ISSN 2348 – 9928
IJAICT Volume 4, Issue 11, March 2018 Doi:01.0401/ijaict.2018.03.12 Published on 05 (04) 2018

EFFICIENT CROP YIELD AND CHEMICAL PREDICTION FOR RISING AGRICULTURAL ECONOMY VICTIMISATION DATA PROCESSING TECHNIQUES

Ms. T.R. Lekhaa
Assistant Professor,
Department of IT, SNS College Of Engineering,
Coimbatore, Tamilnadu, India

Abstract – Agriculture coming up with plays a big role in economic process and food security of agro-based country. Data technology has helped agricultural sector in managing and maintaining the mandatory details. The crop(s) plays a very important role in agriculture coming up with. Several researchers studied prediction of yield rate of crop, prediction of weather, soil classification and crop classification for agriculture turning out with. The planned technique resolves in choosing the crop(s) supported prediction yield rate influenced by parameters (e.g. weather, soil type, water density, crop type, pesticides) and additionally focuses on crop yield prediction, soil classification, chemical prediction and on-line mercantilism supported agriculture commodities.

Keywords – Soil Composition, Climate, Plant Composition, Ph Content, Crop Prediction Algorithm, Pesticide Prediction.

I INTRODUCTION

Agriculture plays a major role in economy in most of the countries. the most important goal of agriculture designing is to attain most yield rate with restricted land resource in agro-based trade like Bharat. Majority of analysis works in agriculture concentrate on biological mechanisms to spot crop growth and improve its yield. the most important downside in agriculture space is that the information acquisition and economical information exploitation. It had been understood that crop yield usually thrives solely in specific region or country, whereas few crops fail at yield in few regions. Antecedent determination of issues related to crop yield indicators will facilitate to extend yield rate of crops.
There ar 2 forms of factors that influence yield rate of crop: 1st is seeds quality which may be improved by genetic development victimization hybridizing technology, and second is crop choice management supported favorable or unfavorable conditions. the end result of crop yield primarily depends on parameters like style of crop, seed kind and environmental parameters like daylight (Temperature), soil (ph), water (ph), precipitation and wetness.

Many analysis supposed to agriculture coming up with is administrated, wherever the goal is to urge AN economical and correct model for crop yield prediction, crop classification, soil classification, weather prediction, crop sickness prediction, classification of crops supported growing part. Crop growth and yield square measure functions of an outsized variety of metabolic processes, that square measure full of environmental and genetic factors. A applied math and machine learning each techniques were sculptured. This paper develops a replacement technique known as Crop Prediction algorithmic program to maximise web yield rate for AN agro-based trade.

1.1 Artificial Neural Network (ANN)

An ANN is Associate in Nursing interconnection of weighted process unit. A process unit takes input from previous process unit or from outer unit and transfer output to alternative process unit. Associate in Nursing ANN may be a topological machine learning technique. most generally used topological algorithmic rule is multi-layered perception and back-propagation algorithmic rule to implement neural networks for crop yield prediction. sadly, there's no any automatic technique to work out an acceptable topology for knowledge sample house. Therefore, topology is by trial and error selected for appropriate crop yield prediction. a synthetic neural network is employed once range of input attributes is lesser.

1.2 Support Vector Machine (SVM)

Support vector machine used for crop yield prediction is termed support vector regression. The goal of the support vector technique is to get non-linear operate victimisation kernel operate (a linear operate or polynomial function). The radial basis operate and also the polynomial operate area unit wide used kernel operate.The advantage of support vector regression is to avoid difficulties of victimisation linear operate in massive input samples area and optimisation of a posh issues reworked into easy linear operate optimisation.

1.3 K-Nearest Neighbors (K-NN)

k-NN is termed sample primarily based learning technique, wherever it holds all past knowledge sample house whereas predicting target price for brand new input sample predictor. It applies distance perform (eg. Euclidean, Manhattan, mathematician distance function) to reason distance from new input sample predictor to any or all coaching sample predictors so k nearest (or smallest) distances square measure selected with corresponding target values. Target price for

brand new sample predictor is weighted add of the target values of all the k neighbors. choice of k may be a puzzle of the strategy, price of k is directly proportional to the prediction instability (i.e. additional knowledge sensitive). Smaller price of k means that high variance and low bias and better price of k means that the other way around. the benefits of k-Nearest Neighbors square measure it doesn't need coaching and improvement. It works on neighbourhood thought and it's used for nonlinear and extremely elastic downside. As k-NN uses all knowledge sample throughout prediction of recent knowledge case, its time and house complexness each square measure comparably high. it's a laziest technique among all the machine learning techniques. Also, increase the spatial property of input vector creating additional puzzling. the employment of k-NN in agriculture is studied in.

1.4 Decision Tree Learning

Decision tree learning splits entire sample area recursively into smaller sub-sample area that is enough to be developed by a straightforward model. the foundation node (first node) within the tree holds entire sample area. ripping sample area into smaller sub-sample area means that forking root node into kids nodes wherever every kid node is also recursively split into leaf nodes (a node on that more split isn't possible). The Nodes except leaf node within the tree, split sample area supported a collection of condition(s) of the input attributes worths and also the leaf node assign AN output value for those input attributes that area unit on the trail from root to the leaf within the tree. the final word goal of sub-sample exploitation call tree technique is to mitigate commixture of various outputs worths and assign single output value for sub-samples area. The ripping criteria of a node area unit} AN impurity measure (e.g. variance employed in ID3 algorithm; Gini Index employed in C4.5 algorithm) and Node size (number of information gift on a node). There area unit several algorithms to create call tree are: CART, M5, and M5-Prime. of these algorithms area unit similar in tree generation procedure, however they dissent in following aspects: 1st the impurity live like M5 uses variance and CART uses variance. Second is prune rule wont to avoid over-fitting of a model. Third is that the leaf worth assignment. M5 apply linear model at leaf nodes rather than constant worth. moreover, M5 is easy, sleek and a lot of correct than CART algorithmic program. M5-Prime is sequent version of M5 handling missing values and enumerated attributes.

1.5 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.

1.6 Gradient Boosted Decision Tree (GBDT)

Gradient Boosted call Tree is Associate in Nursing additive call tree algorithmic rule within which a series of boosted call trees square measure created and additively kind a forest as one prophetical model. It uses call tree as a weak learner to create Associate in Nursing additive prophetical model on re-weighted information. GBDT is additionally known as as wrapper approach within which a call tree treated as a base learner or weak learner. Associate in Nursing additive wrapping is finished aboard learner. the benefits of GBDT square measure - initial, base learner are often modified to alternative learner with same wrapper; second, format of predictor variables isn't required not like Adaboost. The disadvantage of GBDT is that the boosting wrapper teats call tree as a recording equipment and it works on tree improvement instead of forest improvement.

1.7 Regularized Greedy Forest (RGF)

Regularized Greedy Forest is AN additive call tree rule during which a series of boosted call trees square measure created and additively kind a forest as one prophetical model. A globally optimized call tree is formed in RGF whereas regionally optimized call tree is formed in GBDT. RGF takes advantage of tree structure because it works on totally corrective regularized steps wherever as GBDT doesn't cash in of tree structure because it works on partly regularized steps. RGF works quicker and a lot of correct than GBDT for regression drawback.

II RELATED WORK

Forecasting agriculture product plays a big role in agriculture designing. It helps in creating product storage, business strategy and risk management. There are a unit 2 ways to forecast agriculture product earlier. 1st is statistics technique like Autoregressive Integrate Moving Average (ARIMA) and Holt-Winter and second is machine learning technique like Support vector machine and artificial neural network. These ways area unit relatively study over Thailand's pacific white shrimp export knowledge and Thailands produced chicken knowledge victimization support vector machine and ARIMA model. Wherever support vector technique offers additional correct result than ARIMA. Moreover, machine learning ways area unit convenient to implement and comparably quicker than statics ways.

Indian agriculture is very obsessed with summer precipitation. The correlation between summer precipitation and agriculture product production is studied in. This paper presents AN analysis of crop-climate relationship victimization past crops knowledge. Correlation analysis tells that the monsoon precipitation, Pacific and ocean sea-surface temperatures and Darwin sea- level pressure directly influence the crop production in Bharat. Result shows that the state-level crop production statistics and sub divisional monsoon precipitation

area unit in keeping with the all-India result, except few cases. Moreover, the impact of sub divisional monsoon precipitation associated with El Niosouthern oscillation and therefore the ocean sea-surface temperatures have seen long-standing a greatest impact within the western to central earth.

A famine prediction application is sculpturesque victimization machine learning technique. Predicting the famine for a part early is employed to mitigate the vulnerability of the society in danger. Machine learning techniques area unit experimented on past knowledge collected between 2004 and 2005 in Uganda. The performance of machine learning ways named Support Vector Machine (SVM), Naive Bayes, k-Nearest Neighbors (k-NN) and call tree classifier in prediction of famine were assessed by trial and error. SVM and k-NN ways offer higher result than the remainder of the ways, furthermore the region of convergence made by Support Vector Machine are often utilized by strategic planner in cut-off determination of famine prone management.

An UChooBoost machine learning technique is sculpturesque for exactness agriculture. The rising technology in agriculture field must method great amount of digital data associated with agriculture field. The UChooBoost could be a supervised learning ensemble-based algorithmic program used for data mining in agriculture knowledge. UChoo classifier is employed as base classifier in bootstrap ensemble. a mixture of weighted major- ity selection is employed for performance analysis in exactness agriculture. UChooBoost is empirically observation ANd error evaluated for an extended knowledge and it shows sensible performance in experiment with agriculture knowledge. The strongest attribute of victimisation UChooBoost is to use for AN extended knowledge expression and works on combining hypotheses that ends up in improve algorithmic program performance.

Artificial neural network is employed as crop yield prediction by sensing varied parameters of climate and soil. Parameters area unit water depth, soil type, temperature, presser, rainfall, humidity, nitrogen, phosphate, metal and organic carbon. The impact of those parameters area unit studied and by trial and error assessed in paper. it's ascertained that the assembly rate of crop is correlate with region parameter, soil kind and soil composition. This paper conjointly suggests appropriate crop supported prediction of crop yield rate earlier. Artificial neural network is employed as powerful tool for modeling and prediction of crop yield rate and improve the effectiveness of crop yield prediction.

III PROPOSED WORK

An agro-based country depends on agriculture for his or her economic process. once population of country will increase de- pendency on agriculture conjointly will increase and consequent economic process of the country is affected. during this scenario, crop yield rate plays a big role in economic process of the country. So, there's a requirement to extend crop yield rate. Some biological approaches (eg. seed quality of crop, crop hybridization) and a few chemical approaches (eg. use of chemical, urea, potash) square measure distributed to resolve this issue.

In this paper, we've planned a technique referred to as crop yield prediction formula to attain internet yield rate of crops. a mixture of prediction and classification formula is employed for providing the higher crop. The crop details and soil details square measure got from the farmer. the main points like pH content, atomic number 7 content, Sulphur content etc square measure got from the farmer concerning the soil and square measure entered. The classification formula classifies the main points of soil in accordance to the crop details that square measure standardized for each crop. The classification formula can cluster similar information along and so the prediction formula can predict the crop details and best suited crop details for it.

In chemical prediction the chemical composting details square measure obtained at the side of the soil details. The classification formula can classify and cluster the similar details along. The prediction formula can predict the simplest suit of chemical for the given soil and crop.

Algorithm

cropSelector(presentTime)
if presentTime ≥ finish of Season
then
Match the hydrogen ion concentration content, chemical element content, sulphur content of soil with the crop
return 0
end if else
if presentTime= sowingTime then
return cropSelector(presentTime + 1)
end if else
cropSowingTable←cropInputTable(presentTime)
L: crop ←max
if (presentTime + crop → platationDay) ≥ finish 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 finish if
else
update(OutputcropTable, crop)
npr ← (crop → prRate +
cropSelector(presentTime+crop→plantationDay))
return npr
end else finish else

end else
end cropSelector

IV CONCLUSION

In this paper we've planned a mix of prediction and classification formula for providing higher yield in agriculture farm. Here the small print of the soil pH content, chemical element content, Sulphur content etc square measure obtained and classified and foretold with plant details. The chemical composting details square measure obtained and once more classified and foretold with soil and crop details. Performance and accuracy of CSM methodology depends on foretold price of influenced parameters, thus there's a desire to adopt a prediction methodology with a lot of accuracy and high performance.

References

[1]. Rumelhart DE, Hinton GE, Williams RJ, ”Learning internal representations 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]. R. Arulmurugan and H. Anandakumar, “Early Detection of Lung Cancer Using Wavelet Feature Descriptor and Feed Forward Back Propagation Neural Networks Classifier,” Lecture Notes in Computational Vision and Biomechanics, pp. 103–110, 2018. doi:10.1007/978-3-319-71767-8_9
[11]. Haldorai, A. Ramu, and S. Murugan, “Social Aware Cognitive Radio Networks,” Social Network Analytics for Contemporary Business Organizations, pp. 188–202. doi:10.4018/978-1-5225-5097-6.ch010
[12]. Haldorai and A. Ramu, “The Impact of Big Data Analytics and Challenges to Cyber Security,” Advances in Information Security, Privacy, and Ethics, pp. 300–314. doi:10.4018/978-1-5225-4100-4.ch016
[13]. H. Anandakumar and K. Umamaheswari, “A bio-inspired swarm intelligence technique for social aware cognitive radio handovers,” Computers & Electrical Engineering, Sep. 2017. doi:10.1016/j.compeleceng.2017.09.016
[14]. Hand D, Mannila H, Smyth P, ”Principles of data mining”. MIT Press, 2001.
[15]. 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.
[16]. 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.
[17]. Breiman L, Friedman JH, Olshen RA, Stone CJ, ”Classification and regression trees”. Wadsworth, Belmont, CA, USA, 1984..
[18]. 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.
[19]. Uysal I, Altay HG, ”An overview of regression techniques for knowl- edge discovery”. Knowl Eng Rev 14: 319-340, 1999.
[20]. L. Breiman, ”Random forests”. Machine Learning, vol. 45, no. 1, pp. 532, 2001.