

A Survey on Data Mining and Image Processing Techniques Used in Agriculture

Nilesh Deshmukh¹, Aarti Gaikwad², Poonam Ghutukade³, Deepti Moogi⁴, Madhura Phadke⁵

¹Assistant Professor, Department of Computer Engineering, Savitribai Phule Pune University

^{2, 3, 4, 5}UG Students, Department of Computer Engineering, Savitribai Phule Pune University

Abstract: Data mining is used to extract the information from the large dataset and used to predict patterns and behavior of an application. Data mining plays a chief role in the fields of e-commerce, healthcare sector, and agricultural sector. Agriculture is the prime occupation in India. Crop productivity mainly depends on weather conditions. Data mining is used in agriculture to predict crop productivity, water management, crop disease management, pesticide recommendation by using different algorithms. Image processing techniques are also widely used to identify crop diseases using images of affected crops. The purpose of this survey is to study data mining and image processing techniques used in crop disease prediction and identification.

Keywords: Disease, Weather conditions, Prediction, Data mining, Image processing.

I. INTRODUCTION

Agriculture is the primary occupation of the people, i.e. about 70% of people are engaged in this activity. Weather conditions play a vital role in the agricultural sector. The impact of changing climatic conditions like temperature, humidity and rainfall is crucial in the growth of crop diseases. It is necessary to prevent the crop diseases in order to enhance the crop production. Earlier the monitoring and analysis of crop diseases were carried out manually by an expert person in that field. The crop disease detection can be done using various data mining and image processing techniques. Data mining is the process of obtaining knowledge from vast data using various data mining algorithms. Predictive analysis is one of the important part of data mining. Image processing helps in agricultural sector for determining the crop diseases and for pesticide recommendation.

II. DATA MINING AND IMAGE PROCESSING TECHNIQUES

A. Prediction of Weather Using Data Mining [1]

Prediction of plant disease is one of the emerging component in the agricultural systems which helps in disease management. Predicted weather data can provide timely information for crop management. To prevent the onset of diseases, future weather predictions is required. Forecasting systems based on weather helps farmers to take the control measures and necessary precautions. An important aim of weather forecasting system is to reduce pesticide use and crop losses. The production of grapes is seen in every continent under variable climatic conditions. India is one of the highest grape producing country. Grapes are affected by various diseases like downy mildew, powdery mildew and anthracnose. In one two data mining algorithms were used for weather prediction. They are-

- 1) *k-NN approach*: The k-NN approach is a pattern recognition method where the target object and its features are matched with the patterns among the object space. Here the pattern of observed weather data for every single day is matched with the same features for the same day in each year of historical recorded weather data. Euclidean distance is calculated to compare the target pattern and historical pattern. The Euclidean distance helps to find the nearest neighbour of the target data. The prediction for the next day is selected from the nearest k patterns (all weather features are considered) which is generated using random sampling process. Here, the modified k-NN approach is used where it considers average of k-days as a representative of forecasted weather data for the next day.
- 2) *Feed Forward Neural Network*: Feed forward neural network is a type of artificial neural network. In feed forward network the connectivity between the units do not form a directed cycle. In this network, the information moves forward in a single direction. The information moves from input nodes to the hidden node and then to the output node. Feed forward network contains many layers. The layers are interconnected with the previous layers. It is used for input to output mapping. There is a hidden layer and many neurons between the input and the output layer.

The proposed method [1] firstly forecasted the weather using the historical weather data with a help of two approaches

Modified k-NN algorithm and 2) Feed Forward Neural Network and then the disease model is developed based on regression analysis. Using these results the grape diseases are predicted.

B. CBIR and Naive Bayes Technique Used to Predict the Crop Disease

The proposed system [2] used Content Based Image Retrieval (CBIR) which is an image processing technique for image comparison. The essential features that CBIR technique uses are - colour, shape, and texture. In this system the user has to upload the image of the crop where the features of image are extracted using CBIR. The database contains similar feature images stored in a form of cluster. The system compares the query image with the stored image using features and gives the most closely matching image as an output to the user. Naive Bayes is a data mining approach which is used in [2] Naive Bayes is a classification algorithm. In [2] the user is expected to enter the crop details like the weather conditions, soil type, area etc which are then processed using the Naive Bayes algorithm to predict the crop disease. This data mining algorithms can work on a very large dataset. The proposed system consists of five modules listed as a) Test data acquisition module, b) Data pre-processing module, c) Image comparison module(CBIR), d) Data mining module (Naive Bayes) and e) Result analysis and the testing module.

C. Image Processing edge Detection Technique used to identify the Cotton Leaf spot Diseases

The system proposed in [3] used the digital images that are obtained from the mobile . Then, image-processing techniques are applied on the acquired images to extract RGB Pixel counting features that are important for analysis. Some perceptive techniques are applied to classify the images according to the specific problem.

This system consists of two phases to identify the disease. First 'Edge Detection Based Image Segmentation' is done. Then image analysis and classification of diseases are performed using HPCDD Algorithm. The aim is to develop an Advanced Computing System that can identify the a cotton leaf spot disease by using the image analysis technique.

This proposed system [3]uses a Homogeneous Pixel Counting technique for Cotton Diseases Detection (HPCDD). The steps followed by the system are,-I. Acquisition of RGB image, II. Creating color transformation structure, III. Converting the color values in RGB , IV. Applying Color Filtering, V. Masking green pixels, VI. Removing the masked cells inside the boundaries of the infected clusters, VII. Edge detection (using Sobel and Canny with Homogeneous operator techniques), VIII. Pixel Ranging function is used to calculate the RGB features (for every disease), IX. Texture Statistics Calculation , X. Disease Reorganization .

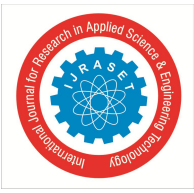
In [3] The image of a leaf is taken as an input image and is converted into a grey-scale image. Then color filter is applied and affected leaf spot color is used for RGB Pixel counting values and features are segmented. After segmentation edge detection is done using Canny and Sobel Edge detection homogeneous techniques. These techniques are used to recognize the clarity for the edges, and the centered pixel of opposite neighbouring one, two, three pixels of clarity of edges having white lightning are obtained. This identified lightning is denoted as the diseases that the plant is affected with, as it is represented in the edges of leaf boundary.

D. Pesticide Recommendation using Data Mining

The aim of the system proposed in [4] is to predict the occurrence of cotton crop disease due to climatic changes. The decision tree data mining algorithm is used. The attributes considered for decision making are crop name, location, pest population, temperature, humidity, rainfall, sunshine value, symptoms and recommended pesticides. These attributes are used to predict the disease and recommend the pesticide for cotton crop. The procedure applied for pesticide recommendation is – 1. Data collection 2. Data Pre-processing 3. Replace the missing values 4. Float to integer 5. Feature selection 6. Prediction using the decision tree. Data pre-processing is applied to convert the data in desired form. Feature selection helps to select the attributes that would give accurate predictions. This is done using alpha investing algorithm. Alpha investing is used to reduce false ratio. Decision tree is built using gain ratio. GUI is designed for pesticide recommendation in this system. This model will help the farmers to increase their food productivity and predict the result in the form of profit and loss percentage for future. In this system, initially the cotton crop pest is predicted and corresponding recommendation is provided for the end user. It is concluded that the decision tree algorithm gives an accuracy of about 85% for a large data-set.

III. CONCLUSION

The various data mining algorithms like k-NN, Feed Forward Neural Network, Naive Bayes and Decision Tree helps the people in agricultural sector to predict and identify the crop diseases according to changing weather conditions at an early stage. This in turn helps them to take the necessary precautions in order to save the crop. The image processing techniques also help to identify the



crop diseases and help in recommending the pesticides. These advanced techniques which are being used in agricultural sector has enhanced the crop productivity.

REFERENCES

- [1] S. Sannakki, V.S. Rajpurohit, F. Sumira, H. Venkatesh, "A Neural Network approach for Disease Forecasting in Grapes using Weather Parameters ", IEEE., 2014
- [2] Samiksha Bhor, Shubha Kotian et.al., "Developing an Agricultural Portal for Crop Disease Prediction", IJIRST –International Journal for Innovative Research in Science & Technology, April 2017
- [3] P. Revathi, M. Hemalatha, "Classification of cotton leaf spot diseases using image processing edge detection techniques"IEEE,2012
- [4] S. Pudumalar, K. S. Suriya, Ramanujam, S. Muthuramalingam "Pesticide recommendation system for cotton crop diseases due to the climatic changes "IJMREM-International Journal of Modern Research in Engineering & Management, January 2018