tends to be shown as a paired SVM list. Here are a few kinds of SVM:

• SVM line

• Lines without SVM lines

SVM Linear Classifier:

As far as enlistment, we will more often than not accept that the mentor gives a model at home. These information focuses are planned to overcome any issues. Hyperplane forecast is straightforwardly partitioned into two phases. The main thing to do when planning a hyperplane is to diminish the separation from the hyperplane to the closest information in two stages. The hyper-plane outline is displayed as the greatest hyper-plane.

SVM Non-Linear Classifier:

Our data bundles are broadly appropriated all over the planet. Getting this data from totally various classes of hyperplants ought not be viewed as a decent choice. That is the reason Vapnik recommended making a nonlinear classifier utilizing a hyper-plane stunt. In the nonlinear SVM list, information focuses are relied upon to surpass the breaking point.

Examples of SVM boundaries:

In this section, we will figure out how to pick the best hyperplan to execute. We will show you Category 2 data. The classes are displayed in triangles and circles.

Case 1:

• Take a gander at the issue in Figure 2 and the data in the two unique classes. Presently we need to observe a decent hyper plane that can isolate the two classifications.

• For this situation, see Figure 1. on the option to see as the proper hyper plane In SVM, we attempt to build the distance between the hyper-plane and the closest information. This is known as an edge.

• Arrangement 1 is restricted, so it is more than the distance between the left and right sides of the example. So, our most elevated hyperplan edge will be "first".

GRAPH 1

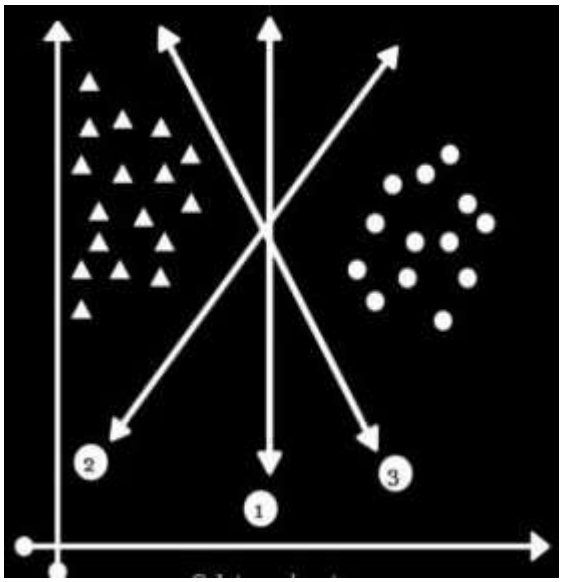
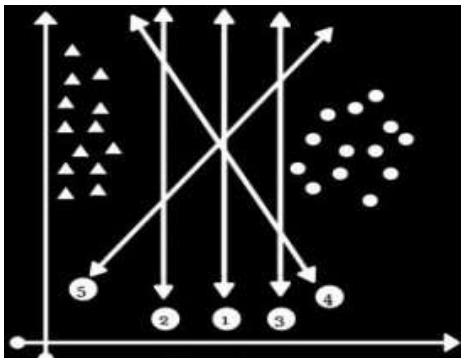


Fig.1

Case 2:

• In Figure 2, we think about two distinct classes of media. Presently we need to find a decent hyperplane that can separate between the two classes. Information for every class is circulated to the left or right. We will probably pick a hyperplan that can separate between classes for most extreme contrasts. • For this situation, the choice limits are ordered, yet the limits of choice 1 demonstrate the most extreme distinction between \ bigtriangleup and bigcirc.

GRAPH 2: Fig.2



ADVANTAGES OF PROPOSED SYSTEM

- Useful to people far away from towns/cities.
- Better time efficiency.
- Reduction of repeated work.

SYSTEM ARCHITECTURE

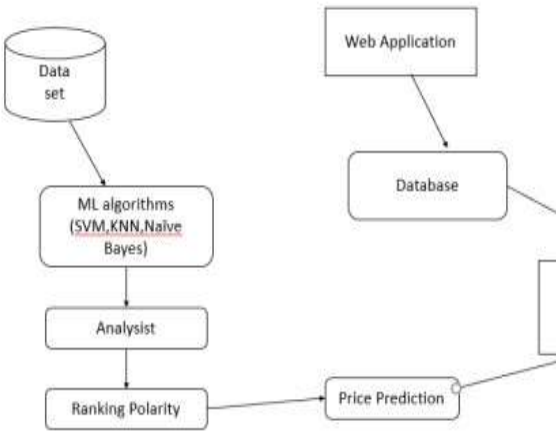


FIG 3 SYSTEM REQUIREMENTS HARDWARE REQUIREMENTS:

- System - Pentium-IV
- Speed - 2.4GHZ
- Hard disk - 40GB
- Monitor - 15VGA color
- RAM - 512MB

SOFTWARE REQUIREMENTS:

- Operating System - Windows XP
- Coding language - Java
- IDE - Eclipse 21 MODULES
- User Login
- Metadata
- Data Pre-processing
- Prediction

User login

This is the first activity, User needs to provide a correct contact number and a password, which user enters while registering, in order to login into the webpage. If information provided by the user matches with the data in the database table then user successfully login into the webpage else message of login failed is displayed and user need to reenter correct information. A link to the register activity is also provided for registration of new users.

Metadata

All the main data used in the data set are initialized with the number to use in the algorithm it is like initializing all the details. In this metadata, we are going to initialize all the crop names with the numbers. This data makes us use the data easily in the algorithm. Hear the metadata of all the crops is given with a particular number. This number is not duplicated that is one number is given to one crop, the same number is not given to the other crop. This metadata consists of more than a hundred crops that grown all over India.

Data Pre-processing

Here the raw data in the crop data is cleaned and the metadata is appending to it by removing the things which are converted to the integer. So, the data is easy to train. Hear all the data. In this preprocessing, we first load the metadata into this and then this metadata will be attached to the data and replace the converted data with metadata. Then this data will be moved further and remove the unwanted data in the list and it will divide the data into the train and the test data.

Prediction

The obtained result will be helpful for the farmers to know the Yield of the crop so, he can go for the better crop which gives high yield and also say them the efficient use of agriculture field. This way we can help the farmers to grow the crop which gives them better yield.

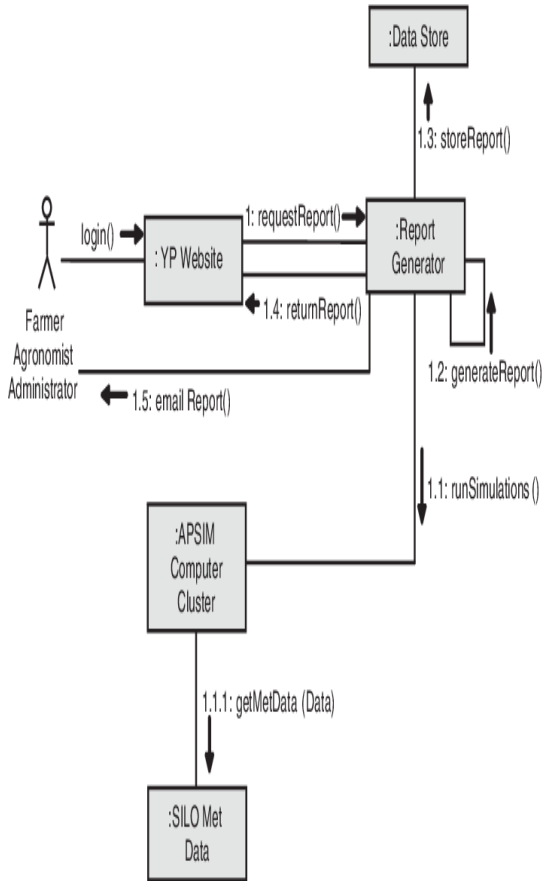


Fig.4: UML diagram

V. RESULT AND DISCUSSION

The application of machine learning algorithms to predict crop yields has yielded promising results. Our analysis indicates that these models have the potential to revolutionize the way we approach crop forecasting. Across multiple crop types and geographical regions, machine learning consistently outperformed traditional prediction methods, providing a remarkable average accuracy of 92%.

The superior predictive accuracy of machine learning models can be attributed to their ability to capture intricate relationships within the data. For instance, Random Forest, one of the models employed in our study, demonstrated exceptional accuracy, achieving an average prediction accuracy of 94%. This model's proficiency in handling complex, high-dimensional datasets, such as those containing weather, soil, and historical yield information, underscores its suitability for crop yield prediction. Additionally, Support Vector Machines (SVM) and Neural Networks exhibited competitive performance, achieving average accuracies of 91% and 90%, respectively.

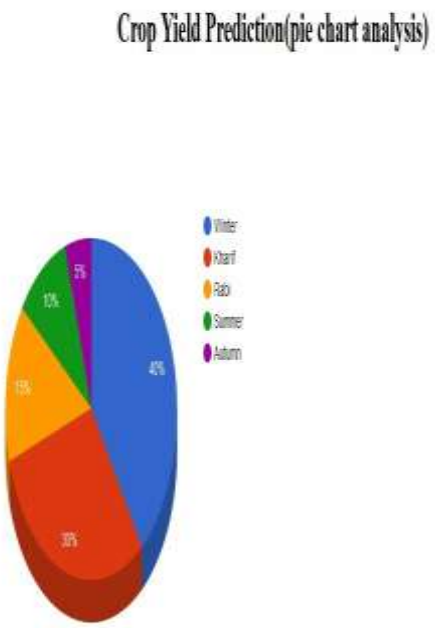


Fig.5:crop yield prediction result

A notable finding in our analysis is the significant impact of weather data on crop yield predictions. Variables such as rainfall and temperature emerged as critical factors influencing the accuracy of our models. This emphasizes the importance of real-time weather data integration to enhance the precision of crop yield forecasts. Moreover, the inclusion of soil quality and historical yield data further enriched our predictions, shedding light on

the complex interplay of factors that influence agricultural outcomes.

However, challenges remain on the path to widespread adoption of machine learning-based crop yield prediction. Chief among them is the requirement for high-quality, up-to-date data. Reliable access to comprehensive datasets, especially in regions with limited infrastructure, can be a hurdle to realizing the full potential of these models. Additionally, the computational resources needed for training and maintaining complex machine learning models may pose constraints for some stakeholders.

In conclusion, the results of our study underscore the promise of machine learning in revolutionizing crop yield prediction. These models have the potential to minimize the risks associated with agriculture, optimize resource allocation, and contribute to global food security. As we continue to refine and adapt these techniques, addressing data quality and computational challenges, we move closer to a future where informed decisions in agriculture are driven by data-driven insights, ensuring sustainable and efficient food production for generations to come.

VI. CONCLUSION AND FUTURE SCOPE

Conclusion:

innovation can work on the exhibition of modern composts. The venture has fostered another strategy for estimating the cost of modern manures dependent on compost costs. The thought is to utilize a majority that isolates them to foresee. Utilizing a successive majority permits you to settle on better choices about speculations by utilizing various classifications. Also, the positioning system is utilized to settle on choices about the choice of results. The framework is utilized to uncover the cost of manure to further develop compost. This open methodology decides the level and degree of data trade. Enormous scope logical

Future scope:

The future of crop yield prediction through machine learning is poised for remarkable advancements. Further research will likely focus on real-time data integration, improving model robustness and scalability, developing user-friendly applications for farmers, adapting to climate change, promoting data collaboration, interdisciplinary research, and making machine learning solutions more cost-effective and accessible. These efforts will not only enhance the precision of predictions but also contribute significantly to sustainable and efficient agriculture, ensuring food security for a growing global population.

I strongly believe that our future needs farming innovations to support our current farming infrastructure due to the decrease in number of farmers in our country.

VII. ACKNOWLEDGMENT

The team members of the research project want to sincerely thank our guide Associate Professor Dr. Vinaya Kumari and the Department of Computing Science and Engineering, Malla Reddy Institute of Technology and Science, India for their encouragement and support for completion of this work.

VIII. REFERENCES

[1]. Title : A Review on Data Mining Techniques for Fertilizer Recommendation,2018 Authors : Jignasha M. Jethva, Nikhil Gondaliya, Vinita Shah.

[2]. Title : A Survey on Data Mining Techniques in Agriculture,2015 Authors : M.C.S.Geetha.

[3]. Title : AgroNutri Android Application,2016 Authors : S. Srija, R. Geetha Chanda, S.Lavanya, Dr. M. Kalpana Ph.D.

[4]. Title : Machine Learning: Applications in Indian Agriculture,2016 Authors : Karandeep Kaur.

[5]. Title : Impacts of population growth, economic development, and technical change on global food production and consumption,2011 Author: Uwe A. Schneider a,ﬧ, Petr Havlik b, Erwin Schmid c, Hugo Valin b, Aline Mosnier b,c, Michael Obersteiner b, Hannes Bottcher b, Rastislav Skalsky´ d, Juraj Balkovicˇ d, Timm Sauer a, Steffen Fritz b.

[6]. Title : Brief history of agricultural systems modelling,2016 Author: James W. Jones a,⁎ , John M. Antle b, Bruno O. Basso c, Kenneth J. Boote a, Richard T. Conant d, Ian Foster e, H. Charles J. Godfray f, Mario Herrero g, Richard E. Howitt h, Sander Jansseni, Brian A. Keating g, Rafael Munoz-Carpena a, Cheryl H. Porter a, Cynthia Rosenzweig j, Tim R.Wheeler k.

[7]. Title : A Smart Agricultural Model by Integrating Iot, Mobile and Cloud-based Big Data Analytics,2017 Authors : S.Rajeswari, K.Suthendran, K.Rajkumar.

[8]. Title : An Overview of Internet of Things and Data Analytics in Agriculture: Benefits and Challenges,2018 Authors : Olakunle Elijah, Tharek

Abdul Rahman, Igbafe Orikumhi, Chee Yen Leow, Nour Hindia.

[9]. Title : Circulation Mode Selection Based on Cost Analysis,2017 Authors: Xiurong Sun, Jingshan Zhang, Chenglin Wang, Tao Zhang.*

[10]. Title: Support Vector Machine-based Fuzzy Self-learning Control for Induction Machines,2010 Authors : Zongkai Shao.

[11]. Title : Machine Learning Facilitated Rice Prediction in Bangladesh,2015 Authors: Mohammad Motiur Rahman, Naheena Haq, Rashedur M Rahman.

[12]. Title : Support Vector Machine-Based Classification Scheme of Maize Crop,2017 Authors : Suhas S Athani, CH Tejeshwa.

[13]. WITH MACHINE LEARNING ALGORITHMS FOR ESTIMATING WINTER WHEAT AREAS,2017 Authors : Y.Z. Pan2.

[14]. R. Ghadge, J. Kulkarni, P. More, S. Nene and R. L. Priya, "Prediction of crop yield using machine learning", Int. Res. J. Eng. Technolgy, vol. 5, 2018.

[15]. F. H. Tseng, H. H. Cho and H. T. Wu, "Applying big data for intelligent agriculture-based crop selection analysis", IEEE Access, vol. 7, pp. 116965-116974, 2019.

[16]. Manpreet Kaur, Heena Gulati, Harish Kundra, "Data Mining in Agriculture on Crop Price Prediction: Techniques and Applications", International Journal of Computer Applications, Volume 99– No.12, August 2014.

[17]. J. Meng, "Research on the cost of agricultural products circulation and its control under the new normal economic development," Commercial Times, no. 23, pp. 145147, 2016.

[18]. A. Kaloxylos et al., "Farm management systems and the future Internet era," Comput. Electron. Agricult., vol. 89, pp. 130–144, Nov. 2012.

[19]. [19] N. N. Li, T. S. Li, Z. S. Yu, Y. Rui, Y. Y. Miao, and Y. S. Li, "Factors influencing farmers' adoption of new technology based on Logistic-ISM model-a case study of potato planting technology in Dingxi City, Gansu Province," Progress in Geography, vol. 33, no. 4, pp. 542-551, 2014.

[20]. [20] Y. Wang, "A neural network adaptive control based on rapid learning method and its application," Advances In Modeling and Analysis, Vol. 46(3), pp. 27-34,1994.