2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, T.N., India. 6 - 8 May 2015. pp.138-145.

Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique

Rakesh Kumar[1], **M.P. Singh**[2], **Prabhat Kumar**[3] and **J.P. Singh**[4]

Dept. of CSE NIT Patna, India
Email: [1]rakeshxaxo@gmail.com, [2]mps@nitp.ac.in, [3]prabhat@nitp.ac.in, [4]jps@nitp.ac.in

***Abstract*— Agriculture planning plays a significant role in economic growth and food security of agro-based country. Selection of crop(s) is an important issue for agriculture planning. It depends on various parameters such as production rate, market price and government policies. Many researchers studied prediction of yield rate of crop, prediction of weather, soil classification and crop classification for agriculture planning using statistics methods or machine learning techniques. If there is more than one option to plant a crop at a time using limited land resource, then selection of crop is a puzzle. This paper proposed a method named Crop Selection Method (CSM) to solve crop selection problem, and maximize net yield rate of crop over season and subsequently achieves maximum economic growth of the country. The proposed method may improve net yield rate of crops.**

***Keywords*— Climate, RGF (Regularized Greedy Forest), Soil composition, CSM (Crop Selection Method), GBDT (Gradient Boosted Decision Tree), regularization, regression problem.**

I. Introduction

Achieving maximum yield rate of crop using limited land resource is a goal of agricultural planning in an agro-based country. Antecedent determination of problems associated with crop yield indicators can help to increase yield rate of crops. Crop selector could be applicable for minimize losses when unfavorable conditions may occur and this selector could be used to maximize crop yield rate when potential exists for favorable growing conditions. Maximizing production rate of crop is an interesting research field to agro-meteorologists which play a significant role in national economic. There are two types of factors which influence yield rate of crop: first is seeds quality which can be improved by genetic development using hybridization technology, and second is crop selection management based on favorable or unfavorable conditions.

Many research intended to agriculture planning is carried out, where the goal is to get an efficient and accurate model for crop yield prediction, crop classification, soil classification, weather prediction, crop disease prediction, classification of crops based on growing phase. A statistical and machine learning both techniques were modeled. This paper develops a new method called Crop Selection Method (CSM) to maximize net yield rate of crops over season.

Crop production rate depends on geography of a region (e.g. hill area, river ground, depth region), weather condition (e.g. temperature, cloud, rainfall, humidity), soil type (e.g. sandy, silty, clay, peaty, saline soil), soil composition (e.g. PH value, nitrogen, phosphate, potassium, organic carbon, calcium, magnesium, sulphur, manganese, copper, iron) and harvesting methods. Different subsets of these influencing parameters are used for different crops by different prediction models. Some of these prediction models for crop production rate are thoroughly studied thorough out the research. Pre- diction models are broadly classified into two types: first is a traditional statistics model (e.g. multiple linear regression model) which formulates a single predictive function holding entire sample space; i.e. it generates a global model over entire sample space. Second is machine learning technique which is emerging technology for knowledge mining that relates input and output variables which is hard to obtain statistically. In traditional statistics methods, structure of data model needs to be assumed priory, whereas machine learning techniques need not assume this structure. This is useful characteristics for machine learning techniques to model complex non-linear behavior in crop yield prediction. Machine learning methods which are widely used in prediction technique are boosting techniques (e.g. RGF, GBDT, and Add boost), regression tree (e.g. ID3, C4.5, and M5-prime regression tree), random forest, support vector machine, k-nearest neighbors and artificial neu- ral network. Among all these prediction techniques boosting techniques (e.g. GBDT, RGF) are still untouched in crop yield prediction. In this research, some machine learning techniques are studied and comparative analysis is presented.

A. Artificial Neural Network(ANN)

An ANN is an interconnection of weighted processing unit.

A processing unit takes input from previous processing unit or from outer unit and transfer output to other processing unit. An ANN is a topological machine learning technique. Most widely used topological algorithm is multi-layered perceptron and back-propagation algorithm [1] to implement neural networks for crop yield prediction [2][3][4].Unfortunately, there is no any automatic technique to determine a suitable topology for data sample space. Therefore, topology is empirically selected for suitable crop yield prediction [5] [6]. An artificial neural network is used when number of input attributes is lesser.

B. Support Vector Machine (SVM)

Support vector machine used for crop yield prediction is called support vector regression. The goal of the support vector technique is to obtain non-linear function using kernel function (a linear function or polynomial function) [7] [8] [9]. The radial basis function and the polynomial function are widely used kernel function.The advantage of support vector regression is to avoid difficulties of using linear function in large input samples space and optimization of a complex problems transformed into simple linear function optimization.

C. K-Nearest Neighbors (K-NN)

k-NN is called sample based learning technique, where it holds all past data sample space while predicting target value for new input sample predictor. It applies distance function (eg. Euclidean, Manhattan, Minkowski distance function) to com- pute distance from new input sample predictor to all training sample predictors and then k nearest (or smallest) distances are selected with corresponding target values. Target value for new sample predictor is weighted sum of the target values of all the k neighbors. Selection of k is a puzzle of the method, value of k is directly proportional to the prediction instability (i.e. more data sensitive). Smaller value of k means high variance and low bias and higher value of k means vice versa. The advantages of k-Nearest Neighbors are it does not require training and optimization [10]. It works on locality concept and it is used for nonlinear and highly adaptable problem. As k-NN uses all data sample during prediction of new data case, its time and space complexity both are comparably high. It is a laziest technique among all the machine learning techniques. Also, increase the dimensionality of input vector making more puzzling. The use of k-NN in agriculture is studied in [11].

D. Decision Tree Learning

Decision tree learning splits entire sample space recursively into smaller sub-sample space which is enough to be formulated by a simple model [12]. The root node (first node) in the tree holds entire sample space. Splitting sample space into smaller sub-sample space means forking root node into children nodes where each child node may be recursively split into leaf nodes (a node on which further split is not possible). The Nodes except leaf node in the tree, split sample space based on a set of condition(s) of the input attributes values and the leaf node assign an output value for those input attributes which are on the path from root to the leaf in the tree. The ultimate goal of sub-sample using decision tree method is to mitigate mixing of different outputs values and assign single output value for sub-samples space. The splitting criteria of a node are an impurity measure (e.g. standard deviation used in ID3 algorithm; Gini-Index used in C4.5 algorithm) and Node size (number of data present on a nodde). There are many algorithms to build decision tree are: CART [13], M5 [12], and M5-Prime [14]. All these algorithms are similar in tree generation procedure, but they differ in following aspects: first the impurity measure such as M5 uses standard deviation and CART uses variance. Second is prune rule used to avoid over-fitting of a model. Third is the leaf value assignment. M5 apply linear model at leaf nodes instead of constant value [12]. Furthermore, M5 is simple, smooth and more accurate than CART algorithm [15]. M5-Prime is subsequent version of M5 dealing with missing values and enumerated attributes [14].

E. Random Forest

Random Forest is a bagging technique which is based on tree ensemble machine learning method[16]. It generates multiple tree of randomly sub-sampled features. The output of forest is evaluated by taking average value of the prediction of individual trees. Since it is using random sub-sampled features, Random Forest can be used in high dimension input predictor.

F. Gradient Boosted Decision Tree (GBDT)

Gradient Boosted Decision Tree is an additive decision tree algorithm in which a series of boosted decision trees are created and additively form a forest as a single predictive model [17]. It uses decision tree as a weak learner to build an additive predictive model on re-weighted data [17]. GBDT is also called as wrapper approach in which a decision tree treated as a base learner or weak learner. An additive wrapping is done on base learner. The advantages of GBDT are - first, base learner can be changed to other learner with same wrapper; second, initialization of predictor variables is not needed unlike Ada-boost. The disadvantage of GBDT is that the boosting wrapper teats decision tree as a black box and it works on tree optimization rather than forest optimization.

G. Regularized Greedy Forest (RGF)

Regularized Greedy Forest is an additive decision tree algorithm in which a series of boosted decision trees are created and additively form a forest as a single predictive model. A globally optimized decision tree is created in RGF whereas locally optimized decision tree is created in GBDT [18]. RGF takes advantage of tree structure as it works on fully corrective regularized steps where as GBDT does not take advantage of tree structure as it works on partially regularized steps [17] [18]. RGF works faster and more accurate than GBDT for regression problem [18].

II. RELATED WORK

Forecasting agriculture product plays a significant role in agriculture planning. It helps in making product storage, business strategy and risk management [19]. There are two methods to forecast agriculture product in advance. First is statistics method such as Autoregressive Integrate Moving Average (ARIMA) and Holt-Winter and second is machine learning method such as Support vector machine and artificial neural network. These methods are comparatively study over Thailand's pacific white shrimp export data and Thailands pro- duced chicken data using support vector machine and ARIMA model [19]. Where support vector method gives more accurate result than ARIMA. Moreover, machine learning methods are convenient to implement and comparably faster than statics methods.

Indian agriculture is highly dependent on summer rainfall [20]. The correlation between summer rainfall and agriculture product production is studied in [20]. This paper presents an analysis of crop-climate relationship using past crops data. Correlation analysis tells that the monsoon rainfall, Pacific and Indian Ocean sea-surface temperatures and Darwin sea- level pressure directly influence the crop production in India. Result shows that the state-level crop production statistics and sub divisional monsoon rainfall are consistent with the all-India result, except few cases. Moreover, the impact of sub divisional monsoon rainfall related to El Niosouthern oscillation and the Indian Ocean sea-surface temperatures have seen long time a greatest impact in the western to central peninsula.

A famine prediction application is modeled using machine learning technique [21]. Predicting the famine for a region early is used to mitigate the vulnerability of the society at risk. Machine learning techniques are experimented on past data collected between 2004 and 2005 in Uganda. The performance of machine learning methods named Support Vector Machine (SVM), Naive Bayes, k-Nearest Neighbors (k-NN) and Decision tree classifier in prediction of famine were assessed empirically [21]. SVM and k-NN methods give better result than the rest of the methods, moreover the region of convergence produced by Support Vector Machine can be used by strategic planner in cut-off determination of famine prone management.

An UChooBoost machine learning method is modeled for precision agriculture [22]. The emerging technology in agricul- ture field needs to process large amount of digital information related to agriculture field. The UChooBoost is a supervised learning ensemble-based algorithm used for knowledge mining in agriculture data. UChoo classifier is used as base classifier in bootstrap ensemble. A combination of weighted major- ity voting is used for performance evaluation in precision agriculture [22]. UChooBoost is empirically evaluated for an extended data and it shows good performance in experiment with agriculture data. The strongest trait of using UChooBoost is to apply for an extended data expression and works on compounding hypotheses which leads to improve algorithm performance.

Artificial neural network is used as crop yield prediction by sensing various parameters of climate and soil [23]. Param- eters are water depth, soil type, temperature, presser, rainfall, humidity, nitrogen, phosphate, potassium and organic carbon. The impact of these parameters are studied and empirically assessed in paper [23]. It is observed that the production rate of crop is correlated with atmospheric parameter, soil type and soil composition. This paper also suggests suitable crop based on prediction of crop yield rate in advance. Artificial neural network is used as powerful tool for modeling and prediction of crop yield rate and improve the effectiveness of crop yield prediction.

Agriculture product depends on climatic, geographical, biological, political and economic factors [24]. Since these factors are highly sensitive, there are some risks which can be measured appropriately. These risks can be quantified mathematically or using learning technique. The accurate information about factors influencing crop yield is important for both farmer and government of the country. Prediction of crop yield based on historical data plays a significant role to mitigate vulnerable risk. The main challenges in agriculture data are to process these huge raw data effectively and accu- rately. Artificial neural network is a learning technique used to mine knowledge of meaningful information from raw data effectively and efficiently. The paper [24] aimed to assess a data mining technique and apply them to big raw data-sets to correlate crop yield rate and influencing factors as mentioned earlier.

An intelligent tool for rice yield prediction is developed using statistics and machine learning techniques. This tool is used in classification and clustering [25]. Support vector machine learning technique is used for classification or rice plantation data. Kernel based clustering algorithm is used for finding cluster in climate data. Kernel methods are applicable for complex, high dimensional and non-linearly separable data. Correlation analysis is performed for evaluating the impact of various influencing parameters on the rice yield and regression analysis is performed for predicting the crop yield rate. Support vector machine is used for noisy data. These features makes tool as an intelligent system for predicting rice yield.

Machine learning techniques are widely used in crop yield prediction. There are many learning techniques proposed for crop yield prediction, and comparatively studied by many researchers seeking for the most accurate technique. But due to the less number of evaluated crops and techniques, an appropriate decision cannot be achieved [26]. A comparative analysis is performed for large number of evaluated crops and technique in the paper [26]. The result shows accuracy percentage of different learning technique on the collected data set and the paper [26] suggest some learning technique for crop yield prediction for different crop data-sets.

The production rate of crop in China is studied by splitting whole region of China into six different regions [27]. Using combination of historical crop yield record, meteorological observations, and 28 CMIP5 (Coupled Model Inter-comparison Project Phase 5) ensemble methods, to evaluate impact of future climate change on crop yields. CMIP5 is a statistics method to build a prediction model. It is seemed that the crop yields in Northwest and Southwest China are positively cor- related with temperature change and little crop (e.g. soybean) production in Northwest depends on precipitation; where as, in East and Central-South China, these crops are positively correlated with both precipitation and temperature change. However, there is no any significant correlation between crop yield and climate parameter in North and Northeast China except for few crops such as wheat as well as rice production in North China is weakly correlated with temperature and soybean production with temperature in Northeast China. It is observed empirically that the spatial pattern among the four crops (e.g. wheat, rice, soybean, maize), the sensitivity to temperature changes increasing from North to South China.

Tow-hidden-layer forward neural network is modeled to de- velop hybrid crop classifier for polar metric synthetic aperture radar (SAR) images [28]. A hybrid feature set is constructed using span images, the H/A/alpha decomposition and gray- level co-occurrence matrix-based texture features. Principle component analysis (PCA) is used to reduce the feature and then a two-layer forward neural network is trained by adaptive chaotic particle swarm optimization. It is empirically observed on flevoland that, in this case, adaptive chaotic particle swarm optimization works better than back propagation (BP), adaptive BP, particle swarm optimization and momentum BP methods. Classifier is regularized using k-fold cross validation.

Hyper spectral wavebands in the range of 400 to 2500 nm is used in satellite crop classification technique to differentiate images of two crops of the same season (e.g. winter crops: clover, wheat and summer crops: rice, maize) in the cultivated land of the Egypt [29]. A spectral reflectance is split into six spectral zones named green, blue, red, near-infrared, shortwave infrared-I, shortwave infrared-II. These spectral zones are mon- itored by hyper spectral ground measurements of ASD field spectrum-III spectra radiometer. Each zone consists of multi- spectral band and each band identifies a crop. The optimal spectral zone is selected to discriminate two crops by ANOVA and Tukeys HSD post hoc analysis. Waveband in the spectral zone is identified by linear regression discrimination (LDA). This paper shows that near-infrared, shortwave infrared-I and shortwave infrared-II spectral zones are more sufficient than red and green spectral zones in the differentiation of two different corps of the same season.

A plant nutrient management system is developed to corre- late needed nutrient of plant with avail fertility of soil [30]. Su- pervised machine learning technique named back propagation neural network (BPN) is used to correlate soil properties such as organic matter, micro-nutrient and essential plant nutrients that affects growth of crops. The whole process is split into three steps such as soil sampling, training back propagation neural network and weight pupation of hidden layer of neuron. BPN technique performs better than multivariate regression technique.

Remote sensing technique is developed for Land Cover Classification (LCC) to monitor land used in crop mapping over large extent geographic area of Queensland, Australia [31]. A new approach named image object-based classification is built instead of tradition pixel-based classification method. Data are collected from Darling Downs region of southern Queensland as a land sat TM image form from 2010 to 2011 cropping season. The developed model is trained with these data using Support Vector Machine (SVM) classification. A polygon object is obtained from collected data images. Image object-based method provides capability to analyze aggregated sets of pixels, shape-related images, textual variation and spectral characteristics. Then support vector machine is trained using 3 shaped-based parameters, 23 textual parameters and 10 spectral parameters of the image objects. It is observed empirically that object-based parameter performs better than traditional pixel-based parameters.

C4.5 decision tree algorithm is used to build Rice Disease Classification (RDC) based on symptoms [32]. The experiment is done over Indian Rice Disease. Decision tree, C4.5, is used to automatically acquire knowledge from empirical data of Indian Rice Disease. The advantage of C4.5 is interpretable. The algorithm, C4.5, can effectively built a tree with high predictive power and gives more accurate result on test data set.

Machine learning techniques are used in prediction of crop diseases classification [33]. Couple of machine learning techniques are studied such as C4.5 decision tree algorithm, support vector algorithm and artificial neural network to de- velop agriculture applications. Paper [33] presents a compar- ative study of crop diseases using above mentioned machine learning techniques.

Machine learning is an emerging technology in knowledge mining. The application of machine learning technique reduces complexity of pattern recognition and gives accurate result. A classification rule extraction tool for rice disease of Egyptian nation is developed using C4.5 decision tree learning algorithm [34]. The larger number of prediction tools have been devel- oped for agriculture application worldwide. The advantage of C4.5 is interpretable i.e. C4.5 is used to create comprehensive tree with greater predictive power. Regularization in it can be achieved by pruning method that improves the capability of predictive power.

III. Proposed Work

An agro-based country depends on agriculture for their economic growth. When population of country increases de- pendency on agriculture also increases and subsequent eco- nomic growth of the country is affected. In this situation, crop yield rate plays a significant role in economic growth of the country. So, there is a need to increase crop yield rate. Some biological approaches (eg. seed quality of crop, crop hybridization) and some chemical approaches (eg. use of fertilizer, urea, potash) are carried out to solve this issue. In addition to these approaches a crop sequencing technique is required to improve net yield rate of crop over season.

This paper proposed a method named Crop Selection Method (CSM) to achieve net yield rate of crops over season. Crop can be classified as :

a) *Seasonal crops*— crops can be planted during a season. eg. wheat, cotton.

b) *Whole year crops*— crops can be planted during entire year. eg. vegetable, paddy, Toor.

c) *Short time plantation crops*— crops that take short time for growing. eg. potato, vegetables, ratoi.

d) *Long time plantation crops*— These crops take long time for growing. eg. sugarcane, canda.

A combination of these crops can be selected in a sequence based on yield rate per day. Fig.1 illustrates sequences of crops with cumulative yield rate over season. CSM method, shown in Fig.2, may improve net yield rate of crops using limited land resource and also increases re-usability of the land.

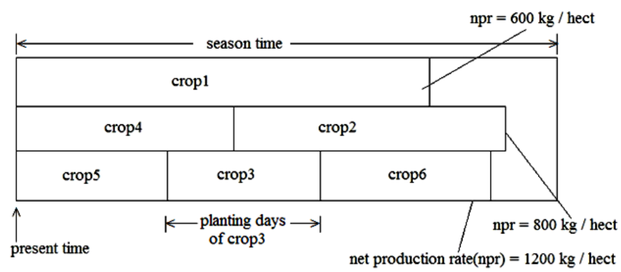

Fig. 1. Selected Crop Sequence: Crop5, Crop3, Crop6

$$\text{cropSelector(presentTime)} = \begin{cases} 0 \quad ; \text{if presentTime} \geq \text{End of season} \\ \text{cropSelector(presentTime+1) ; if presentTime} \neq \text{sowingTime} \\ \text{crop} \leftarrow \max\limits_{\text{crop} \in \text{cropSowingTable}} \{\text{crop} \rightarrow \text{productionRate/crop} \rightarrow \text{plantingDay}\} \\ \quad + \text{cropSelector(presentTime+ crop} \rightarrow \text{plantingDay);} \\ \qquad \text{if presentTime = sowingTime} \end{cases}$$

Fig. 2. CSM Method

A. Crop Data and CSM Algorithm

Season of entire year in India depends on summer rainfall depth of every year [20]. So, it is possible to predict season in advance for entire year. Based on these predicted season and past data of crop yield rate, it is possible to predict yield rate of crop earlier. CSM algorithm works on prediction of crop yield rate based on favorable condition in advance and gives a sequence of crops with highest net yield rate. For example consider crop sowing table, data are gathered from farmer of Patna district, Bihar (India).

Algorithm: Crop Selection Method(CSM)

```
cropSelector(prese
ntTime)
   if presentTime ≥ End of
      Season then return 0
   end if else
      if presentTime= sowingTime then
         return cropSelector(presentTime + 1)
      end if else
         cropSowingTable←cropInputTable(presentT
         ime) L: crop ← max{crop →prRate /
                        crop → plantationDay}
                  crop ∈ cropSowingTable
```

Table 1. Crop Sowing Table

Crop name	*Sowing period*	*Harvesting period*	*Growing day or Plantation day*	*Predicted yield rate(kg/hectare)*
Rice	June -July	Sep - Oct	3.5 month	2000
Soybean(less water)	June - July	Nov - Dec	4 month	1264
Sweet potato(less water)	July - August	Oct - Nov	3.5 month	800
Arhar(Toor)(less water)	July - August	Dec - Jan	5.5 month	1359
Castor seed(less water)	July - August	Feb - March	5 month	1064
Wheat	Oct- Nov	Feb - March	3.5 month	788
Potato	Oct - Dec	Dec - Jan	2.5 month	1650
Ratoi(less water)	Oct- Nov	Dec - Jan	1.5 month	1472
Toria(average water)	Oct - Nov	Jan - Feb	3 month	1800
Sarso(average water)	Oct - Nov	Jan - Feb	3 month	1800
Lineseed(Tisi)(no water)	Oct - Nov	Feb - March	4 month	1182
Masoor	Oct - Nov	Feb - March	4 month	1259
Khesari	Oct - Nov	Feb - March	4 month	1511
Onion	Jan - Feb	April - May	3 month	115
Sugarcane	Feb - March	Jan - Feb	11 month	270
Kanda	March - April	Sep - Oct	6 month	995
Mung	March - April	June - July	3.5 month	1492
Til	March - April	June - July	3.5 month	1287
Ladies finger	All season	All season	2.5 month	1136
Pumpkin	All season	All season	3 month	756
Nenua(Pror)	All season	All season	3 month	865
Coriander	All season	All season	2.5 month	1000

```
      if (presentTime + crop → platationDay) ≥
            End of Season then
         //remove crop from cropSowingTable
         cropSowingTable ← cropSowingTable -
         crop if cropSowingTable is NULL then
            return cropSowingTable(presentTime +
            1)
         end if else
            go to L
         end else end if
      else
         update(OutputcropTable, crop)
         npr ← (crop → prRate +
         cropSelector(presentTime + crop
         →plantationDay))
         return npr
      end else end else
  end else
end cropSelector
```

CSM method retrieves all possible crops that are to be sown at a given time stamp. Yield rate of these crops are evaluated, if yield rate per day of these crops are fair (within tolerance) then those crops are selected for crop sequences. Further, time after harvesting time of considered crop is taken as a given time stamp for further selection of crop. Therefore, multiple sequences of crops with different yield rate are obtained. The sequence of crops with highest net yield rate is finally considered for crop sequencing. For better understanding consider following examples illustrated in Fig.3. Thus, final obtained sequence consist Rice, Wheat and Mung for given year.

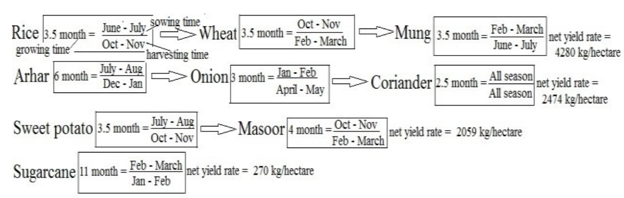

Fig. 3. Selected Sequence of Crops: Rice, Wheat, Mung

IV. CONCLUSION

This paper presents a technique named CSM to select sequence of crops to be planted over season. CSM method may improve net yield rate of crops to be planted over season. The proposed method resolves selection of crop (s) based on prediction yield rate influenced by parameters (e.g. weather, soil type, water density, crop type). It takes crop, their sowing time, plantation days and predicted yield rate for the season as input and finds a sequence of crops whose production per day are maximum over season. Performance and accuracy of CSM method depends on predicted value of influenced parameters, so there is a need to adopt a prediction method with more accuracy and high performance.

References

[1] Rumelhart DE, Hinton GE, Williams RJ, ”Learning internal repre- sentations by error propagation”. vol. 1, chapter 8. The MIT Press, Cambridge, MA (USA), pp:418-362, 1986.

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

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] Vapnik V, Lerner A,” Pattern recognition using generalized portrait method”. Automat Remote Contr 24: 774-780, 1963.

[8] Smola A, Schlkopf B,” A tutorial on support vector regression”. Stat Comput 14(3): 199-222, 2004.

[9] 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.

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

[11] Zhang L, Zhang J, Kyei-Boahen S, Zhang M,”Simulation and prediction of soybean growth and development under field conditions”. Am-Euras J Agr Environ Sci 7(4): 374-385, 2010.

[12] Quinlan JR, ”Learning with continuous classes”. Proc. AI92, 5th Aust. Joint Conf. on Artificial Intelligence (Adams & Sterling, eds.), World Scientific, Singapore, pp: 343-348, 1992.

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

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

[15] Uysal I, Altay HG, ”An overview of regression techniques for knowl- edge discovery”. Knowl Eng Rev 14: 319-340, 1999.

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

[17] J. Friedman, ”Greedy function approximation: A gradient boosting machine”. Ann. Statist., vol. 29, no. 5, pp. 11891536, 2001.

[18] Rie Johnson, Tong Zhang, ”Learning Nonlinear Functions Using Reg- ularized Greedy Forest”. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 36, No. 5, May 2014.

[19] Thoranin Sujjaviriyasup, Komkrit Pitiruek, ”Agricultural Product Fore- casting Using Machine Learning Approach”. Int. Journal of Math. Analysis, Vol. 7, no. 38, 1869 1875, 2013.

[20] 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.

[21] 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.

[22] Anastasiya Kolesnikova, Chi-Hwa Song, Won Don Lee, ”Applying UChooBoost algorithm in Precision Agriculture”. ACM International Conference on Advances in Computing, Communication and Control, Mumbai, India, January 2009.

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

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

[25] 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, Disember 2006.

[26] 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): 313-328, 2014.

[27] Zhihong Zhuo, Chaochao Gao, Yonghong Liu., "Regional Grain Yield Response to Climate Change in China: A Statistic Modeling Approach". IEEE Journal of Selected topic in Applied Earth Obeservation and Remote Sensing, VOL. 7, NO. 11, November 2014.

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

[29] 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.

[30] Shivnath Ghosh, Santanu Koley, "Machine Learning for Soil Fertil- ity and Plant Nutrient Management using Back Propagation Neural Networks". International Journal on Recent and Innovation Trends in Computing and Communication Volume: 2 Issue: 2, 292 297 ISSN: 2321-8169.

[31] 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 Photogramme-try, Remote Sensing and Spatial Information Sciences, Volume XXXIX- B7, 2012 XXII ISPRS Congress, Melbourne, Australia, September 2012.

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

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

[34] Mohammed El- Telbany, Mahmoud Warda, Mahmoud El-Borahy, "Min- ing the Classification Rules for Egyptian Rice Diseases". The Interna- tional Arab Journal of Information Technology, Vol. 3, No. 4, October 2006.