Early Detection of Grapes Diseases Using Machine Learning and IoT

Suyash S. Patil[1], Sandeep A. Thorat[2]
[1, 2] Department of Computer Science and Engineering
Rajarambapu Institute of Technology, Rajaramnagar
Islampur, Sangli (MS) 415409 India
[1]suyashshpatil@gmail.com, [2]sandip.thorat@ritindia.edu

Abstract— **Grape cultivation has social and economic importance in India. In India, Maharashtra ranks first in grapes production. Over the last few years the quality of grapes has degraded because of many reasons. One of the important causes is diseases on grapes. To prevent diseases farmers spray huge amount of pesticides, which result in increasing the cost of production. Also farmers are unable to identify the diseases manually. The diseases are identified only after the infection, but its takes up a lot of time and have adverse effects on vineyard. The proposed work is to develop a monitoring system which will identify the chances of grape diseases in its early stages by using Hidden Markov Model provides alerts via SMS to the farmer and the expert. The system includes temperature, relative humidity, moisture, leaf wetness sensor and Zig-Bee for wireless data transmission.**

Keywords— Wireless Sensor Network, Vineyard, Hidden Markov Model, Zig-Bee

I. INTRODUCTION

In 2014-15 about 2 lakh tons of grapes were exported from India. Indian farmers lack in knowledge about the use of pesticides because of which for last four years European Union rejected Indian grapes due to excess use of pesticides. In [1] [2], it is found that more than 40 % agriculturists utilize an overdose of pesticides.

Contamination of infections can bring about lessening in both quality and amount of grapes. At each stage development of grapes we have seen different diseases. Many of the farmers can make visual inspection for detection and identification of grape diseases, also some of the systems use image processing technique [3]. This requires constant observing by specialists, which is tedious, costly and less precise. Therefore, farmers and experts want some quick some snappy, programmed, less costly yet exact strategy to identify diseases infected on grapes.

A. *Wireless Sensor Networks*

According to Jason Lester Hill, Wireless Sensor Network (WSN) is Sensing + CPU + Radio = Thousands of potential applications [4]. It consists of autonomous devices called as sensors, which monitor physical conditions such as sound, also environmental conditions like temperature, pressure, humidity, etc. In current situations WSN is an important technology for automation in agriculture.

B. Internet of Things (IoT) in Agriculture

IoT is network of Internet enabled an object which uses web services and interact with those objects [5]. In IoT WSN plays an important role. One of the hottest trend in the current agriculture is precision agriculture [6]. In agriculture, we can increase productivity of crop by using IoT. Also IoT is used for crop water management, precision agriculture, integrated pest management control, food production and safety [7]. Due to IoT in agriculture, we can easily collect sensor data, manage cloud services, easy to get recommendations about weather forecast data etc. with a fast speed [8].

C. Hidden Markov Model

A Hidden Markov Model is the statistical model in which the states are hidden from the user. In basic Markov model states are noticeable to spectator, yet in HMM states are undetectable to client, be that as it may, just the watching states are unmistakable [9]. HMM consists of two stochastic processes. The first stochastic process is states and its transition probabilities called as a Markov chain. And the second stochastic process is emission observation which depends on a probability distribution. With the help of HMM we develop our own model which will fulfil all the necessary conditions for accurate classification of diseases.

II. RELATED WORK

In [10], Vittorio Rossi, Francesca Salinari, Stefano Poni, Tito Caffi, Tiziano Bettati delivered a decision support supportive network (DSS) called vite.net. This was predominantly created for administration of vineyards and it is planned for vineyard chiefs. The exhibited DSS has two fundamental parts, I) An incorporated framework for observing ongoing segments like soil, air, bothers, sicknesses, and so forth. II) An electronic apparatus that can dissect this information and give up and coming data to overseeing vineyard supervisors as cautions and choices.

The paper exhibits the all-encompassing methodology towards the concentrating on critical issues in vineyards. Creator conveyed a framework by means of s web entry so it empowers both consistent redesigning by the supplier and adaptable access by the client. This paper gives a full computerization and joining of information gathering. A. Matese, et al. [11] displayed NAV (Advanced Vineyard Network) which is for the most part dealing with expert slave units of remote correspondence. This framework was actualized in Italy to collect and checking ongoing micrometeorological parameters. This framework is sorted out in three parts, for example, equipment usefulness and information securing, vitality utilization and correspondence.

In [12], Nattapong Tongrod et al. are focused on the development of high quality grapes. For that preparation is needed like preparation of soil, cultivation, and irrigation management and harvesting. In current days precision agriculture is fast growing area of smart agriculture. This paper discusses various tools for developing software solutions for improving quality of the vineyard. The image array module monitors the vineyard using a webcam; also the web module is used to generate web pages automatically. Xufeng Ding et al. in [13] shows early cautioning framework taking into account Internet, which gathers soil dampness, nitrogen focus, pH esteem, temperature, air stickiness and CO2 fixation, and sends it to a focal server situated outside the field. The central server analyzes these parameters and sends necessary warning message to related persons.

In [14], A. Meukaewjinda developed a system in Thailand for identification of grape diseases using image processing. This framework is sorted out into principle three stages I) grape leaf shading division, II) grape leaf ailment division, and III) examination and characterization of maladies. Here back propagation neural network is used for recognizing the colour of grape leaves. This system categorizes the image taken as input into three classes: scab disease, rust disease and no disease. The support vector machine is used for this classification. Sai Kirti Pilli [15] has built up an aGROBOT (model) which recognizes and screens plant ailments, supplements insufficiency, controlled watering system and controlled utilization of composts and pesticides. Proposed system identifies diseases using image processing and the robot will spray pesticides accordingly.

III. Methodology

Here we used two methods for the data analysis. The first one is statistical method and second is Hidden Markov Model. These two methods discussed as follows,

A. Statistical Method

In this method the input data is strictly checked with rules given by National Research Centre for Grapes (NRCG) [16] and Vikaspedia [17]. Due to strict rule checking the results are not correctly classified, so needs to some another method for classification. By doing a literature survey [18-22], we propose new system which is based on a concept of Hidden Markov Model.

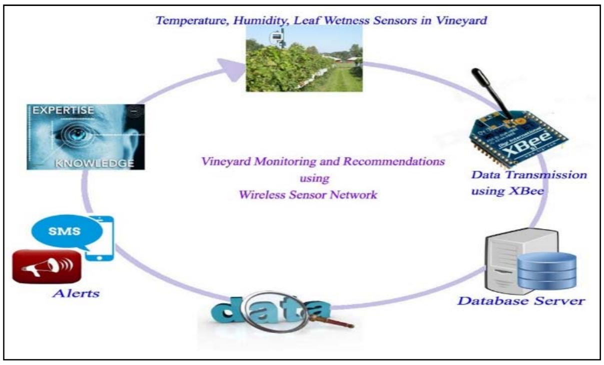

Figure1. Proposed Vineyard Monitoring System Architecture

By using Hidden Markov Model we analyse the input data and classify the labels of diseases. In the next section we elaborate the actual procedure to assign the disease labels. This system is organized into five phases which is shown in Fig. 1.

B. Data Transmission

Temperature, relative humidity and leaf wetness sensors are put in the vineyard. These sensors generate data at continuous intervals. We need to transfer data towards the server for that, using Zig-Bee module. Zig-Bee or IEEE 802.15.4 protocol is a wireless standard designed for the wireless data transmission which requires low energy.

In our design we use Zig-Bee co-ordinator which is responsible for the establishing network and end device which collects the sensor data. All sensors are connected to Arduino board which have inbuilt analog to digital convertor i.e. ADC. Sensors send data which are in analog format, by using ADC that data is converted into digital format and that digital data is collected for transmission. The end device Zig-Bee sends that data to Zig-Bee co-ordinator which is placed outside the vineyard. Now at receiving end Zig-Bee co-ordinator has all sensors data in digital format. By using USB serial adaptor that data is serially transferred to the server.

TABLE I. Favourable Conditions for Growth of Diseases

Disease Name	Temp (°C)	RH (%)	LW (Hours)
Bacterial Leaf Spot	25-30	80-90	-
Powdery Mildew	21-27	More than48	-
Downy Mildew	17-32.5	More than 48	2-3
Anthracnose	24-26	-	12
Bacterial Cancer	25-30	>80	-
Rust	24	75	-

C. Data Analysis

The input data is stored in MySQL database, which is

ID	Temperature	Relative Humidity	Leaf Wet ness	DateTime
4691	28	51	NO	20-11-2015 08:52:00
4692	25	53	NO	20-11-2015 08:52:00
4693	25	57	NO	20-11-2015 08:53:00
4694	25	55	NO	20-11-2015 08:53:00
4695	25	52	NO	20-11-2015 08:54:00
4696	25	51	NO	20-11-2015 08:54:00
4697	25	48	NO	20-11-2015 08:55:00
4698	25	53	NO	20-11-2015 08:55:00
4699	25	53	NO	20-11-2015 08:56:00
4700	25	57	NO	20-11-2015 08:56:00

Table II. Dataset Generated by System

shown in Table II. The next phase is to analyse that data for classification of diseases. The National Research Centre for Grapes (NRCG) gives us the conditions responsible for spreading of diseases. The TABLE II shows name of disease and favourable conditions for growth of disease. For classification of diseases from input data and NRCG data, we proposed Hidden Markov Model which is shown in Fig. 2.

D. Proposed Hidden Markov Model

In our proposed model we use a Baum Welch algorithm which is also called as forward backward algorithm. We have training data of temperature, relative humidity and leaf wetness duration. A proposed model consists of three hidden states such as S1, S2 and S3 and it also has seven observing states like AN (Anthracnose), PM (Powdery Mildew), RU (Rust), DM (Downy Mildew), BLS (Bacterial Leaf Spot), BLC (Bacterial Leaf Cancer) and No disease which are the names of the diseases shown in Fig. 2.

Each state has a certain conditions like, the leaf wetness period is more than 2 hours, or the state change from initial to S1,or when relative humidity is between 48 to 75 then state change to S2 and when relative humidity is more than 75 at that time stage is S3. In HMM current state probability depends on previous state probability.

$$State\ Transition\ Probability\ (a_{ij}) = P\ (S_j(t) \mid S_j(t-1)) \quad (1)$$

If the state changes from initial state to hidden state it is called as transition probability. When the state changes from hidden state to observed state, it's called as emission probability. The probability of observing state depends on its previous states means hidden state, such as hidden state probability. After the hidden stages we have observing stages which are visible to observe.

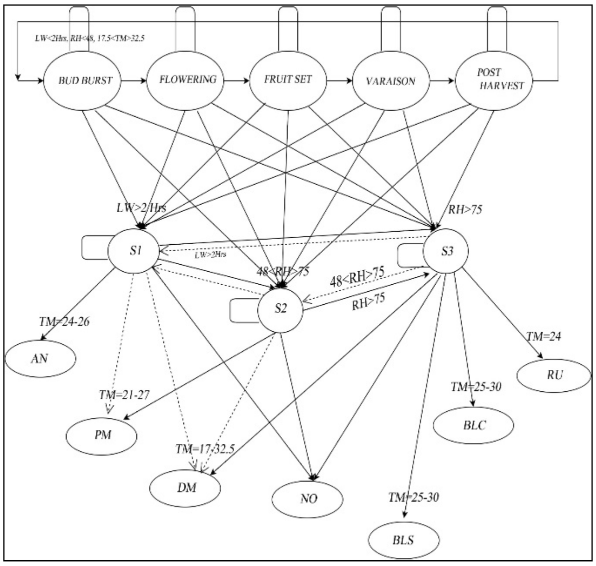

Figure.2. Proposed Hidden Markov Model

The algorithms works as follows,

1. Take input from temperature, humidity and moisture sensors.
2. Convert it into digital format using inbuilt analog to digital convertor placed on Arduino board.
3. Transfer data to server using Zig-Bee wireless communication protocol and store in database.
4. Find Euclidean distance at each point. In our system for calculating probabilities at hiding states we need to first calculate the Euclidean distance by using equations shown below,

$$E.D.\ at\ hidden\ state = \sqrt{(RH - \overline{RH})^2 + (LW - \overline{LW})^2} \quad (2)$$

Where,
RH= Input relative humidity,
LW= Input leaf wetness duration,
$\overline{RH}$= Mean relative humidity at hidden state,
$\overline{LW}$ = Mean leaf wetness duration at hidden state.

At observing state the Euclidean Distance is calculated using the formula,

$$E.D.\ at\ Observing\ state = \sqrt{(TM - \overline{TM})^2} \quad (3)$$

5. Find the state transition probability at each point. After calculating E.D. we need to calculate the probability of each state. Equation (4) calculates probability at hidden and observing state and state changes with highest probability.

$$P_{i,j\,\varepsilon\, all\ possible\ states} = 1 - \min_{0<i<n}\left(\frac{d_{i,j}}{d_{i,0}+d_{i,1}+\cdots d_{i,j}}\right) \quad (4)$$

IV. RESULTS AND DISCUSSION

We had installed this system in Walwa (MS, India) since 2nd November 2015. Since then, we have been taking data check the performance of each method for finding suitable method which will give better and accurate results. These results are shown on a web page and also on text message on mobile.

The analysis is taken from 2nd November 2015 to 15th February 2016, which accounts to a total of 11 weeks. As discussed earlier, we analyse the given training data by two methods. Table III shows the performance analysis based on accuracy. The analysis test is performed on result of both the methods. By analyzing results of both the methods we conclude that performance and accuracy of the HMM is better than the old method.

TABLE III. Performance Analysis based on Accuracy

Parameter	Statistical Method	Hidden Markov Model
Correctly Classified in %	63.63	90.9
Absolute Error	4	1
Relative Error	0.37	0.09
% Error	36.36	9.09
True positive records	2	5
True negative records	5	5
False positive records	0	0
False negative records	3	0
Classification Error Rate	0.3	0
Kappa	0.46	1

The results are checked against the NRGC's Weather forecast based weekly advisory, local newspapers and local agriculture experts. Difference is shown in Table III i.e. it shows the comparative analysis of both the methods. From that we conclude that the performance and accuracy of HMM is higher than statistical method.

A. Hypothesis Testing

In early phase we are doing performance analysis only in terms of accuracy. But accuracy alone is bad measure for classification task. So we take hypothesis testing on the results. Equation (5) calculates the t- test score, Equation (6) z- test score and Equation (7) calculates chi square test score. Also, we measure precision, recall and f-measure for the better performance analysis. In hypothesis testing firstly we define null hypothesis (H_0) and alternate hypothesis (H_a) statements.

H_0= Grape disease detection by using statistical method is more accurate than Hidden Markov Model

H_a = Grape disease detection by using Hidden Markov Model is more accurate than statistical method.

$$T = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{S_1}{n_1}+\frac{S_2}{n_2}}} \quad (5)$$

$$Z = \frac{\mu - \mu_0}{\sqrt{\frac{\mu_0(1-\mu_0)}{n}}} \quad (6)$$

$$X^2 = \Sigma \ \frac{(observed\ value - expected\ value)^2}{expected\ value} \quad (7)$$

TABLE IV. Hypothesis Testing Results

Test	Statistical method	Hidden Markov Model	Confidence	Critical Value
t-test (score)	2.64	1.25	0.95	2.09
z-test (score)	2.64	1.11	0.95	1.96
Chi-square test	4	1	0.95	3.9

By analyzing Table IV we found that in t test and z test scores are above the critical region, so the null hypothesis is rejected and alternate hypothesis i.e. "Grape disease detection by using Hidden Markov Model is more accurate than statistical method" is accepted.

$$P = \frac{TP}{TP+FP} \quad (8)$$

$$R = \frac{TP}{TP+FN} \quad (9)$$

$$\text{F- Measure} = 2*\frac{P*R}{P+R} \quad (10)$$

TABLE V. Precision, Recall and F-Measure

Method	Precision	Recall	F- Measure
Statistical method	1	0.33	0.50
Hidden Markov Model	1	0.83	0.91

Precision (equation (8)) is a measure of result pertinence, while recall (equation (9)) is a measure of what number of really significant results are returned. The high scores for both demonstrate that the classifier is returning exact results (high precision), and also giving back a larger part of all positive results (high recall). F measure (equation (10)) reaches its best value at 1 and worst at 0. According to Table V we identify that the HMM has high precision, high recall and high f_1 score. By observing Table III, IV, and V grape disease detection using HMM is more accurate and more precise than statistical method.

V. CONCLUSIONS

This system is designed for early and accurate detection of diseases by using machine learning and suggests pesticides to protect the crop from those diseases and reduce the manual disease detection efforts. Due to this, it helps farmers to improve the quality of farming and increase the production of grapes. Using this system, farmers get the information about a schedule of fertilizers, pesticides spraying, irrigation, etc. This will be helpful to farmers to increase their profits, and protect the vineyards from affecting the diseases. By accurate identification of diseases and providing correct spraying and irrigation schedule, improvement in the quality and quantity of grapes is achieved along with reduction in excessive use of pesticides. In future we will use more number of sensors to cover large area of vineyard.

References

[1] DEVI P, I.N.D.I.R.A., 2010. Pesticides in Agriculture-A Boon or a Curse? A Case Study of Kerala. Economic and Political Weekly, pp.199-207.

[2] Meng, Yang. "The Damaging Truth About Chinese Fertiliser And Pesticide Use".Chinadialogue.net. N.p., 2012. Web. 1 Apr. 2016.

[3] Moran, M.S., 2000, January. Image-based remote sensing for agricultural management–Perspectives of image providers, research scientists and users. In Proceedings of the Second International Conference on Geospatial Information in Agriculture and Forestry (Vol. 1).

[4] Jason Lester Hill. System Architecture for Wireless Sensor Networks. Dissertation, Doctor of Philosophy, 2003.

[5] Zhao, J.C., Zhang, J.F., Feng, Y. and Guo, J.X., 2010, July. The study and application of the IOT technology in agriculture. In Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on(Vol. 2, pp. 462-465).

[6] Application Of Internet Of Things (Iot) In Agriculture - Case Study. Aditi.com. N.p., 2016. Web. 1 Apr. 2016.

[7] Mahendra, Ashish, Yash Mehta, and Jeevana Mounika. "Agriculture Internet Of Things (Iot) Technology | Applications". Iotworm. N.p., 2015. Web. 1 Apr. 2016.

[8] Huang, Ryan. "Internet Of Things: 5 Applications In Agriculture". HWTrek Blog. N.p., 2014. Web. 1 Apr. 2016.

[9] Sukkar, R., Katz, E., Zhang, Y., Raunig, D. and Wyman, B.T., 2012, August. Disease progression modeling using hidden Markov models. In Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE (pp. 2845-2848). IEEE.

[10] Togami, T., Yamamoto, K., Hashimoto, A., Watanabe, N., Takata, K., Nagai, H. and Kameoka, T., 2011, September. A wireless sensor network in a vineyard for smart viticultural management. In SICE Annual Conference (SICE), 2011 Proceedings of (pp. 2450-2454). IEEE.

[11] Rossi, V., Salinari, F., Poni, S., Caffi, T. and Bettati, T., 2014. Addressing the implementation problem in agricultural decision support systems: the example of vite. net®. Computers and Electronics in Agriculture, 100, pp.88-99.

[12] Matese, A.D.G.S.F., Di Gennaro, S.F., Zaldei, A., Genesio, L. and Vaccari, F.P., 2009. A wireless sensor network for precision viticulture: The NAV system. Computers and electronics in agriculture, 69(1), pp.51-58.

[13] Ding, X., Xiong, G., Hu, B., Xie, L. and Zhou, S., 2013, July. Environment monitoring and early warning system of facility agriculture based on heterogeneous wireless networks. In Service Operations and Logistics, and Informatics (SOLI), 2013 IEEE International Conference on (pp. 307-310). IEEE.

[14] Meunkaewjinda, A., Kumsawat, P., Attakitmongcol, K. and Srikaew, A., 2008, May. Grape leaf disease detection from color imagery using hybrid intelligent system. In Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 2008. ECTI-CON 2008. 5th International Conference on (Vol. 1, pp. 513-516). IEEE.

[15] Pilli, S.K. and George, S.J., 2014, February. eAGROBOT-A robot for early crop disease detection using image processing. In Electronics and Communication Systems (ICECS), 2014 International Conference on (pp. 1-6). IEEE.

[16] P.G. Adasule, Director National Research Center for Grapes, Pune, “Good Agricultural Practices For Production of Quality Table Grapes”.

[17] Grapes: Diseases And Symptoms — Vikaspedia. Vikaspedia.in. N.p., 2016. Web. 17 May 2016.

[18] Eddy, Sean R. "What Is A Hidden Markov Model?". Nat Biotechnol 22.10 (2004): 1315-1316. Web. 14 May 2016.

[19] El-Hanjouri, M., Alkhaldi, W., Hamdy, N. and Alim, O.A., 2002. Heart diseases diagnosis using HMM. In Electrotechnical Conference, 2002. MELECON 2002. 11th Mediterranean (pp. 489-492). IEEE.

[20] Brittain, J.S., Probert-Smith, P. and Aziz, T.Z., 2010, August. Demand driven deep brain stimulation: regimes and autoregressive hidden Markov implementation. In Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE (pp. 158-161). IEEE.

[21] Sukkar, R., Katz, E., Zhang, Y., Raunig, D. and Wyman, B.T., 2012, August. Disease progression modeling using hidden Markov models. In Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE (pp. 2845-2848). IEEE.

[22] Thorat, S.A. and Kulkarni, P.J., 2014, July. Design issues in trust based routing for MANET. In Computing, Communication and Networking Technologies (ICCCNT), 2014 International Conference on (pp. 1-7). IEEE.