ISSN:0975-9646

Comparative Study of Knowledge in Crop Diseases Using Machine Learning Techniques

P. Revathi , R. Revathi and Dr.M.Hemalatha
Dept of Software Systems ,Karpagam University, Coimbatore, Tamil Nadu, India.

Abstract: In this article a challenge has been made to analysis the explore studies on significance of data mining techniques in the field of agriculture. Couple of the techniques, such as decision algorithms C4.5, the SVM machine learning, ABC algorithm, and artificial neural networks applied in the field of agriculture was presented. Data mining in application of agriculture is a relatively new approach for forecasting / predicting of agricultural crop/animal management/soil management. This article explores the applications of data mining techniques in the field of farming and similar sciences.

Keybords: ABC algorithm, Machine Learning Techniques, SVM

1. INTRODUCTION

Information on new crop varieties is important to farmers when assessing whether to adopt these varieties. This information can be used as part of the farmer's decision-making process to help to improve crop production. Often changing to a newer crop variety will result in greater yields with little or no change in farm-resource outlays. Thereby, it is important to both the farmer and seed marketer that this variety information is accurate. Data mining is the extraction of hidden diagnostic information from huge databases, is a great new technology with immense potential to help companies focus on the most significant information in their data warehouses. Data mining tasks predict future trends and behaviors, allowing businesses to make proactive, knowledge-driven decisions. The automated, potential analysis presented by data mining move past the investigation of past procedures provided by retrospective tools typical of decision support systems. Agriculture and correlated activities constitute the only major component of India's gross domestic product, contributing nearly 25% of the total and nearly 60% of Indian population depends on this profession. Due to vagaries of climate factors the agricultural productivities in India are continuously decreasing over a decade. In this paper an attended dispute has been made to compile the research findings of different researcher's earlier capitalized data.

2. BENEFICIAL OF MINING TECHNIQUES IN FARMING

Presented the potential of three machine learning techniques viz. [1]DT induction using C4.5, RS and hybridized rough set based decision tree induction (RDT) in comparison to standard LR method. RS offers mathematical tools to discover hidden patterns in data and therefore its application in forewarning models needs to be investigated. A DT is a classification scheme which generates a tree and a set of rules representing the model of different classes from a given dataset. A java implementation of C4.5 (CJP) is used for DT induction. A variant of RDT called RJP, combines merits of both RS and DT induction algorithms. *Powdery mildew of Mango* (PWM) is a devastating disease and has assumed a serious threat to mango production in India resulting in yield losses of 22.3% to 90.4%. As a case study, prediction models for forewarning PWM disease using variables viz. temperature and humidity have been developed. The results obtained from machine learning techniques viz. RS, CJP and RJP are compared with the prediction model developed using LR technique. The techniques RJP and CJP have exposed better performance over LR approach.

Previous studies rice diseases decision trees widely used learning method and do not require any prior knowledge of data distribution, works well on noisy data .It has been applied [2] to classify Rice disease based on the symptoms. This paper intended to discover classification rules for the Indian rice diseases using the c4.5 decision trees algorithm. Expert systems have been used in agriculture since the early 1980s. Several systems have been developed in different countries including the USA, Europe, and Egypt for plant disorder diagnosis, management and other production aspects.

Web Based diseases [3] and treatment to the diseases which were effected to the chilly plants by using the mechanism of Rule based system and Artificial Bee Colony[1] (ABC) algorithm. The rules in the database is processed by the rule based system and if the required rules are not present in the database, then the system goes to the Machine learning algorithm technique used expert system. Thus, by applying machine learning techniques, resulting to best global optimized solution for recognizing the diseases in chilly plants.

The experiments were done using Radial Basis Function (RBF), polynomial and sigmoid kernel function. The SVM method based on RBF kernel function made the best performance for classification of cucumber leaf diseases [4].

Sweet Orange Crop Information Expert System deals with different varieties of Sweet Crop, Identification of various diseases diagnose used Rule based techniques and ABC Optimization Algorithms Algorithm generally occurs to Sweet Orange crop based on the symptoms. It contains two main parts one is Sweet Orange Information System and the other is Sweet Orange Crop Expert System where information system, the user can get all the static information about different species, Diseases, Symptoms, chemical controls, Preventions, Pests, Virus of Sweet Orange

fruits and plants. Depends on the response by the user the expert system decides the disease and displays its control measure of disease [5].

Unsupervised diseases pattern identification and classification algorithm that is based on a modified Hierarchical Dynamic Artificial Neural Network which provides an adjustable sensitivity-specificity herbs diseases detection and classification from the analysis of noise-free colored herbs images. It is also to future diseases treatment algorithm that is capable to provide a suitable treatment and control for each identified herbs diseases. [6] Firstly, noise removal algorithm for removal of wide range impulse noises and Gaussian noise that operating on the same image, which produces a hard damage. Secondly, machine learning system that can be used to identify the visual symptoms of herb plant diseases and this may have a particular application for farmer or crop producers in remote locations. Thirdly, the expert system that is capable to suggest the proper treatment and control of the identified diseases.

Images of the affected rice plants are captured by digital camera and processed using image growing, image segmentation techniques to detect infected parts of the plants. Then the infected part of the leaf has been used for the classification purpose using neural network. [7]The methods evolved in this system are both image processing and soft computing technique applied on number of diseased rice plants.

Early detection of diseases in tomato crops furthermore, Multi-Layer Perception (MLP), Learning Vector Quantization (LVQ) and Radial Basis Function (RBF) based Artificial Neural Network (ANNs) were used to learn to classify and hence categories the datasets. Using the [8] RBF, MLP and LVQ techniques we achieved 94, 96 and 98% classification accuracy for the healthy, powdery mildew (Oidium lycopersicum) and spider mite infected plants respectively. From these results it is evident that EN is capable of discriminating between the healthy and artificially infected tomato plants and hence may be deployed as a potential early disease detection tool for tomato crops in commercial greenhouses.

Early paper presented an Exploration Support System we are currently developing that is based on reinforcement learning and genetic algorithm techniques. [9]It defines interesting strategies in an incremental manner, and appears to be a promising approach to assist users to explore a wider range of solutions.

The warning and prediction system for crop diseases and pests based on Super Map IS. NET geographic information system (GIS), which was developed by Super Map Company. In this system, [10] the author used GIS and remote sensing (RS) technology. The system could transform data information into a geographical information map to show the occurrence degree and distribution on various diseases and pests.

Early detection and classification of plant diseases with Support Vector Machines based on hyper spectral reflectance discrimination between [11] healthy and inoculated plants as well as among specific diseases can be achieved by a support vector machine learning with a RBF function as kernel. The discrimination between healthy sugar beet leaves and diseased leaves resulted in classification accuracies up to97%.the multiple classifications between healthy leaves and leaves with symptoms of the three diseases still achieved accuracy higher than 86%.

Previous done using crop diseases risk helping gather and dissect data used by the ANN, [12] various tools and methods have been developed to quickly test different ANN models. Automation of processes such as variable selection and derivation through an object oriented framework has the potential to measurably improve the accuracy of the ANN model, and allow for easier implementation of similar models for a variety of crop diseases.

3. METHODS AND MATERIALS

3.1 ABC algorithm

ABC algorithm a population based algorithm, the position of a food source represents a possible solution to the optimization problem and the nectar amount of a food source corresponds to the quality (fitness) of the associated solution. The number of the employed bees is equal to the number of solutions in the population. At the first step, a randomly distributed initial population (food source positions) is generated. After initialization, the population is subjected to repeat the cycles of the search processes of the employed, onlooker, and scout bees, respectively. An employed bee produces a modification on the source position in her memory and discovers a new food source position. Provided that the nectar amount of the new one is higher than that of the previous source, the bee memorizes the new source position and forgets the old one. Otherwise she keeps the position of the one in her memory. After all employed bees complete the search process; they share the position information of the sources with the onlookers on the dance area.

3.2 Artificial Neural Network

Neural networks are also similar to the biological neural networks in the sense that functions are performed collectively and in parallel by the units, rather than there being a clear delineation of subtasks to which various units are assigned (see also connectionism). Recently the term Artificial Neural Network (ANN) tends to refer mostly to neural network models employed in statistics, cognitive psychology and artificial intelligence. NNmodels designed with emulation of the central nervous system (CNS) in mind are a subject of theoretical neuroscience and computational neuroscience.

3.3 C4.5 Algorithm

C4.5 is an algorithm used to generate a decision tree developed by Ross Quinlan. C4.5 is an extension of Quinlan's earlier ID3 algorithm. The decision trees generated by C4.5 can be used for classification, and for this reason, C4.5 is often referred to as a statistical classifier.

3.4 RDT Algorithm

RDT Algorithm, the decision table is given as input and predominant attributes called reduct is obtained as output. If the

data is large, vertical fragmentation is to be done and RDT can be applied to each fragment after adding the decision attribute. The predominant attributes for all fragments are obtained and they are grouped together with fragment information and decision attribute. To this RDT Algorithm is once again applied giving rise to, a new set of attributes called composite reduct.

3.5 Genetic Algorithm

GA,a population of strings (called chromosomes or the genotype of the genome), which encode candidate solutions (called individuals, creatures, or phenotypes) to an optimization problem, evolves toward better solutions. Traditionally, solutions are represented in binary as strings of 0s and 1s, but other encodings are also possible. The evolution usually starts from a population of randomly generated individuals and happens in generations. In each generation, the fitness of every individual in the population is evaluated, multiple individuals are stochastically selected from the current population (based on their fitness), and modified (recombined and possibly randomly mutated) to form a new population. The new population is then used in the next iteration of the algorithm. Generally, the algorithm terminates when either a maximum number of generations has been produced, or a satisfactory fitness level has been reached for the population. If the algorithm has terminated due to a maximum number of generations, a satisfactory solution may or may not have been reached.

4. RECOMMENDATIONS

As per above algorithms used my suggestion is that to use ABC Algorithm, SVM With RBF,ANN .Since using this algorithm various types of diseases were identified and they are being also controlled. Using these algorithms we can obtain accuracy and performance optimizations. These can be applicable in the field of medical and agriculture and financial evaluate.

5. CONCLUSION

Plenty of data is growing number of applications of data mining techniques in agriculture and a emergent amount of data that are currently available from lots of resources. This is relatively a novel research field and it is estimated to develop in the future. This is fairly a novel research field and it is expected to grow in the future. There is several of work to be done on this up-and-coming and interesting research field. The multidisciplinary approach of integrating computational with agriculture will help in forecasting/managing agricultural crops effectively.

REFERENCE

1. Rajnijain, Sonajharia Minz1 and Ramasubramanian .V, ***Machine Learning for Forewarning Crop Diseases.*** J. Ind. Soc. Agril. Statist.63 (1), 2009, PN: 97-107.
2. Nithya.A, And Dr.V.Sundaram**,** *Classification Rules For Indian Rice Diseases***,** IJCSI International Journal Of Computer Science Issues, Vol. 8, Issue 1, January 2011,ISSN (Online): 1694-0814,PN:444-448.
3. Prasad Babu.M.S, Sridhar. DVPR Etc. Al. / (IJCSE) International Journal On Computer Science And Engineering, ***Implementation Of Web-Based Chili Expert Advisory System Using ABC Optimization Algorithm*** Vol. 02, No. 06, 2010, PN: 2141-2144.
4. Zhangjian, Zhangwei, Inst. Of Intel. Machines, Chinese Acad. Of Sci., Hefei, China, ***Support Vector Machine For Recognition Of Cucumber Leaf Diseases,*** Advanced, and Issue Date: 27-29 March 2010 PN: 264 – 266.
5. Prasad Babu.M.S Etc. Al. / International Journal of Engineering Science and Technology, ***A Web Based Sweet Orange Crop Expert System Using Rule Based System And Artificial Bee Colony Optimization algorithm*** Vol. 2(6), 2010, PN: 2408-2417.
6. Lili N. A / International Journal on Computer Science and Engineering**,** ***Classification Of Herbs Plant Diseases Via Hierarchical Dynamic Artificial Neural Network After Image Removal Using Kernel Regression Framework,*** ISSN: 0975-3397, Vol. 3 No. 1 Jan 2011, PN: 15-20.
7. Phadikar, S.; Sil, J.; Dept. Of CSE, West Bengal Univ. Of Techno, Kolkata. Computer and Information Technology, 2008. ICCIT 2008. 11th International Conference. ***Rice Disease Identification Using Pattern Recognition Techniques,*** Issue Date: 24-27 Dec. 2008, PN: 420 – 423.
8. Ghaffari, R.; Fu Zhang; Iliescu, D.; Hines, E.; Leeson, M.; Napier, R.; Clarkson, J.; Univ. Of Warwick, Coventry, UK. Neural Networks (IJCNN), the 2010 International Joint Conference, ***Early detection of Diseases in Tomato Crops: An Electronic Nose And Intelligent Systems Approach***, Issue Date: 18-23, July 2010, PN: 1 - 6, ISSN: 1098-7576.
9. Attonaty.J.M, Chatelin.M.H, Garcia.F, Ndiaye.S, *Using Extended Machine Learning and Simulation Technics to Design Crop Management Strategies.*
10. Juhua Luo, Wenjiang Huang, Jihua Wang And Chaoling Wei, ***The Crop Disease And Pest Warning And Prediction System***, Computer And Computing Technologies In Agriculture II, Volume 2 IFIP Advances In Information And Communication Technology, 2009, Volume 294/2009, PN: 937-945,
11 Rumpf.T, Mahlein.A.K, Steiner.U, Oerke.E.C, Dehne.H.W and Plümer.L, ***Early Detection And Classification of Plant Diseases With Support Vector Machines Based On Hyper spectral Reflectance***, Vol: 74, Issue 1, October 2010, PN: 91-99
12 .Jason A Smith, Kathleen M Baker, Robert Trenary, 7th World Congress On Computers In Agriculture Conference Proceedings, Mining Derived Weather Forecast Variables For Crop Disease Risk Prediction, American Society Of Agricultural And Biological Engineers, St. Joseph, Michigan Www.Asabe.Orgjune 2009, PN:22-24.

P,Revathi was born an brought up coimbatore. She received her Bachelor of Business Administration from Madras University. and Master degree in Computer Applications from Bharathiyar University. She completed her M.Phil from P,S,G Krishnammal colleges for women. Currently pursuing Ph.D., in computer science at Karpagam University under the guidance of Dr.M.Hemalatha Head, Dept of Software System, Karpagam university, Coimbatore. Area of Research is Data mining and Machine learning .

R,Revathi was born an brought up Trichy She received her Bachelor of Computer Science from Bharathidasan University. and Master degree in Information Technology from Bharathidasan University. She completed her M.Phil from Bharathidasan University. Currently pursuing Ph.D., in computer science at Karpagam University under the guidance of Dr.M.Hemalatha Head, Dept of Software System, Karpagam University, Coimbatore. Area of Research is Image Processing and Networks.

M. Hemalatha completed MCA, MPhil., PhD in Computer Science and currently working as an Assistant Professor and Head, Department of software systems in Karpagam University. Ten years of Experience in teaching and published Twenty seven papers in International Journals and also presented seventy papers in various National conferences and one international conferences Area of research is Data mining, Software Engineering, bioinformatics, Neural Network. She is also reviewer in several National and International journals.