ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 10, October 2015

An Identification of Variety of Leaf Diseases Using Various Data Mining Techniques

Dr. N. Sasirekha, N. Swetha

Rathinam College of Arts and Science

Abstract: Diseases in plants cause major production and economic losses as well as reduction in both quality and quantity of agricultural products. Now a day's plant diseases detection has received increasing attention in monitoring large field of crops. Farmers experience great difficulties in switching from one disease control policy to another. The naked eye observation of experts is the traditional approach adopted in practice for detection and identification of plant diseases. This paper review the need of simple plant leaves disease detection system that would facilitate advancements in agriculture. Early information on crop health and disease detection can facilitate the control of diseases through proper management strategies. This technique will improves productivity of crops. This paper surveys the data mining techniques to predict the various leaf diseases of various types of leaf. Different types of leaf diseases of also discussed in this paper.

Keywords: Data mining, Biology diversity, Agriculture.

I. INTRODUCTION

India is an agricultural country wherein most of the population depends on agriculture. Research in agriculture is aimed towards increase of productivity and food quality at reduced expenditure, with increased profit. **Agricultural production system is an outcome of a complex interaction of soil, see**d, and agro chemicals. Vegetables and fruits are the most important agricultural products. In order to obtain more valuable products, a product quality control is basically mandatory. Many studies show that quality of agricultural products may be reduced due to plant diseases. Diseases are impairment to the normal state of the plant that modifies or interrupts its vital functions such as photosynthesis, transpiration, pollination, fertilization, germination etc. These diseases are caused by pathogens viz., fungi, bacteria and viruses, and due to adverse environmental conditions. Therefore, the early stage diagnosis of plant disease is an important task. Farmers require continuous monitoring of experts which might be prohibitively expensive and time consuming. Leaf presents several advantages over flowers and fruits at all seasons worldwide. This paper provides an introductory part includes importance of leaf disease detection; plant leaves analysis, various types of leaf diseases. By concludes this paper along with possible future directions.

II. DATA MINING

Data mining refers to extracting or mining the knowledge from large amount of data. The term data mining is appropriately named as 'Knowledge mining from data' or "Knowledge mining". Data mining is used to extract implicit and previously unknown information from data. Data collection and storage technology has made it possible for organizations to accumulate huge amounts of data at lower cost. Exploiting this stored data, in order to extract useful and actionable information, is the overall goal of the generic activity termed as data mining.

Data mining is the process of exploration and analysis, by automatic or semiautomatic means, of large quantities of data in order to discover meaningful patterns and rules. Data mining is an interdisciplinary subfield of computer science which involves computational process of large datasets patterns discovery. The goal of this advanced analysis process is to extract information from a data set and transform it into an understandable structure for further use. Data Mining is about solving problems by analyzing data already present in databases. Data mining is the process which provides a concept to attract attention of users due to high availability of huge amount of data and need to convert such data into useful information. Data mining is also stated as essential process where intelligent methods are applied in order to extract the data patterns.

An Overview of the Research on Plant Leaves Analysis

Plant leaf diseases identification has a broad application prospective in agriculture and medicine, and is especially significant to the biology diversity research. Plant leaf classification finds application in botany and in tea, cotton and other industries. Plants are vitally important for environmental protection. However, it is an important and difficult task to recognize plant species on earth. Many of them carry significant information for the development of human society. The urgent situation is that many plants are at the risk of extinction. So it is very necessary to set up a database for plant protection. The first step is to teach a computer how to classify plants. Leaf recognition plays an important role in plant classification. Plants are basically identified based on flowers and fruits. However these are three dimensional objects and increases complexity. Plant identification based on flowers and fruits require morphological features such as number of stamens in flower and number of ovaries in fruits. Identifying plants using such keys is a very time consuming task and has been carried out only by trained botanists. However, in addition to this time intensive task, there are several other

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 10, October 2015

drawbacks in identifying plants using these features such as the unavailability of required morphological information and use of botanical terms that only experts can understand. However leaves also play an important role in plant identification. Moreover, leaves can be easily found and collected everywhere at all seasons, while flowers can only be obtained at blooming season. Shape of plant leaves is one of the most important features for characterizing various plants visually. Plant leaves have two-dimensional nature and thus they are most suitable for machine processing. Our paper presents survey of different leaf diseases identification.

III. TYPES OF DISEASES IN PLANT LEAF

Generally the diseases of plant in two forms the diseases can be:

- Infectious diseases caused by Fungi, Bacteria, Viruses, etc.
- Non-infectious diseases or disorders caused by mineral toxicities, soil acidity, nutrient deficiencies, or environmental factors.

Bacterial disease symptoms

The disease is characterized by tiny pale green spots which soon come into view as water- soaked. Then it appear as dry dead spots as e.g. bacterial leaf spot have brown or black water-soaked spots on the foliage, sometimes with a yellow halo, generally identical in size. Under dry conditions the spots have a speckled appearance.

Viral disease symptoms

Among all plant leaf diseases, those caused by viruses are the most difficult to diagnose. Viruses produce no telltale signs that can be readily observed and often easily confused with nutrient deficiencies and herbicide injury. Aphids, leafhoppers, whiteflies and cucumber beetles insects are common carriers of this disease, e.g. Mosaic Virus, Look for yellow or green stripes or spots on foliage. Leaves might be wrinkled, curled and growth may be stunted. (a) Bacterial leaf spot (b) mosaic virus Bacterial and Viral disease on leaves.

Fungal disease symptoms

Among all plant leaf diseases, those caused by fungus some of them are discussed below and e.g. Late blight caused by the fungus Phytophthora infesters. It first appears on lower, older leaves like water-soaked, gray-green spots. When fungal disease matures, these spots darken and then white fungal growth forms on the undersides.

Early blight is caused by the fungus Alternariasolani. It first appears on the lower, older leaves like small brown spots with concentric rings that form a bull's eye pattern. When disease matures, it spreads outward on the leaf surface causing it to turn yellow. In downy mildew yellow to white patches on the upper surfaces of older leaves occurs. These areas are covered with white to greyish on the undersides. (a) Late blight (b) early blight (c) downy mildew Figure

(a)Bactria leaf spot

(b) Vangi affected plant

Bacterial and Viral disease on leaves

(a) Late blight

(b) early blight

(c) downy mildew

Fungal disease on leaves

DOI 10.17148/IJARCCE.2015.41098

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 10, October 2015

Table 1: Different diseases and their techniques

PLANT SPECIES	DISEASES	TECHNIQUES	AUTHOR & YEAR
Rice	a) Bacterial leaf blight b) Bacterial leaf streak c) Blast disease; d) Brown spot disease e) Leaf scald; f) Narrow brown leaf spot g) Red stripe h) Sheath blight; i) Tungro.	Image processing, Chaos Theory, fractal dimension.	V.Surendrababu Research Scholar Dr.C.P.Sumathi AsstProff Jan 2014
Cottan	a) Bacterial blight b) Fusarium wilt c) Leaf blight d) Root rot e) Micro nutrient f) Verticillium wilt	Cross Information Gain Deep forward Neural Network (CIGDFNN) classifier.	P.Revathi Jan 2014
Wheat	a)Powdery mildew b)Leaf rust c)Septoriatriticileaf blotch d)Stagonospora(Septoria) nodorum leaf e) Glume blotch, f) Fusariumhead scab.	Decision Trees, Pruning, Datamining, Classification, Expert System, Neural Networks	A.Nithya, Dr.V.Sundaram September-2011
Tomato	a)Altercank b) Spot powder c) Stemphyllium d) peppery spot	Content-based image retrieval (CBIR), Principle Component Analysis(PCA)	D.N.D.Harini Research Scholar D.LalithaBhaskari AsstProff (2011)
Wheat	a)Powdery mildew b)Leaf rust c)Septoriatriticileaf blotch d)Stagonospora(Septoria) nodorum leaf e) Glume blotch, f) Fusariumhead scab.	Decision Trees, Pruning, Datamining, Classification, Expert System, Neural Networks	A.Nithya, Dr.V.Sundaram September-2011
Sugar beet leaf	a) Cercospora leaf spot b) Sugar beet rust c)Powdery mildew	Classification, Artificial neural networks, Decision Trees, Support Vector Machines	T. Rumpf (2010)
Coffee	a) Coffee leaf rust b) Antestia, c) Lepidopterous miners d)Defoliators e)Berry borer f) Stphanorders	Coffea, Colletotrichumkahawae, Hemileiavastatrix, coffee breeding, resistance.	FuadAbafitaAbabulgu June 2010
Coconut	a)Promecothecacumingii b)Brontispalongissima c) Plesispareichei d) Oryctes rhinoceros e)Artonacatoxantha	Classification, SVM	Mat Hassan Othman Nov 2009
Coca	a)Pathogens b)Frosty pod c)Witches broom d)Black pod	Sptial temporal distribution	Heinrich Lehmann Germany April 15, 2003

DOI 10.17148/IJARCCE.2015.41098

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 10, October 2015

IV. RESULT AND DISCUSSION

All plant leaf and their diseases are finding by using their related techniques and the diseases affected by the plants are generally in bacterial, fungal and virus. Merely all the leaf diseases finding techniques are processed by Classification and Association rule. Leaf changes their color by diseases will affect the plant. And the diseases are occurred in seed, root, stem and leaf. Images of plant will easy to find the energy balance, diseases and their changes. The overall approach used in this study follows the Knowledge Discovery in Databases (KDD) process to identify disease co-occurrences at different levels of abstraction from a public health dataset consisting of codes.

The goal of this preliminary study was to explore the impact of generalizing data and the effect of using different concept hierarchies in the plant leaf diseases. Using domain knowledge to create a concept of hierarchy and incorporating the hierarchy into the data mining process by abstracting data to a higher level concept can help constrain the search space and enhance performance and results of the mining process. This study examined two concept hierarchies for generalization and involved maintaining links between the generalized and raw data. The results differed based on the concept hierarchy used for generalizing the data.

Once validated, disease-specific knowledge obtained using automated methods such as association rule mining could be used for applications such as quality of care, decision support, and hypothesis generation. The associations found in this study show the potential for discovering potentially new and interesting knowledge related to disease co-occurrences at different levels of abstraction. The identity associations found could help flag possible irrelevant disease co-occurrences and the new associations could serve as a starting point for further exploration. Established medical knowledge sources and more in depth review by domain experts will be done in future studies to determine the clinical validity of the new associations generated in this study. In addition, calculating other statistical or "interestingness" measures could be used to further assess the strength of each association.

V. CONCLUSION

This study explored the use of generalized rule mining for generalizing data to a higher level of abstraction in a public health dataset and compared disease co-occurrences at the different levels. The results suggest that maintaining links from the generalized data to the raw data may help constrain the search space and rule space, minimize information loss, and discover new associations. In addition, use of different concept hierarchies may have an impact on the outcome.

REFERENCE

[1] H. L. Khatibl, F. Hawels, H. Hamdi, and N. L. Mowelhi. "Spectral Characteristics Curves of Rice Plants Infected by Blast," IEEEGeoscience and Remote Sensing Symposium, vol. 2, pp. 526-528,1993.

[2] Bernardes.A.A, J.G.Rogeri, N.Marranghello, A. S. Pereira, A.F. Araujo and João Manuel R. S. Tavares. "Identification of Foliar Diseases in Cotton Crop". SP, Brazil.

[3] Bettencourt, A.J. and Podrigues C.J. (1985). Principles and practice of coffee breeding for resistance to Rust and other diseases.**In:** coffee agronomy vol.4 (Clarke R.J and Macrae R. eds) 43-90pp. Booker Agriculture International Limited, Thame, Oxon., UK.

[4] Anon. 1999. Control of Plesispareicheiin coconut nurseries, Report of Crop Protection Division, Annual Report of the Coconut Research Institute of Sri Lanka.p. 151.

[5] Brown, Charles, Frank Chaloupka, and Michael Grossman. "The Demand for Cocaine by Young Adults: A Rational Addiction Approach." NBER Working PaperSeries.Working Paper 5713. August 1996.

[6] Lili N. A / International Journal on Computer Science and Engineering,Classification Of Herbs Plant Diseases Via Hierarchical Dynamic ArtificialNeural Network After Image Removal Using Kernel RegressionFramework, ISSN: 0975-3397, Vol. 3 No. 1 Jan 2011, PN: 15-20

[7] Lili N. A / International Journal on Computer Science and Engineering,Classification Of Herbs Plant Diseases Via Hierarchical Dynamic ArtificialNeural Network After Image Removal Using Kernel RegressionFramework, ISSN: 0975-3397, Vol. 3 No. 1 Jan 2011, PN: 15-20

[8] Devraj and Jain R (2011) PulsExpert: An expert system for the diagnosis and control of diseases in pulse crops. AgriSci J 38(9).

[9] Subbaiah, K. M., In Proceedings of the First International Workshop on Biological Control of Chromolaenaodorata (ed)

[10] Pretorius ZA, Singh RP, Wagoire WW, Payne TS (2000) Detection of virulence to wheat stem rust resistance gene Sr31 in Pucciniagraminisf. sp. triticiin Uganda. Plant Dis 84:203

[11] Attonaty.J.M, Chatelin.M.H, Garcia.F, Ndiaye.S, Using Extended Machine Learning and Simulation Technics to Design Crop Management Strategies.

[12] Shalini Sharma, Vishal Shrivastava, International Journal on Recent and Innovation Trends in Computing and Communication , ISSN 2321 – 8169 Volume: 1 Issue: 4, March 2013.

[13] UshaSinha, and HooshangKangarloo, "Principal Component Analysis for Content-based Image Retrieval" in September 2002 RadioGraphics, 22, 1271-1289.

[14] Jan Knorn, Andreas Rabe, Volker C. Radeloff, Tobias Kuemmerle, JacekKozak and Patrick Hostert, "Land cover mapping of large areas using chain classification of neighboring Landsat satellite images", ELSEVIER Science Direct 2009.