Machine learning techniques for selecting the crop to increase the yield

Siddharth H. E.
sidduppilakshmi@gmail.com
S. J. C. Institute of Technology,
Chikkaballapur, Karnataka

Bharathi M.
bharathim5597@gmail.com
S. J. C. Institute of Technology,
Chikkaballapur, Karnataka

Navya G.
navyagowda255@gmail.com
S. J. C. Institute of Technology,
Chikkaballapur, Karnataka

Praveen Kumar L.
praveenkumarpraveen713@gmail.com
S. J. C. Institute of Technology,
Chikkaballapur, Karnataka

Harshavardhana Doddamani
hdoddamani@gmail.com
S. J. C. Institute of Technology,
Chikkaballapur, Karnataka

ABSTRACT

Agriculture is the backbone of our country for economic growth and it depends on selecting the crops, weather conditions, soil classification, marketing price, etc. In an economy where most of the produced food is from agriculture, selection of crops plays an important role. In light of the decreasing crop produce and shortage of food across the country which also has been a consequence of bad crop selection and thus, leading to increasing farmer suicides. So that we suggest one method which would help the farmers in selecting the crops in order to maximize the yield by summing up the analysis of all the affecting parameters. These affecting parameters can be economical, environmental as well as related to yield in nature. Selecting and maintaining suitable crops for maximum profit involves a sequence of tasks. These tasks should be carried out with basic knowledge and experience. The crop advisor expert system then suggests with a consultation with the farmer (through a graphical user interface) a suitable agricultural crop that can be grown in a land unit with reasoning.

Keywords— Economical growth, Selecting the crops, Weather conditions, Soil classification, Marketing price, Crop advisor, Demand of the crop

1. INTRODUCTION

Agriculture plays an important role in the economic growth of the country. Basically, the crops were depending on the soil classification and demand in the Market for the crop. Most of the farmers they know when to grow where to grow and how to grow but they don't know that grown crops will be getting the good yield so In our country population was more and dependency on agriculture is also increases then affect the economic growth. In this situation, we have to focus on increasing the crop yield rate. Before, if we want to grow a crop, we consider certain approaches like seed quality, crop hybridization, use of fertilizer etc. to increase the yield rate. But this approach is not enough so, here we used a method called crop selection method to increase the net yield rate of the crops. Many research is being carried out on agriculture planning where the ultimate motive is to obtain an efficient and accurate model for prediction of crop yield, classification of the crop, classification of soil, prediction of weather, prediction of crop disease.

Some of the widely used machine learning predicting techniques are artificial neural network, Support vector machines, random forest, k-nearest neighbours, linear regression and boosting techniques. Out of these techniques boosting techniques (eg. Regularized Greedy Forest, GBDT) are still not used in prediction crop yield.

In this research, a study on a few machine learning techniques are carried out and comparative analysis is presented.

1.1 Artificial Neural Network (ANN)

An artificial neural network is an information processing models that are inspired by the way biological nervous systems, such as the brain, process information. It uses processing of the brain as a basis to develop algorithms that can be used to model complex patterns and prediction problems. ANNs are considered as nonlinear statistical data modelling tools where the complex relationship between inputs and outputs are modelled or patterns are found.

ANN is a connected network where the processing units are interconnected which can be used for the further analysis.so the processing units takes the input from the earlier processing units and transfer output for other processing units. ANN is used for classification and prediction. Weights in ANN are generated such that the outputs depend on the actual inputs and the internal state of the network. Multilayered perceptron and deep learning are the widely used topological techniques of ANN to implement neural networks for prediction of crop yield.

1.2 Support Vector Machine (SVM)

In machine learning, support vector machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. The crop yield prediction which is used from the support vector machines is called support vector regression. The main objective of a support vector machines technique is to obtain the non-linear function using kernel function.

Using the linear function in large input sample space and optimization of complex problems are difficult, to avoid these difficulties support vector regression is used and the complex problems are transformed into simple linear function optimization.

1.3 K-Nearest Neighbors (K-NN)

K-nearest Neighbour is a non-parametric method used for classification and regression. In both cases, the input consists of the k closest training examples in the feature space. The output depends on whether K-NN is used for classification or regression.

K-NN is one of the simplest machine learning technique. When a new situation occurs, it scans through all past experiences and looks up the k closest experiences. Those experiences are what we call the K nearest neighbours. In this K-NN technique, it holds all the previous data sample space while predicting the target value of the new input sample data predictor.

1.4 Linear Regression

Regression is a method of modelling a target value based on independent predictors. This method is mostly used for forecasting and finding out cause and effect relationship between variables. Regression techniques mostly differ based on the number of independent variables and the type of relationship between the independent and dependent variables.
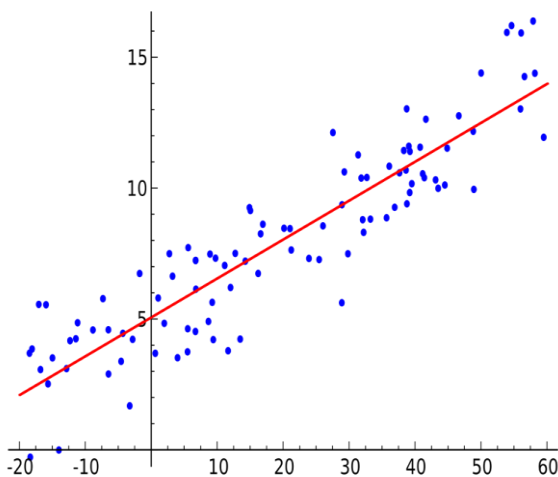

Fig. 1: Linear regression graph

Simple linear regression is a type of regression analysis where the number of independent variables is one and there is a linear relationship between the independent(x) and dependent(y) variable. The red line in the above graph is referred to as the best fit straight line. Based on the given data points, we try to plot a line that models the points the best. The line can be modelled based on the linear equation shown below.

$$y = a\_0 + a\_1 * x \quad \text{\#\# Linear Equation}$$

The motive of the linear regression algorithm is to find the best values for a_0 and a_1.

2. PROPOSED WORK

In India as per the survey was done in 2011, 61.5% of the population depends on agriculture (from 2011 to now population is rapidly increasing dependency on agriculture also increases.

Due to the informal practice of agriculture farmers are not getting proper yield. This leads to the shortage of food and it also affects economic growth. But nowadays, though the farmers are using advanced technologies they are getting maximum yield. But they are not getting demand. Farmers never had an idea about whether the cultivated crop will have more demand or not.
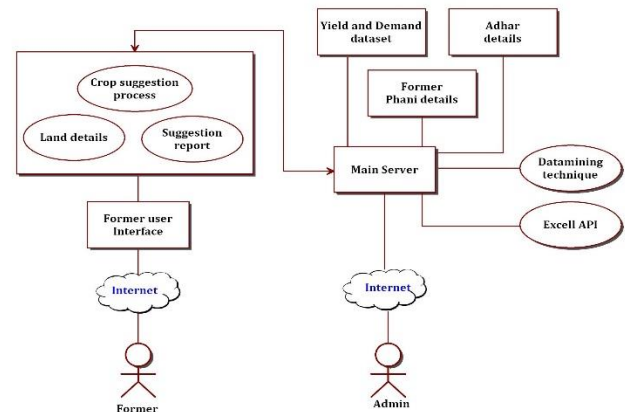

Fig. 2: System architecture

In the given above architecture, the system will carry out to perform the yield and demand prediction and give the suggestion to the farmers whether to cultivate the right crop or not. In our paper, we have made into two sections one is for the government purpose (admin) and another for the farmers (user) to know the crop has to grow in their land.

In the admin section, it controls all the details of the farmers including Aadhar number details and land details with phani. We will be having 10 years of a dataset of the crop which is distributed in a district wise crops and seasonable crops. By using data mining technique we will predicting the crop yield and for the particular yield whether the demand is there or not in the market price will be predicted. Whereas in the farmer's section that crop suggestion will be done based on the user selecting the crop and suggesting that crop can be cultivated.

3. MODULES
3.1 PHANI registration
In this module, admin can add the former land information details like Aadhaar number and survey number of the land and Number of Acers Land.

3.2 Upload the dataset
In this module, an admin will upload the previous 10-year dataset like
- vegetable demand,
- crops demand
- vegetable yield dataset monthly wise
- vegetable demand dataset monthly wise
- District wise crops data season wise(Kharif, Summer, Winter)

Admin will upload all the above datasets using the Excel API

3.3 New demand calculation process
In this module we predicting the Yield from Yield Dataset Using the Linear Regression Technique Let NYLD and No acres former going to cultivate NACR and calculate the new demand.

$$NewDemand1 = NYLD * NACR$$

3.4 Demand prediction process

In this module we predicting the demand From Demand Dataset Using the Linear Regression Technique Let NewDemand2.

3.5 Land suggestion process

We are comparing the NewDemand1 and NewDemand2 if NewDemand2 is greater than NewDemand1 admin will give permission to cultivate.

4. RESULTS

4.1 Input

- The area (land information) with size should be given as the input and the crop name.
- Information on the current market rate and demand.

4.2 Output

- It will process the datasets given and it will calculate
- and predicate the required output as:
- Whether the crop which is given as input is correct or no will you get proper yield?
- How much land should be kept for that particular crop?
- How much will be the profit for a given crop in the given land?
- Which crop will be the best crop for a particular season?

5. CONCLUSION

This paper presents a crop selection method (linear regression) technique to select the crop for maximizing the yield and even checking the crop of the demand in the market.so crop selection technique can increase the yield rate of the crop which is planted over seasons. The proposed method can resolve the problem in selecting the crop through influencing some parameters are soil classification, market price, weather conditions etc. So there is a need to adopt a prediction method with more accuracy and high performance. Through this approach farmer's death rate can be reduced.

6. FUTURE SCOPE

In Future work, as we know 50% of the population is dependent upon agriculture. Hence our project helps the farmers to grow right crops. Vegetables/grains wasting can be reduced by guiding the farmers in a correct way during cultivation and simultaneously economic growth of country increases

7. REFERENCES

[1] Drummond ST, Sudduth KA, Joshi A, Birrel SJ, Kitchen NR, "Statistical and neural methods for site-specific yield prediction". T ASABE 46 (1): 5-14, 2003.

[2] Safa B, Khalili A, Teshnehlab M, Liaghat A," Artificial neural networks application to predict wheat yield using climatic data. Proc". 20th Int. Conf. on IIPS, Jan. 10-15, Iranian Meteorological Organization, pp: 1-39, 2004.

[3] Irmak A, Jones JW, Batchelor WD, Irmak S, Boote KJ, Paz JO," Artificial neural network model as a data analysis tool in precision farming". T ASABE 49(6): 2027-2037, 2006.

[4] Miss Snehal, S. Dahikar, Dr. Sandeep V. Rode," Agricultural Crop Yield Prediction Using Artificial Neural Network Approach". International Journal of Innovative Research in Electrical, Electronic, Instrumentation and Control Engineering, Vol. 2, Issue 1, January 2014.

[5] M. Amini, N. Sadat Safavi, S. Sohaei, and S. M. Noorbakhsh, Agricultural Development In IRAN Base On Cloud Computing Theory, Int. J. Eng. Res. Technol. IJERT, vol. 2, no. 6, pp. 796–801, 2013.

[6] Liu J, Goering CE, Tian L," Neural network for setting target corn yields". T ASAE 44(3): 705-713, 2001.

[7] Sudduth K, Fraisse C, Drummond S, Kitchen N" Integrating spatial data collection, modeling and analysis for precision agriculture". First Int. Conf. on Geospatial Information in Agriculture and Forestry, vol. 2, pp: 166-173, 1998.

[8] K. Krishna Kumar, K. Rupa Kumar, R. G. Ashrit, N. R. Deshpande and J. W. Hansen, "Climate Impacts on Indian Agriculture". International Journal of Climatology, 24: 13751393, 2004.

[9] Hand D, Mannila H, Smyth P," Principles of data mining". MIT Press, 2001.

[10] Raorane A.A., Kulkarni R.V.,"Data Mining: An effective tool for yield estimation in the agricultural sector". International Journal of Emerging Trends and Technology in Computer Science (IJETTCS) Volume 1, Issue 2, July August 2012.

[11] P. Revathi, R. Revathi, Dr. M. Hemalatha,"Comparative Study of Knowl- an edge in Crop Diseases Using Machine Learning Techniques". Inter- national Journal of Computer Science and Information Technologies (IJCSIT), Vol. 2 (5), 2180-2182, 2011.

[12] Sayed M. Arafat, Mohamed A. Aboelghar, Eslam F. Ahmed," Crop Discrimination Using Field Hyper Spectral Remotely Sensed Data". Advances in Remote Sensing, scientific research, 2, 63-70, 2013.

[13] Yudong Zhang, Lenan Wu.,"Crop Classification by forwarding Neural Network with Adaptive Chaotic Particle Swarm Optimization". Sensors, 4721-4743; doi: 10.3390/s110504721, 2011.

[14] Raorane A.A., Kulkarni R. V.," Data Mining: An effective tool for yield estimation in the agricultural sector". International Journal of Emerging Trends and Technology in Computer Science (IJETTCS) Volume 1, Issue 2, July August 2012.

[15] A.Nithya, Dr.V.Sundaram,"Classification rules for Indian Rice dis- ease". International Journal of Computer Science Issues, Vol. 8, Issue 1, ISSN (Online): 1694-0814, January 2011.

[16] K. Krishna Kumar, K. Rupa Kumar, R. G. Ashrit, N. R. Deshpande and J. W. Hansen, "Climate Impacts on Indian Agriculture". International Journal of Climatology, 24: 13751393, 2004.

[17] L. Breiman,"Random forests". Machine Learning, vol. 45, no. 1, pp. 532, 2001.

[18] Uysal I, Altay HG,"An overview of regression techniques for knowledge discovery". Knowl Eng Rev 14: 319-340, 1999.

[19] Breiman L, Friedman JH, Olshen RA, Stone CJ," Classification and regression trees". Wadsworth, Belmont, CA, USA, 1984.

[20] Wang Y, Witten I,"Inducing model trees for continuous classes". Proc. 9th Eur. Conf. Machine Learning (van Someren M and Widmer G, Eds), pp: 128-137, 1997.

[21] Hand D, Mannila H, Smyth P," Principles of data mining". MIT Press, 2001.

[22] Washington Okori, Joseph Obua,"Machine Learning Classification Technique for Famine Prediction". Proceedings of the World Congress on Engineering 2011 Vol II WCE 2011, July 6 - 8, London, U.K, 2011.

[23] Vapnik V, Golowich S, Smola A," Support vector method for function approximation, regression estimation, and signal processing". MIT Press, Cambridge, MA, USA, pp: 281-287, 1997.

[24] Mohd. Noor Md. Sap, A. Majid Awan, "Development of an Intelligent Prediction Tool for Rice Yield based on Machine

Learning Techniques ". Jurnal Teknologi Maklumat, Jilid 18, Bil. 2, December 2006.

[25] Alberto Gonzalez-Sanchez, Juan Frausto-Solis, Waldo Ojeda- Bustamante," Predictive ability of machine learning methods for massive crop yield prediction". Spanish Journal of Agricultural Research, 12(2): 313328, 2014.

[26] R. Devadas, R. J. Denham, M. Pringle," Support Vector Machine Classification of Object-based Data for Crop Mapping Using Multi-Temporal Landsat Imagery".

International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXIX-B7, 2012 XXII ISPRS Congress, Melbourne, Australia, September 2012.

[27] Raorane A.A., Kulkarni R. V.," Data Mining: An effective tool for yield estimation in the agricultural sector". International Journal of Emerging Trends and Technology in Computer Science (IJETTCS) Volume 1, Issue 2, July August 2012.