

Real-Time Monitoring Model for Early Detection of Crop Diseases

Patrick K. Toroitich (*Author*)
Strathmore University
Faculty of Information Technology
Nairobi, Kenya
patrick.toroitich@strathmore.edu

Dr. Joseph Orero (*Author*)
Strathmore University
Faculty of Information Technology
Nairobi, Kenya
jorero@strathmore.edu

Abstract— The agricultural sector has been a key backbone to Kenya's economy. Agriculture has played a key role in the economy through agricultural farm produce exports and job creation hence improving and maintaining good farming practices is critical in ensuring agricultural yields. Potato (*Solanum tuberosum* L.) is a major food and cash crop in the Kenyan highlands, widely grown by small-scale farmers. However, early detection of potato diseases still remains a challenge for both farmers and agricultural extension officers. Consequently agricultural extension officers who play a critical role in training and creating awareness on sound agricultural practices are few and often lack sufficient knowledge and tools. Current techniques used for determining and detecting of crop diseases have relied upon use human vision systems that try to examine physical and phenotypic characteristics such as leaf and stem color. This technique is indeed important for diagnosis of crop diseases, however the use of this technique is not efficient to support early detection of diseases. This study proposed the use of internet of things technology and machine learning techniques for the prediction of potato late blight disease. Temperature and humidity sensor probes placed on the potato were instrumental in monitoring conditions for potato late blight disease on a farm. These parameters constituted abiotic factors that favor the development and growth of *Phytophthora infestans*. Back propagation neural network model was suitable for the prediction of potato late blight disease. In designing the potato late blight prediction model, historical weather data, potato variety tolerance on late blight disease was used to build an artificial neural network disease prediction model. Incoming data streams from the sensors was used to determine level and risk of blight. This study focused on a moderate susceptible cultivator of potato in developing the model. The algorithm was preferred due to its strengths in adaptive learning. The developed model achieved an accuracy of 94%. (*Abstract*)

Keywords— *machine learning, internet of things, android, data mining, potato farming, crop disease prediction, weather forecast, late blight, predictive analytics, gps, gsm (key words)*

I. INTRODUCTION

Agriculture is the backbone of Kenya's economy and central to the Government of Kenya's development strategy. More than

75% of Kenyans make some part of their living in agriculture, and the sector accounts for more than a fourth of Kenya's gross domestic product (GDP). The sector contributes about 26% of the country's GDP. The sector is a major source of revenue with agricultural produce exports accounting for nearly two thirds of total domestic export. (Republic of Kenya, 2005).

Pests and diseases cause heavy losses through deaths, reduced productivity and loss of markets for products. Crop pests and diseases reduce yields substantially, sometimes by over 50 per cent or even total crop failure. Measures to prevent, control and eradicate diseases and pests in livestock and crops play a major role in improving productivity. In the livestock subsector, notifiable, communicable, zoonotic, transboundary and trade-sensitive diseases are of major economic importance. (Government of Kenya (GOK), 2010)

According to (Ghaiwat & Arora, 2014) plant disease diagnosis is very essential at an earlier stage in order to cure and control them. The authors' further note that human vision systems are mostly used to identify crop diseases. Such techniques are prone to inaccuracy as diagnosis of the diseases are based on the perception and experiences of the farmer or agricultural extension worker. Consequently, (Divya, Manjunath, & Ravindra, 2014) identified that climate and weather conditions are very significant in identifying the actual epidemiology of uprising of diseases or pests.

(Hong, Kalbarczyk, & Iyer, 2016) Identified some of the challenges facing data driven agriculture. They include Crop management decisions and data collection systems need to be designed to meet the needs of specific farms, Automated and user friendly systems need to be developed for users with less software experience, the introduction of expert knowledge must be possible. Systems should allow the inclusion of new automated methods for user defined terms, devices need to be affordable and scalable for large farm deployment.

In addition, agro-input companies and extension agents often lack suitable platforms on which to record farm and crop information that could be beneficial in any self-sustaining

agricultural value chain system. The net negative result of these identified issues lead to misinformation, poor utilization of resources, loss in productivity and poor crop yield while incurring high input costs. (Ousmane, & Collins, 2016).

Early and accurate detection and diagnosis of plant diseases are key factors in plant production and the reduction of both qualitative and quantitative losses in crop yield. Early detection of crop diseases still remains a challenge for farmers in Kenya. This is because majority of the agricultural extension officer lack knowledge in plant disease diagnostic and are prone to prescribe ineffective management options to farmers (Otupa, et al., 2015).

Existing systems for forecasting the disease mostly rely on visual based image processing systems. The limitation of such systems is that they can only be utilized when symptoms appear on a plant, thus such type of systems are unable to assist farmers in treating diseases at an early stage. (Sarika & Sanjeev, 2014).

According to (Soon, Yong, Kyu, Sung, & Eun Woo, 2010), disease forecasts play an important role in determining when to use pesticides. They mention that a weather detection, monitoring and early warning system can provide reliable and timely information to the farmers to deal with weather and climate variability and changes.

Early detection of crop diseases can contribute to better farm productivity. Given micro-weather variables, crop symptoms, a model can be created to inform a farmer on the particular disease and also offer treatment recommendations to stop the spread of the disease. As such for this study, continuous monitoring of the soil, environment variables and weather conditions from sowing to harvest will be required to be able to make informed smart decision.

II. LITERATURE REVIEW

A. Agriculture and Potato Farming in Kenya

Agriculture is the mainstay of Kenya's economy, currently contributing 24 percent of GDP directly, which is valued at Kenyan shillings 342 billion and another 27 percent indirectly, which is valued at Kenyan shillings 385 billion. The sector also accounts for 65 percent of Kenya's total exports and provides more than 18 percent of formal employment. (Government of Kenya, 2009).

Potato farming is one of the cultivated crops in Kenya. It is ranked second to maize in consumption. It is a source of livelihoods for the vast majority of many rural farmers in Kenya. Potato farming in Kenya has been vital in improved national food and nutrition security. It's a source of income generation to actors involved in the potato industry value chain. (Janssens, Wiersema, Goos, & Wiersma, 2013) In their assessment of the value chain for seed and ware potatoes in Kenya, the researcher mentions that approximately 500,000 small scale farmers in Kenya practice potato farming. Approximately 90% of these farmers are said to have less than 1 hectare with a cumulative average of 7.7 tons per hectare from potato farming.

In spite of its importance to the country, the potato industry is plagued by several challenges. Lack of clean seeds, inefficient pest and disease management, inefficient marketing system and a lack of clear packaging policies are some of the challenges that have been identified. (Riungu, 2011).

Management of potato diseases and pests are complicated by lack of reliable clean seed sources as well as heavy reliance on farm saved seed potatoes for planting in the coming season. This accelerates reinfection of fields in the farm.

B. Late Blight Forecasting Models

Various forecasting models have been developed and utilized over the years for predicting late blight of potato across the globe. A summary of the late blight models as reviewed by (Vaibhav, Shailbala, & Pundhir, 2013) are presented below.

BLITECAST is a computerized forecast model for potato late blight developed by Krause and colleagues at Pennsylvania State University. BLITECAST works by using two late blight forecasting techniques i.e. the concept of blight favorable days and the severity values for potato late blight forecasting in Pennsylvania State.

PhytoPRE is a computer based information and decision support system for potato late blight in Switzerland which consists of an epidemiological forecast model, a set of decision rules and an information system.

JHULSACAST is a computerized forecast of potato late blight in Western Uttar Pradesh for rainy and non-rainy year. Weather data included temperature, relative humidity and rainfall on hourly basis.

C. Crop Disease Determination and Prediction Models

(Maina, 2016) Explored use of a vision based model in identifying maize diseases. The author utilized artificial neural network in identifying maize leaf disease by implementing back propagation learning algorithm. The author points out that the algorithm was preferred due to its strengths in adaptive learning, its fast processing speed and the accuracy of its output. The author's work concentrated on examining of phenotypic characteristics to determine the type of disease. The author's model utilizes a smartphone camera to take an image a plant after which pixels are extracted and used as input to determine a particular disease. The researcher recommends use of abiotic stress factors such as pH of the soil, weather should be considered as inputs to the system so as to give a more accurate classification.

(Sandika, Bhushan, Bir, & Mehi, 2014) Proposed a system for severity identification of potato Late Blight disease from crop Images captured under uncontrolled environment. The key contribution of the study was an algorithm to determine the severity of Potato Late Blight disease using image processing techniques and neural network. The algorithm consists of two steps i.e. fuzzy c-mean clustering to separate the disease affected area along with background and a neural network to extract affected leaf area from background. The authors note that the algorithm they utilized achieves an accuracy of 93% for 27 images captured in different light

condition, from different distances and at different orientations along with complex background.

According to (Sannaki, Rajpurohit, Sumira, & Venkatesh, 2013) metrological parameters such as temperature and humidity are important in agricultural systems. The researchers proposed a model to predict weather using a modified k-Nearest Neighbor approach and Feed Forward Neural Network and then utilized parameters such as humidity and temperature to predict the disease outbreaks in grapes. Their proposed system uses historical weather data for forecasting weather data. Their approach consisted of five steps: collection of historical weather data, preprocessing of weather data, building the various disease models, weather forecasting and disease prediction. The figure below illustrates the authors' design of the system. Figure 1 below highlights the authors' model.

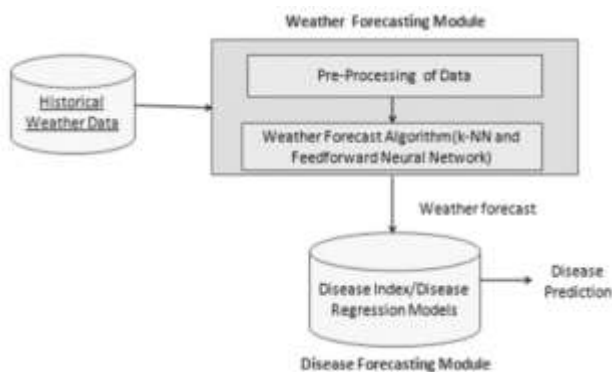


Figure 1: Crop disease prediction system design (Sannaki, Rajpurohit, Sumira, & Venkatesh, 2013)

According to (Milos, Petar, Mladen, & Abdolkarim, 2015) automatic methods for an early detection of plant diseases can be vital for precise fruit protection. The researchers proposed a data mining model for early fruit disease detection. The researchers focused their study on weather variables and microscopic data about registered spores. The authors pointed out that an active pathogenic spore in appropriate weather conditions can lead to fruit tree infection.

D. Internet of Things in Agriculture

The Internet of Things (IoT) technologies can support precision agriculture, a form of agriculture whose goal is to maximize return on investment in agriculture. Irrigation / water detection / soil detection sensors give alerts to help protect a farmer's crop and relay information wirelessly to water reserve points on when to irrigate. Furthermore, farmers can adopt automated drip irrigation in areas where water is scarce. This can be achieved by linking data from various sensors which controls not only where water is released but how much is needed (Dlodlo & Kitwe, 2015).

E. Working Principles of a Soil Hygrometer with Arduino

Arduino is an open-source electronics platform based on easy-to-use hardware and software. Arduino can be used to sense and log information on soil temperature, soil moisture,

soil humidity among other environmental parameters. Taking a soil moisture probe also referred to as a hygrometer as an example, the instrument can be used to detect levels of moisture in soil. The working principle for this probe is that it measures conductivity or resistivity in soil. The sensor consists of two probes that pass current through the soil. The sensor then reads conductance in soil. For a dry soil, the moisture is low hence higher resistance.

F. Rapid Miner

Rapid Miner is a complete business analytics workbench with a strong focus on data mining, text mining, and predictive analytics. It uses a wide variety of descriptive and predictive techniques to give you the insight to make profitable decisions. Rapid Miner together with its analytical server Rapid miner also offers full reporting and dash boarding capabilities and, therefore, a complete business intelligence solution in combination with predictive analytics.

(Ahamed, et al., 2015) Applied data mining techniques using rapid miner platform to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh. The study considered the effects of environmental (weather), abiotic (pH, soil salinity) and area of production as factors towards crop production in Bangladesh. Taking these factors into consideration as datasets for the various districts, they applied clustering techniques to divide regions; and then applied suitable classification techniques to obtain crop yield predictions. The researchers further recommended as part of future works the inclusion of geospatial analysis and also factoring in of time between seeding and harvesting.

G. Mobile Applications Used in Agriculture

Kilimo Salama is an insurance product protecting farmers' investment in farm inputs (seed, fertilizer and chemicals) against extreme weather risk (Drought or excess rainfall) using solar powered weather stations to monitor rainfall and mobile payment technology to collect premiums and payout to farmers. (Helen, et al., 2015).

Wefarm is a free peer-to-peer service that enables farmers to share information via SMS, without the internet and without having to leave their farm. Farmers can ask questions on farming and receive crowd-sourced answers from other farmers around the world in minutes. (Temperton, 2016).

Esoko provides a suite of applications that a network can use to push and pull information to targeted and profiled users. The service started as a piece of software to push market prices out to farmers via SMS alerts. Esoko now targets agribusinesses, smallholder farmers, network operators, NGOs, and ministries. The basic aim is to reduce the cost of communication and improve value chain management for stakeholders in the agricultural sector. The service was officially launched in 2008, and is currently operating in ten countries across East and West Africa. (Gichamba & Lukandu, 2012).

III. SYSTEM DESIGN

The Figure below indicates how the conceptual model for this study is intended to work. The model utilizes weather information from a weather station and disease symptoms to train an artificial agent on the risk of late blight. Farmers will capture planting season data details such as seed variety, location and date of sowing via a smartphone app. An artificial agent will be trained to predict potato disease based on well-known symptoms and provide optimal recommendation of the best fungicide to be applied. Both farmer and extension worker will receive information via the smartphone application.



Figure 2: Conceptual design of the proposed model

A. Arduino Based Soil Probe

The Arduino based soil probe will comprise of an Arduino Uno R3 board fitted with a 3G/GPRS/GSM Shield. A SHT-10 soil probe will then be connected to the set. The soil probe captures both temperature and humidity sensor. The 3G/GPRS/GSM Shield module shield will be used to transmit data to the server. The shield has a slot for a normal subscriber sim card that enables for data transmission via a mobile telephone network.

B. Android client

The android client represents an application that will allow farmers to receive near real time information from the soil probes. The android client is composed of an Android mobile handset with an application that will allow farmers to interact with the system. Alert messages or notifications will be transmitted via an application programming interface to the mobile phone warning farmers on the possibility of a crop disease and possible control mechanism such as type of fungicide application. The android application will also support real time monitoring for farmers who may not be physically residing on the farm. A farmer will use the android application to register their details via a simple registration form and the data will be sent to the server for storage.

Agricultural extension officer will also utilize the developed android application to receive alerts on where a certain crop disease has been flagged by the system. This will allow them to be able to schedule a farm visit to carry out further inspection. A geographic map interface showing the location of the farm where the disease has been flagged is

desired to enable faster and easier locating by the agricultural extension officer. Server

The server will act as a host for both the farm and disease database. The farm database will store farm information received from the farmers. The disease database will store information on various crop diseases and their symptoms. The server will as well host the data mining application that will use data from both the farm database and disease database to determine the likelihood of occurrence of a crop disease. This study proposes to use rapid miner a data mining application that will be installed on the server to be used as part of the crop disease prediction module.

C. Restful Api

A restful API is an application programming interface that adheres to the principles of REST i.e. Representational State Transfer. A restful application use HTTP request to post, read and delete data via CRUD operations i.e. create, read, update and delete. The restful API will be vital in data transfer from the android phone to the server and vice versa.

D. Architecture

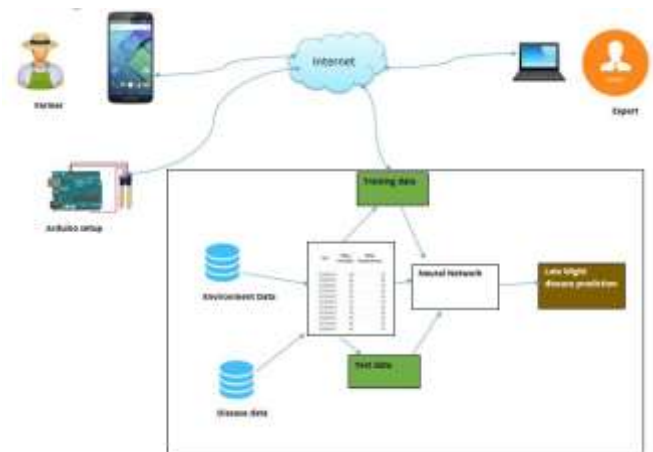


Figure 3: System Architecture of the Proposed Model

IV. IMPLEMENTATION AND TESTING

The complete bill of materials are listed below:

- i. Arduino Uno R3 microprocessor
- ii. DHT11 Temperature and humidity probe
- iii. 3G/GPRS/GSM Shield
- iv. A Cloud based server with rapid miner application
- v. Android Mobile device
- vi. Laptop

The DHT11 sensor probe and gsm shield was fitted to the Arduino uno using jumper pin cables. Temperature, geographical coordinates and humidity readings were transmitted to a cloud server via a restful protocol. Geographical coordinates is essential in determining where late blight incidence is very high and also will allow agricultural extension officers in easily locating farmers.

In designing the potato late blight prediction model historical weather data, potato variety tolerance on late blight disease and disease data was used to build an artificial neural network disease prediction model. Incoming data streams from the sensors will be used to determine level and risk of blight.

Back propagation algorithm was implemented in the forecasting of late blight of potato. The algorithm constituted several elements as documented below.

Input Layer- The input layer consisted of mean daily humidity, mean temperature and daily blight units accumulated. The values that were used in the prediction were based on the scaled values using minmax scaler algorithm.

Hidden Layer - In a neural network, the hidden layer is used to ensure better results for the output are achieved during the prediction process given input values. For the prediction of potato blight disease, two hidden layer was used.

- i. **Initial weight range-** The weights were initialized in a range between -1 and + 1.
- ii. **Number of hidden Layers-** Two hidden layer was used in the prediction of late blight
- iii. **Number of nodes in the hidden layer-** The model was constituted of eight hidden nodes in its hidden layer.
- iv. **Number of Epochs-** An epoch refers to one clean sweep through all records in the training set. Increasing the number of epochs increases the accuracy of the model. 2500 iterations were used in the training thus increasing the accuracy of the model.
- v. **Step size or learning rate for gradient descent-** This is the multiplying factor for error correction during back propagation. Low step size produces slow but steady learning. High value produces rapid but erratic value. The value of the step size ranges from 0.1 to 0.05. A learning rate of 0.05 was used in the model. This ensured steady learning by the network.
- vi. **Hidden layer sigmoid-** The outputs of the hidden node passed through the sigmoid function. The range of the sigmoid function was between 0 and 1.

Output Layer- This was the last layer on the neural network. The output layer for this research was comprised of potato late blight disease presence as high, medium and low.

Figure 4 below represents software flow of the implementation of the prototype.

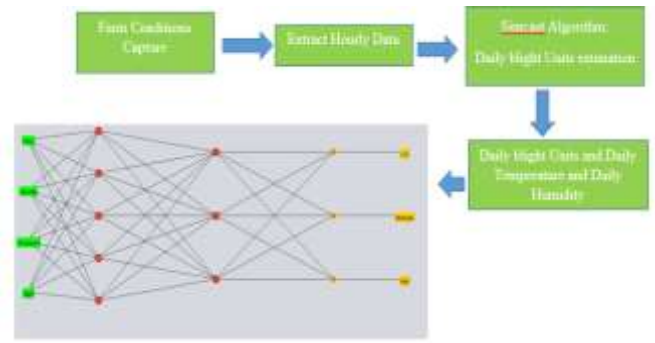


Figure 4: Software Flow diagram

V. RESULTS

This research validated the developed model for accuracy, precision, recall ratio using the confusion matrix. A cross validation of 5 folds was used to test the model. 1367 out of 1461 instances presented to the network were correctly classified. This resulted to an accuracy 93.57% as indicated in table 1 below.

Table 1: Model Classification Output

Correctly Classified Instances	1367	93.5661 %
Incorrectly Classified Instances	94	6.4339 %
Kappa statistic	0.7531	
Mean absolute error	0.0643	
Root mean squared error	0.1807	
Relative absolute error	34.3659 %	
Root relative squared error	9.1763 %	
Total Number of Instances	1461	

VI. CONCLUSION AND FUTURE WORKS

The model presented for the study outlines how information technology can be applied in early detection of crop diseases. The model builds upon vision based techniques in crop disease determination. The model seeks to address weaknesses of vision based model that heavily concentrate on phenotypic characteristics which can only be noticed at later stages of a crops growth. The model for this study seeks to capture environmental variables that promote growth, spread and development of crop diseases.

Internet of things and machine learning techniques can prove vital in coming up with more accurate and practical techniques in disease identification. Utilizing these aspects farmers can well be informed on the status of their farms as well as accorded on farm services on control measures that can curb spread of the crop disease.

The study recommends the following as part of future studies aimed at extending the works of this research.

- i. The researcher recommends use of soil fertility parameters should be considered as part of the model input
- ii. The research recommends the application be extended to support text or short message service

iii. The researcher recommends use of leaf and stem image texture extractions to be part of the model so that an ideal and adequate prediction of late blight can be provided.

iv. Majority of small holder farmers practice mixed crop farming hence the researcher recommends the extension of the model to consider cultural and farm practices that promote disease infection of crops

In conclusion, there is need to use information technology in addressing problems on early detection of crop diseases.

REFERENCES

- [1] Republic of Kenya. (2005). Economic Survey 2005. Nairobi: Government Printer.
- [2] Government of Kenya (GOK). (2010). Agricultural Sector Development Strategy 2010- 2020. Retrieved November 9, 2016, from <http://faolex.fao.org/docs/pdf/ken140935.pdf>
- [3] Ghaiwat, S. N., & Arora, P. (2014). Detection and Classification of Plant Leaf Diseases Using Image processing Techniques: A Review. *International Journal of Recent Advances in Engineering & Technology (IJRAET)*,2(3),2347-2812.
- [4] Divya, M., Manjunath, N. T., & Ravindra, H. S. (2014). A Study on Developing Analytical Model for Groundnut Pest Management using Data Mining Techniques. 2014 Sixth International Conference on Computational Intelligence and Communication Networks (pp. 691-696). IEEE.
- [5] Hong, Z., Kalbarczyk, Z., & Iyer, R. K. (2016). A Data-Driven Approach to Soil Moisture Collection and Prediction. 2016 IEEE International Conference on Smart Computing (SMARTCOMP), (pp. 1-6). St. Louis, MO. doi:10.1109/SMARTCOMP.2016.7501673
- [6] Ousmane, B. M., & Collins, T. J. (2016). Regional Developments; Africa. Chapter 9. 2016 Global Food Policy Report. IFPRI.
- [7] Otupa, M., Karanja, T., Wendot, P., Owiti, E., Ringera, E., Oduor, H., . . . Oronje, M. (2015). Plant Clinics in Kenya: An Innovative Way of Controlling Plant Diseases. Nairobi, Kenya.
- [8] Sarika, D., & Sanjeev, W. (2014, February). Monitoring and Detection of Agricultural Disease using Wireless Sensor Network. *International Journal of Computer Applications*, 87(4). Retrieved November 18, 2016, from <http://research.ijcaonline.org/volume87/number4/pxc3893573.pdf>
- [9] Soon, S. H., Yong, K. H., Kyu, R. K., Sung, K. K., & Eun Woo, P. S. (2010). A Web-based Information System for Plant Disease Forecast Based on Weather. *The Plant Pathology Journal*, 37-48.
- [10] Government of Kenya. (2009). AGRICULTURAL SECTOR DEVELOPMENT STRATEGY (2009-2020). Retrieved from http://www.fao.org/fileadmin/user_upload/drought/docs/Agricultural%20Sector%20Development%20Strategy.pdf
- [11] Janssens, S., Wiersema, S., Goos, H., & Wiersma, W. (2013). *The value chain for seed and ware potatoes in Kenya : Opportunities for development*. Retrieved from www.wageningenUR.nl/en/lei
- [12] Riungu, C. (2011). No easy walk for potatoes. *Horticultural News. The East African Fresh Produce Journal*, 19,16-17.
- [13] Vaibhav, K. S., Shailbala*, & Pundhir, V. (2013). Forecasting Models For Potato Late Blight Management- A Review. *Agricultural Research Communicatio Centre*, 34(2), 87-96. Retrieved from <http://www.arcjournals.com/uploads/articles/R3421.pdf>
- [14] Maina, C. N. (2016). Vision-Based Model for Maize Leaf Disease Identification: A Case Study in Nyeri County.
- [15] Sandika, B., Bhushan, J., Bir, P. S., & Mehi, L. (2014). Severity Identification of Potato Late Blight Disease from Crop Images Captured under Uncontrolled Environment. *Humanitarian Technology Conference - (IHTC), 2014 IEEE Canada International*, (pp. 1-5.). Montreal. doi:10.1109/IHTC.2014.7147519
- [16] Sannaki, S., Rajpurohit, V., Sumira, F., & Venkatesh, s. (2013). A Neural Network approach for Disease Forecasting in Grapes using Weather Parameters. *Computing, Communications and Networking Technologies (ICCCNT),2013 Fourth International Conference on, Tiruchengode*, (pp. 1-5). Retrieved from <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6726613&isnumber=6726462>
- [17] Milos, I., Petar, S., Mladen, V., & Abdolkarim, A. (2015, November 24-26). Data mining model for early fruit diseases. 23rd Telecommunications forum TELFOR 2015.
- [18] Dlodlo, N., & Kitwe, J. K. (2015). The internet of things in agriculture for sustainable rural development. 2015 International Conference on (pp. 13-18). Windhoek, Namibia.: ETNCC 2015.
- [19] Ahamed, A., Mahmood, N. T., Nazmul, H., Mohammad, T. K., Kallal, D., Faridur, R., & Rashedur, M. R. (2015). Applying Data Mining Techniques to Predict Annual Yield of Major Crops and Recommend Planting Different Crops in Different Districts in Bangladesh. 2015 16th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), (pp. 1-6). Takamatsu. doi:10.1109/SNPD.2015.7176185
- [20] Helen, G., James, H., Samantha, G., Rahel, D., Sari, B., Margot, L., . . . Daniel, O. (2015). Scaling up index insurance for smallholder farmers:Recent evidence and insights.CCAFS Report No. 14 . CGIAR Research Program on Climate Change, Agriculture and Food Security (CAAFS), Copenhagen. Retrieved from www.ccafs.cgiar.org
- [21] Temperton, J. (2016, January). The peer-to-peer SMS network helping farmers connect. Retrieved from [http://www.wired.co.uk/article/wefarm-sms-farmers-peer-to-peer-network](http://www.wired.co.uk: http://www.wired.co.uk/article/wefarm-sms-farmers-peer-to-peer-network)
- [22] Gichamba, A., & Lukandu, I. A. (2012). A model for designing M-agriculture applications for dairy farming. *The African Journal of Information Systems*, 4(4).