



# Medicinal Analysis Based On Agri And Human Healthcare Under Mining Techniques

PAVITHRA.A

DEPARTMENT OF CSE  
MASTER OF ENGINEERING  
ARJ COLLEGE OF ENGG AND TECH  
MANNARGUDI, THIRUVARUR(DT)  
TAMILNADU, INDIA  
Paviuma296@gmail.com

**Abstract**— The medical environment is still information rich but knowledge weak. However, there is a lack of powerful analysis tools to identify hidden relationships and trends in data for plants crops. Disease for plant is based on soil moisture, temperature, water level, pesticide and other environmental conditions and for human it is a based on health care conditions related to Human. These medical conditions describe the unexpected health conditions that directly control all its parts based on the location, weather and some attributes.

**Keywords**— Support vector machine, artificial neural networks

## I. INTRODUCTION

Data mining (the analysis step of the "Knowledge Discovery in Databases" process, or KDD), an interdisciplinary subfield of computer science, is the computational process of discovering patterns in large data sets ("big data") involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use. Aside from the raw analysis step, it involves database and data management aspects, data pre-processing, model and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures, visualization, and online updating.

A major aim of many forecasting system is to reduce fungicide, and accurate prediction is important to synchronize the use of disease control measures to avoid crop losses. A prediction model based on the relationship between environmental conditions at the time of management and late-season disease severity could be used to guide management decisions the factors that govern disease epidemics and the design of control systems that minimize yield losses. The same models have potential to guide breeding programs and

work to develop strategies that will prolong the usefulness of disease-resistance genes.

## II. SYSTEM MODEL

This paper investigates integrating support vector machine (SVM) based Fuzzy logic with decision tree in the diagnosis of disease for Farmer and Human. It also investigates different methods of initial centroid selection of the clustering value such as inlier, outlier, range, random attribute values at risk, and random row methods in the diagnosis of diseases in plant and human based on the mean, median and mod values. It also showed that the inlier initial centroid selection method could achieve higher accuracy than other initial centroid selection methods in the diagnosis of plant and human.

### A. Proposed technique

Decision Tree is one of the data mining techniques used in the diagnosis of value at risk of disease showing considerable success. Fuzzy logic clustering is one of the most popular clustering techniques; however initial centroid selection strongly affects its results. Various techniques of computer modeling and simulation viz. machine learning techniques like artificial neural networks and the conventional multiple regression approaches are being used to help synthesize and develop understanding of this pathogen environment relationship. Support vector machine can also be used as the technique to predict the disease in both crop and also in human health conditions.

## III. IMPLEMENTATION

### A. Disease data collection in crop and human

A web database is an organized collection of database in website. The data is typically organized to model aspects of reality in a way that supports processes requiring information. In this database consists of inadequate value to predict

diseases complications. The benign and malignant is the paradigm tumor marker for management of patient diseases value at risk. For more than certain decades, benign and malignant has been the most important biomarker benign and malignant testing in disease value is collected and stored this module for the input data set for input processing for comparison and prediction.

### B. Query processing

Title User Query is activity of obtaining information resources relevant to an information need from a collection of information resources. Searches can be based on metadata. An information retrieval process begins when a user enters a query into the system. Queries are formal statements of information needs, for example search strings in collected database. In information retrieval a query does not uniquely identify a single object in the collection. The database are typically transformed into a suitable representation. Each retrieval strategy incorporates a specific model for its database representation purposes. The patient needs to retrieve the information for (ADC) value in differentiating benign and malignant disease and evaluating the detection accuracy of the disease value at risk extension. Exact value predict as per the data querying among the request.

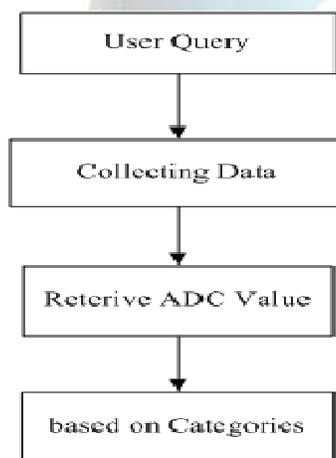


Fig: user query about disease

### C. Support vector machine

A benign tumor is a mass of cells that lacks the ability to invade neighboring tissue. These characteristics are required for a tumor value at risk to be defined as diseases. But it may identified by precautions using SVM. In the field of artificial intelligence algorithms, which refers to combinations of artificial neural networks and fuzzy logic. Using this fuzzy logic algorithms, to comparisons the normalized and transformed values like mean, median, and mod of criteria are

being determined by value at risk from the defined group of possible value are evaluated.

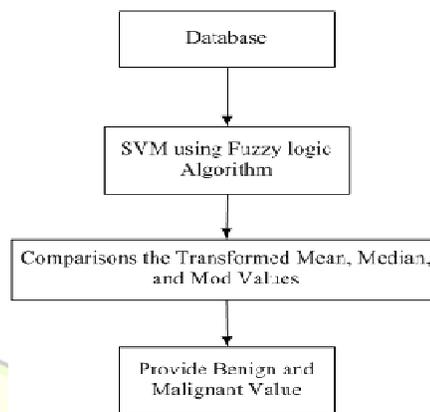


Fig: svm based prediction

### D. Prediction of diagnostic value

In order to predict the benign and malignant concentration at the end of each interval by using the SVM based Fuzzy logic. After proper matching, the output from the fuzzy algorithm will be the predicted benign and malignant concentration at the end of the interval. The experimental results of this proposed system by using different wavelet families in the hidden layer. The data stream mining case, where SVM systems are sequentially updated with new incoming samples. These benign and malignant are rather small, indicating a promising performance, and comparison between these different input selection methods and predictor based on value at Risk.

### E. System architecture

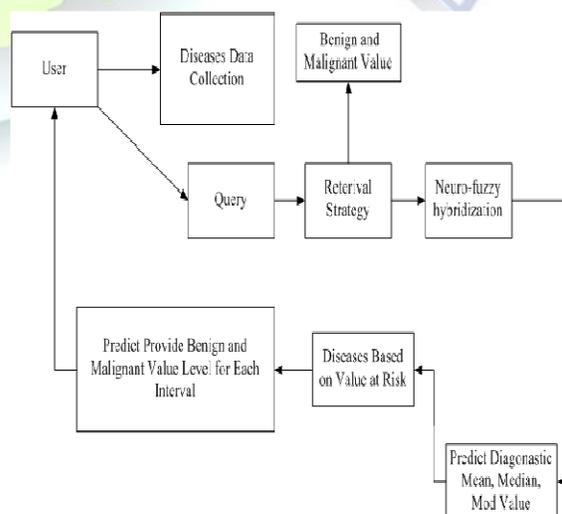




Fig:3 system architecture for prediction of disease.

#### F. Context Analysis Algorithm

1) *Fuzzy logic algorithm*: Fuzzy logic is a form of many-valued logic in which the truth values of variables may be any real number between 0 and 1. By contrast, in Boolean logic, the truth values of variables may only be 0 or 1. Fuzzy logic has been extended to handle the concept of partial truth, where the truth value may range between completely true and completely false. Furthermore, when linguistic variables are used, these degrees may be managed by specific functions.

#### G. Support vector machine

In machine learning, support vector machines (SVMs, also support vector networks) are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. Given a set of training examples, each marked for belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other, making it a non-probabilistic binary linear classifier. An SVM model is a representation of the examples as points in space mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on. This paper introduces a new prediction method based on a powerful machine learning technique Support Vector Machines (SVM) is a very effective method for general-purpose supervised predictions. It has been shown in the past that this machine learning technique is very effective in the classification of proteins particularly in discriminating membrane proteins, prediction of subcellular localization, solvent accessibility, CTL epitopes, binding peptides, protein structures, protein-protein binding sites and gene expression level. SVM provides an alternative or complement to the present ANN and REG based approaches for model development. The SVM learns how to classify from a training set of feature vectors, whose expected outputs are already known. The training enables a binary classifying SVM to define a plane in the feature space, which optimally separates the training vectors of two classes. When a new feature vector is fed, its class is predicted on the basis of which side of the plane it maps. To the best of authors' knowledge, there is no report of using support vector machines in understanding the relationship between disease severity and its associated environmental conditions.

#### H. Pseudocode for sequence matching in database

Pseudo code for to Binary(SeqA )

// input: character sequence SeqA

// output: binaryA, the binary form of SeqA

- i. initialize binaryA = [ ] //empty string
- ii. for each character m in the sequence SeqA  
if (m == 'A' || m=='a') suffix=[0 , 0 , 0]  
elseif (m=='C' || m=='c') suffix=[0 , 0 , 0]  
elseif (m=='G' || m=='g') suffix=[0 , 0 , 0]  
elseif (m=='T' || m=='t') suffix=[0 , 0 , 0]  
else suffix=[1 , 0 , 0]  
endif
- iii. append suffix to end of binary
- iv. return binary

Pseudo code for BA\_Sequence Match (SeqA, SeqB )

// input: character sequences SeqA and SeqB

//output: the percentage of similarity of given sequences SeqA and SeqB

- i. binaryA = toBinary (SeqA)
- ii. binaryB = toBinary (SeqB)
- iii. compute result = XNOR (binaryA, binaryB)
- iv. initialize final-result = [ ] // empty string
- v. for i = 1 to length (result), step-increment=3if  
(result[i] == result[i+1]== result[i+2] ==  
1)append 1 to final-result else append 0 to final-  
result  
endif
- vi. compute count = number of 1's in final-result
- vii. match-percent = (count \* 100) / length (final-result)
- viii. return match-percent.

#### IV. CONCLUSIONS

The In this way we concluded that SVM is beneficial tool for crop prediction. In this paper includes the parameter of their regional soil parameter and disease data prediction. Then it is analyses by using neuro fuzzy logic algorithm using SVM. In an analyses of using the algorithm of fuzzy logic algorithm is a Crop simulation models vary greatly between them. Some of them are rather hard to use and parameterize. The need for calibration can be data extensive and applicable to some developing areas. However, this problem is hard to overcome in the near-future because of the inability of SVM to estimate yield in mixed agriculture. But, the increased availability of high-spatial resolution SVM at a reasonable cost make this technique a possible interesting alternative for yield forecast.

#### ACKNOWLEDGMENT

First of all I thank the almighty for giving knowledge.

I thank our honorable chairperson Smt. RAJAKUMARI AYYANATHAN, A.R.J College of Engineering and Technology, managed for having given us the opportunity to do the course. I wish to extend my sincere gratitude to the Head of the Department of Computer science and Engineering



and the Project coordinator, Mrs. V. SNEHALATHA M.E., who has always a constant source inspiriting and encouragement to me during the course and my heartfelt thanks to my project coordinator for her guidance from the starting of the project was really grateful without which I might not have been able to succeed in my work.

#### REFERENCES

- [1] China, 'Support Vector Machine For Recognition Of Cucumber Leaf Diseases, Advanced', and Issue Date: 27-29 March 2010 PN: 264 – 266.
- [2] 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.
- [3] 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.
- [4] 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.
- [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] 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.

