ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

A Survey on Predicting Suitable Crops for Cultivation Using IoT

Shridhar Mhaiskar[1], Chinmay Patil[2], Piyush Wadhai[3], Aniket Patil[4], Vaishali Deshmukh[5]

B. E Students, Dept. of Computer Engineering, Smt. Kashibai Navale College of Engineering, Pune, India[1234]

Asst. Professor, Department of Computer Engineering, Smt. Kashibai Navale College of Engineering, Pune, India[5]

ABSTRACT: The Internet of Thing (IoT) is playing a vital role in everyday life of human being. It is being used in almost every domain. The aim of the system is to predict a crop which can be cultivated in order to give maximum yield after harvesting. In order to get maximum yield it is very necessary that the crop which is going to be cultivated have suitable climatic and soil conditions. The proposed system works on two domains, IoT and Machine Learning. The system predicts the values of moisture, temperature and Ph value from the historical data which is generated by respective sensors. Crop production depends on various factors which includes biological characteristics, environmental conditions, soil attributes. Applying appropriate data mining techniques on data collected from sensor devices will be used for the recommendation of the suitable crops for cultivation.

KEYWORDS: IoT, Machine learning, prediction, K means, ARIMA Model, KNN algorithm, Precision agriculture, Raspberry pi3, sensors.

I. INTRODUCTION

Precision agriculture is an upcoming field which provides an alternative for traditional crop selection techniques. Traditional crop selection method mostly relies on farmer's crop instincts and intuition's. Combining IoT and data mining with traditional method would increase the efficiency of crop selection and thus will improve the efficiency of selecting crop**.**

Agriculture plays a vital role in Indian economy. [1] Agriculture is considered as a primary means of livelihood for about 58% of the rural India. The green revolution which introduced various high yielding seeds and fertilizers undoubtedly leaded to increase in crop productivity. However, for the past 20 years scientific contribution in fields of agriculture is low compared to the technological inventions in services and manufacturing industries. [2] Agriculture is now currently 15% of GDP as per Government of India Statistics.

Indian farmers still follow the traditional way for selecting crops for cultivation which was passed onto them by their ancestors. There is no proper guidance available to assist them for cultivating appropriate type of crop using modern technologies. Thus using various data mining techniques, the proposed system provides the end user with a variety of crops suitable for cultivation. It is a Cross domain system which uses IoT for collecting temperature value, moisture value and Ph value from different sensors further uses data mining algorithms such as k-means clustering and KNN to predict the type of crop which is appropriate for the given conditions**.**

II. LITERATURE SURVEY

[3] Proposes three recommendation system based on past data. Since efficiency of random forest algorithm is higher than naïve bayes and ID3, it uses random forest algorithm to predict appropriate crop based on current NPK value of soil. However random forest algorithm doesn't deal with large number of categories in categorical variable. Another crop rotation recommendation system is described in this paper which uses FP tree. Paper also provides recommendation for appropriate fertilizers using sufficiency method. However this method would put additional pressure on soil to match its fertility with required conditions thus degrading the soil condition.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

[4]Proposes a system which is a recommendation engine that not only makes recommendations based on type and quantity of nutrients present in the soil, weather conditions and irrigation facilities present in the farm but also suggests the type and quantity of fertilizer to be used. The data is collected though the sensors present in the farm and the data on weather is obtained from the weather institute with the wireless sensor network. The system uses statistical predictive modeling for prediction and recommendation of crops. The advantage of this system is that it is easily available and uses gsm network for communication which is cheap and available to farmers nationwide.

[5] Proposes the development of a mobile application which recommends crops to be cultivated in a year in such a way that the fertility of soil will be maintained. For this the previous crops cultivated are also taken into consideration so that the next crop suggested will be different following the crop rotation principle. The system is highly efficient as it uses random forest algorithm which has an efficiency of 80% compared to that of " naïve bayes" which is 40% and that of "id3" is 60%.Apart from just recommending the crops the system provides recommendation for fertilizers and an online purchase system using apriori gen algorithm for frequent item sets so that the soil quality maintenance will be simplified by the use of appropriate fertilizers.

[6]proposes recommendation system for predicting suitable crops based on NPK values of soil .The system provides the notification using SMS service which include the recommended crop. Naive bayes algorithm is used to predict most suitable crop. However efficiency of naïve bayes is less compared to KNN algorithm. The system forces the farmers to provide input to system manually in contrast to proposed system which collects data automatically and provides the result on single login. The system also provides only single crop which is less compared to proposed system which provides with a list of N different suitable crops.

[7]ARIMA model is autoregressive integrated moving average model which was basically designed for predicting the stock prices in future. Has used the ARIMA model for tropical cyclone prediction using the historical data i.e. from (2007-2011).The proposed model in is made up of two entities firstly it creates the model for time series, and secondly it applies the generated model to produce forecast.

Paper	Parameters Considered	Data considered for prediction	Algorithms proposed	Architecture	Devices Required by end User
[3]	Year of cultivation, market price, output ratio	Training dataset :year 1998 to 2009 Test dataset: year 2010	1.Random forest 2. FP tree	Multitier/N-Tie- Data Tier Business Tier Presentation Tier	NPK sensors, computer for WebApp
[4]	NPK values, temperature, humidity, moisture, ph	Live sensor values , historical data	1.Statistical predictive modeling	Client server ,web app and mobile application	NPK sensors ,raspberry pi, temperature ,moisture, Ph
[5]	NPK values, past purchase	training dataset: past NPK value Past purchases, items purchased together.	1.Random forest algorithm 2.apriory algorithm.	Client server mobile application	NPK sensors, raspberry pi ,Android mobile.
[6]	Nitrogen, potassium, Sodium(N,P,K)	Current Value of soil condition	1.Naïve Bayes	Client server ,WebApp	N,P,K sensors, Raspberry pi, computer for WebApp
Proposed system	Temperature, Moisture content, Ph content(T,M,P)	Historical data stored in DB	1.Linear regression 2.k-means 3.knn algorithm	System uses web api	Temperature,Moisture,Ph sensors, Raspberry pi, computer to access WebApp.

Table 1: Literature Survey Table

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

III. PROPOSED ARCHITECTURE

Figure 1 shows the proposed system architecture of recommendation system. The sensors are connected to raspberry pi3 using wireless connection. Three types of sensors are used to collect biological as well as atmospheric data. Moisture sensor SHT-10 denoted by MS is used to collect moisture reading of soil. Temperature sensor DHT-22 denoted by TH is used to read current temperature and humidity of the atmosphere. The Temperature sensor DHT22 is a basic, low-cost digital temperature and humidity sensor. It uses a capacitive humidity sensor and a thermostat to measure the surrounding air, and spits out a digital signal on the data pin Ph sensor CS526 probe denoted by PS is used to measure current Ph ratings of soil. Raspberry pi3 model is basically used to connect various sensor using wireless adapter. Thus N numbers of sensors which are located in field are connected to centralized raspberry pi 3 wirelessly. Sensor data is collected from raspberry pi3 and stored into intermediate server using mqtt protocol.

Intermediate server consists of Mysql database which stores the collected data in structured format. Further this collected data is sent to main data server which stores data to predict top N suitable crops for cultivation.. Authentication process is performed to allow access to user. On successful login user obtains list of suitable crops for cultivation. The proposed system works on algorithms such as ARIMA model for prediction of values received from temperature, Moisture and Ph Sensors. The ARIMA model takes the values from database as input and then predicts what will be the value of that particular parameter after 1 month. The ARIMA model does this for all the three parameters separately that is for temperature, moisture and Ph value and predicts the value for each. Then the predicted values are sent to K means algorithm for classification based on Ph value thus creating k clusters of crops having similar Ph value. KNN algorithm is used to predict top N suitable crops which are displayed to user.

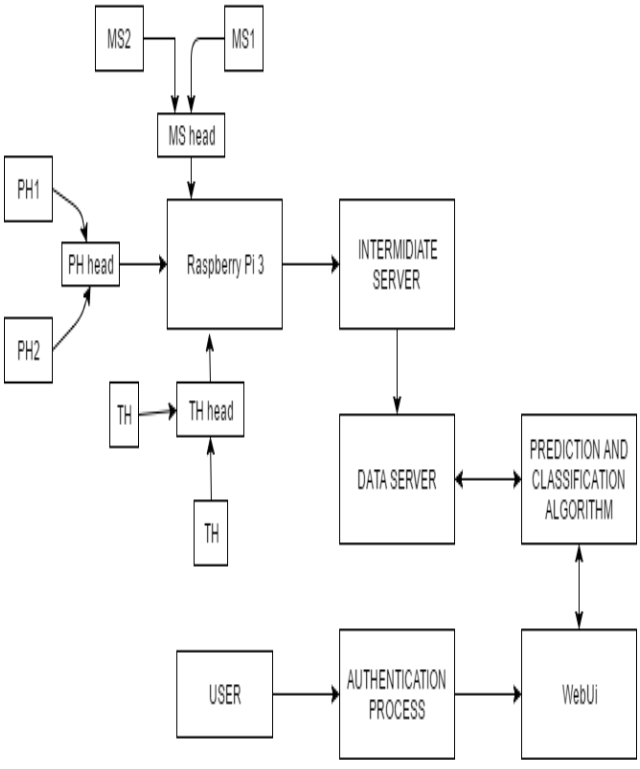

Figure 1: Proposed system architecture

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

IV. SYSTEM EXECUTION FLOW

STEP 1: Sensor data is collected using temperature, moisture and Ph sensors and sent to raspberry pi3.
STEP 2: Collected data is sent to intermediate server and stored in MySQL database.
STEP 3: Collected data in intermediate server is further sent to ARIMA model which predicts future value.
STEP 4: Authentication process is used to allow access to legitimate user.
STEP 5: Prediction and classification algorithm is performed on collected data.
STEP 6: Top N recommended crops are

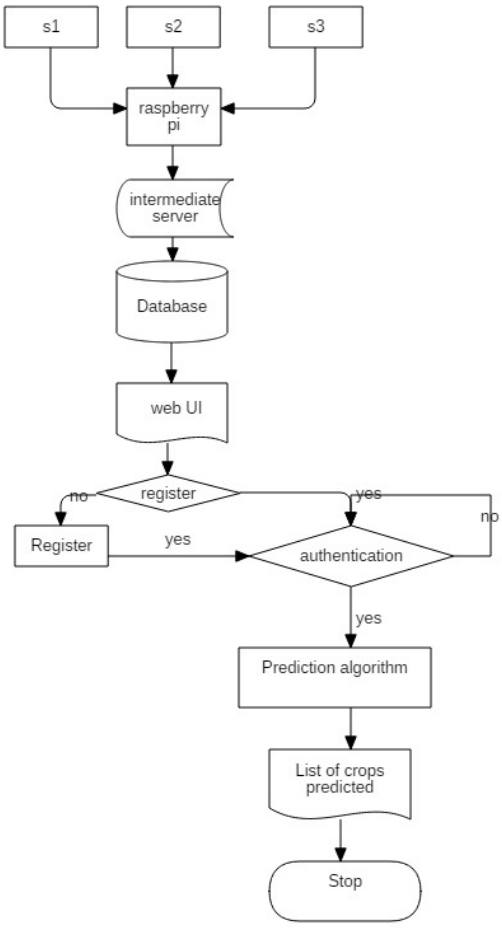

Figure 2: System Execution Flowchart

V. ALGORITHM

K-MEANS:
[11]K means algorithm is popular as a typical clustering algorithm. It uses information appeared in the data set to do clustering .Accuracy can be improved by selecting value of K according to the number of available labeled data in dataset.
INPUT: Data set X and assigned cluster number K

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

OUTPUT: K clusters C1, C2…..CK
STEP 1: Select K data objects as the initial clustering centers c1, c2,….cK randomly
STEP 2: According to the Euclidean distance formula, calculate the distances between each data object and each clustering center and assign each data object to the cluster whose distance is the smallest.
STEP 3: Recalculate the clustering centers of each cluster
STEP 4: If the clustering centers do not change or the iterations has reached the assigned iterations, the algorithm end. If not, turn back to Step 2 and go on.
KNN (K NEAREST NEIGHBOR):
[10] K-Nearest Neighbor (KNN) is used to develop a system that uses numeric historical data to determine crops which are suitable for current soil condition. Euclidean distance formula is used to calculate top K closest neighbors.
INPUT: Selected cluster
OUTPUT: List of k suitable crops.
STEP 1: A cluster is selected such that distance between soil's current Ph value and centroid is smallest.
STEP 2: Euclidean distance formula in used on temperature, moisture and Ph value.
STEP 3: K smallest values are considered.
STEP 4: List of K suitable crops is obtained.

VI. FUTURE WORK

More number of parameters can be introduced to the system such as Nitrogen, Phosphor, potassium so that the accuracy of the system can be increased. Also analyzing the soil condition will help in recommending the fertilizer to increase the soil fertility. One of the limitations of the system is predicting temperature. It can be predicted, but in that case more powerful algorithms and sensors will be required to implement which will increase the cost of whole system. In order to avoid this, in future the system can be merged with the governmental metrological department to get the input of the predicted temperature value by the metrological department. [9] Farms can be automated using raspberry pi and embedded Linux system which provides access to user for irrigation situated remotely. [8]A large amount of uncertainty might exist in agriculture during prediction thus fuzzy logic can be implemented for better precision. Location based prediction can be used to predict crops which are suitable for specific region.

VII. CONCLUSION

Thus the proposed system uses various data mining techniques and IoT to provide list of possible crops for cultivation based on climatic factors such as humidity and atmospheric temperature and biotic factors which include soil moisture status and current Ph value of soil. The proposed system thus provides solutions for current soil condition and thus implicitly maintains soil fertility by rotating crops.

VIII. ACKNOWLEDGEMENT

We would like to acknowledge the support lent by our guide Prof. Vedashree Takalkar. She was always available to address our queries and lead us through the right direction. Her encouragement and faith in us proved to be extremely helpful for us. Her knowledge and experience came to rescue us from the times when we were stuck at a point. We thank her for guiding us.

REFERENCES

1. http://www.ibef.org/industry/agriculture-india.aspx
2. https://mkhari21.wordpress.com/2012/06/11/what-are-the-reasons-for-agricultural-growth-decline-in-indias-gdp/
3. Kiran Shinde, Jerrin Andrei, AmeyOke, *"Web Based Recommendation System for Farmers", International Journal on Recent and Innovation Trends in Computing and Communication 2014 ISSN: 2321 -8169 ,Volume: 3 Issue: 3,2014.*
4. Daryl H. Hepting, Timothy Maciag, Harvey Hill,*"Web-Based Support of Crop Selection for Climate Adaptation", 45th Hawaii International Conference on System Sciences, 2012.*

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

5. Lokesh.K, Shakti.J , Sneha Wilson , Tharini.M.S, *"Automated crop prediction based on efficient soil nutrient estimation using sensor network", National Conference on Product Design (NCPD 2016), July 2016.*
6. Mansi Shinde1, Kimaya Ekbote2, SonaliGhorpade3 ,Sanket Pawar4 ,Shubhada Mone5 ,*"Crop Recommendation and Fertilizer Purchase System" ,(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 7 (2) , 665-667,2016.*
7. A. Geetha and G. M. Nasira ,*"Time series modeling and forecasting: Tropical Cyclone prediction using ARIMA model", 3rd International Conference on Computing for Sustainable Global Development (INDIACom),2016.*
8. MohdNajib,MohdSalleh, *"A Fuzzy Modelling of Decision Support System for Crop Selection",2012 IEEE Symposium on Industrial Electronics and Applications (ISIEA2012), September 23-26, 2012.*
9. Pandurang H. Tarange, Rajan G. Mevekari, Prashant A. Shinde,*" Web based Automatic Irrigation System using wireless sensor network and Embedded Linux board" ,International Conference on Circuit, Power and Computing Technologies [ICCPCT],2015.*
10. Sagar S. Badhiye, Nilesh U. Sambhe, P. N. Chatur, *"KNN Technique for Analysis and Prediction of Temperature and Humidity Data", International Journal of Computer Applications (0975 – 8887)Volume 61– No.14, January 2013.*
11. Ye Hanmin, Lv Hao, Sun Qianting,*"An Improvised Semi-Supervised K-means Clustering Algorithm",IEEE,2016.*