A Model for early detection of potato late blight disease: a case Study in Nakuru County

Patrick Kiplimo Toroitich
Faculty of Information Technology (FIT)
Strathmore University

A Model for Early Detection of Potato Late Blight Disease:

A Case Study in Nakuru County

Toroitich Patrick Kiplimo

Submitted in partial fulfillment of the requirements for the Degree of Master of Science in Information Technology at Strathmore University

Faculty of Information Technology

Strathmore University

June, 2017

Declaration

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the thesis contains no material previously published or written by any other person expect where due reference is made in the thesis itself.

Student Name	Toroitich Patrick Kiplimo
Signature	………………………………………………..
Date	………………………………………………...

Approval

The thesis of Toroitich Patrick Kiplimo was reviewed and approved by the following:

Dr. Joseph Orero (PhD)

Senior Lecturer, Faculty of Information Technology

Strathmore University

Signature	………………………………………………..
Date	………………………………………………...

Dr. Joseph Orero (PhD)

Dean of Faculty of Information Technology

Strathmore University

Professor Ruth Kiraka

Dean, School of Graduate Studies

Strathmore University

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 for the country, widely grown by small-scale farmers in the Kenyan highlands. However, early detection of potato diseases such as potato late blight 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 heavily 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 in supporting early detection of crop diseases.

This study proposed use of sensors and back propagation algorithm for the prediction of potato late blight disease. Temperature and humidity sensor probes placed on the potato farms were instrumental in monitoring conditions for potato late blight disease. These parameters constituted abiotic factors that favor the development and growth of ***Phytophthora infestants***. 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 93.89% while the precision obtained was 0.949. The recall ratio from the neural network was 0.968 and an F-measure of 0.964.

Key words: machine learning, internet of things, android, data mining, potato farming, crop disease prediction, weather forecast, late blight, predictive analytics, gps, gsm

Table of Contents

List of Tables

List of Figures

List of Equations

Abbreviations/Acronyms

AFSIS	Africa soils information service
API	Application Programming Interface
AUPDC	Area under the disease progress curve
CSA	Climate Smart Agriculture
GDP	Gross domestic product
GPRS	General Packet Radio Service
GPS	Global Positioning System
GSM	Global System for Mobile Communications
HTTP	Hypertext Transfer Protocol
IBM	International Business Machine
ICT	Information Communication Technologies
IEEE	The Institute of Electrical and Electronics Engineers
IOT	Internet of Things
MHZ	Megahertz
NGO	Non-governmental organization
NVDI	Normalized vegetation index
WSN	Wireless Sensor Network
PA	Precision Agriculture
RS	Remote Sensing
REST	Representational State Transfer
SPSS	Statistical Package for the Social Sciences
VRT	Variable Rate Fertilizers

Dedication

I dedicate this research to the almighty God for the grace, the strength and the good health. To all the people who believed in me: my parents Norah and William Toroitich, my sisters Cheryl and Laureen Toroitich, brother Timothy Toroitich, cousins Karen Tallam, and all my friends who did not cease to encourage me.

Acknowledgement

I would like to sincerely appreciate the efforts and time of Dr. Josesph Orero for his commitment and time in guiding me in this research. His positive feedback went a long way in ensuring success for this project. I would like also to thank Prof. Ateya, Dr. Omwenga and Dr. Wekesa for the constructive feedback and scrutiny they accorded this research. I would also like to acknowledge the assistance in software engineering from my colleagues' Lastly, I wish to appreciate the efforts and feedback provided by Prof. Adam Sparks on Potato late blight disease and modeling approaches.

Chapter One: Introduction

1.1 Background Study

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 agricultural sector in Kenya accounts for approximately two thirds of total domestic exports. (Republic of Kenya, 2005).

The Kenya National Climate Change Action Plan 2013-2017 identifies that farming in Kenya is primarily small scale. It's estimated that about 75% of the total agricultural outputs produced on farms averaging 0.33 hectares heavily depend on rainfall. (Government of Kenya, Kenya National Climate Change Action Plan 2013 -2017.Nairobi, 2012 a). Small scale farmers practice a variety of farming systems. They include crop-tree systems, crop-livestock systems, rice-fish integrated systems and fish-poultry systems. (Thorpe, Omare, Owango, & Staal, 2000) Identified these strategies being critical in increasing productivity and sustaining incomes of farmers in the wake of increasing human population growth and pressure, particularly in the densely populated areas in Kenya.

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)

Early information on crop health and disease detection can facilitate the control of diseases through proper management strategies such as vector control through pesticide applications, fungicide applications, and disease-specific chemical applications; and can improve productivity.

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).

1.2 Problem statement

Early, 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 farmers and agricultural extension officers lack adequate knowledge in plant disease diagnostic and are prone to prescribe ineffective management options to farmers (Otipa, et al., 2015).

According to (Muthoni, Shimelis, & Melis, 2013), farmers are also faced by poor advisory services from agricultural extension staff as well being in under developed remote regions which makes it difficult for farmers to access market for certified seeds which are vital in reducing occurrences of diseases such as bacterial wilt and late blight. Late blight of potato disease is most damaging in areas with high rainfall and low temperatures which happens to be areas suitable for rain fed potato production in the tropics (Kaguongo, et al., 2008).

Further, (Elliott, 2015) identified that limited access to agricultural advisory services, technical knowledge and market information are some of the chief obstacles which smallholder farmers in sub Saharan countries face.

There is indeed need to use information technology to address this problem of early and accurate detection of crop diseases. (Maina, 2016) Developed a vision based model to classify maize diseases, however, the use of such models are applicable in situations where the crop has fully matured and phenotypic characteristics are visible hence such models may not be feasible for early and accurate determination of crop diseases. This is because the disease will have been at advanced stage and would have spread.

1.3 Aim

The aim of this research is to find out the immediate needs of farmers and agricultural extension staff that will help them in early detection of crop diseases. The study aims to empower farmers with smart intelligent tools for better decision making during planting, growth and harvesting seasons. The proposed study will provide a model that will enhance use of ICT's and more specifically use of internet of things together with artificial intelligence techniques for early detection of crop diseases.

1.4 Research objective

i. To investigate the problems associated with the current methods applied in the prediction of crop diseases.
ii. To review conditions and symptoms for potato late blight disease development
iii. To review the existing models, mobile applications, techniques and architectures designs for disease identification and prediction in crops
iv. To design a model for crop disease prediction using neural network technique
v. To validate the model

1.5 Research questions

i. What problems are associated with the current methods applied in prediction of crop diseases?
ii. What are the symptoms and conditions associated with potato late blight disease?

iii. What are the existing models, mobile applications and architectures used for crop disease prediction?
iv. How will the prediction based model be designed?
v. How will the prediction based model be validated?

1.6 Justification

Crop diseases prediction is vital in enhancing yield. Current methods and techniques for forecasting plant disease rely on visual based image processing models. However, the limitation with such systems is that they can only be utilized when phenotypic symptoms and characteristics emerge, thus such type of systems or models are unable to assist farmers in treating diseases at an early stage. (Sarika & Sanjeev, 2014)

According to (Soon , Yong , Kyu , Sung, & Eun Woo , 2010), forecasting of crop diseases plays an important role in determining when to use pesticides. They mention that monitoring of weather conditions and use of early warning systems can be vital in providing reliable and timely information to farmers. Farmers can be able to be empowered with knowledge on the correct action to undertake with regards to the changes in weather and climate.

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.

1.7 Scope and limitation

The scope of this study is limited to smallholding farmers in Kenya and more specifically on small scale potato farming. The study will lay its focus on use of temperature and humidity for real time monitoring and prediction of the occurrence of late blight of potato disease. These input readings will be used to train artificial agent such as Rapid miner to deduce risk of late blight disease infection and offer advice on how to mitigate and control the situation. Further the study is proposed to be carried out in Nakuru County. Correlations between farm environment variables and potato disease symptoms will be modeled to assist farmers on early detection of disease. The parameters of interest for this study will comprise of potato cultivator resistance factor, humidity and temperature conditions which are key in inoculation of ***Phytophthora infestants***.

Chapter Two: Literature Review

2.1 Introduction

This chapter will focus on the contribution of agriculture in Kenya, the challenges facing farmers in monitoring crops and mitigating spread of diseases. The chapter will lay a focus in discussing in detail related literature on crop disease prediction techniques, smart farming and precision agriculture techniques. Diseases affecting potato farmers will also be elaborated in detail. The chapter will provide more details on mobile applications used by farmers. This chapter will illustrate machine learning techniques, precision agriculture models that can be used to monitor crop health and agricultural productivity. A conceptual model will be presented to summarize the proposed working elements of the solution.

2.2 Agriculture 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)

2.2.1 Climate Smart Agriculture in Kenya

According to (Lipper, et al., 2014), climate smart agriculture is defined as an approach and process of changing and reorienting agricultural development under the current realities of climate change. Climate smart agriculture is firmly anchored on three key pillars, i.e. increasing agricultural productivity, adopting and building resilience of farming systems to climate change and reducing of greenhouse gas emissions. A key practice in climate smart agriculture is soil management. Maintaining or improving soil health is essential for sustainable and productive agriculture.

2.2.2 Agricultural Extension and Challenges Related to Potato Farming in Kenya

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.

2.2.3 Potato Production Regions in Kenya

Potato cultivation in Kenya is practiced in the highland regions where temperatures ranges from 15 – 24 degree Celsius. There are thirteen major potato producing counties. They include: Bomet, Bungoma, Elgeiyo-Marakwet, Kiambu, Meru, Nakuru, Narok, Nyandarua, Nyeri, Taita-Taveta, Trans-Nzoia, Uasin-Gishu and West Pokot.

2.2.4 Potato Varieties in Kenya and Tolerance to Late Blight

Table 2.1 below illustrates the various potato cultivator varieties grown in Kenya and their tolerance to late blight disease. The table is based on information retrieved from the National Potato Council of Kenya

Table 2.1: Potato Cultivator Varieties in Kenya

Potato Variety	**Tolerance To Blight**
Tigoni	Tolerant
Asante	Fairly Tolerant to blight
Kenya Mavuno	Tolerant to late blight and other potato diseases
Purple Gold	Moderately resistant to late blight,

2.3 Diseases Affecting Potato Farming

Potato farming like any other agricultural crops are affected by several diseases pests. Late blight, early blight, bacterial wilt, soft rot and potato viruses have been identified as some of the common diseases. The table below summarizes important diseases and their control as per the crops extension pocket handbook provided by the Kenyan Ministry of Agriculture. Table 2.2 below highlights the different diseases affecting potatoes.

Table 2.2: Potato diseases and their control

Disease	Symptom	Control
Late blight	Water soaked spots on leaflets or stems which later turn brown and black on underside of leaves, white mold growth at edge of spots or underside of leaves	Resistant varieties, preventive fungicides e.g. Sancozeb,Greemzeb at emergence followed by a single systematic fungicide e.g. Ridomil
Early blight	Dark brown spots on oldest leaves which enlarge when wet.	Resistant varieties, spray with fungicide e.g. Ridomil, Antracol
Bacterial wilt	Wilting without yellowing even when soil has enough moisture, rotting of tuber-white bacteria mass oozes when tuber is cut and squeezed.	3 year crop rotation cycle, use certified seed, field sanitization, avoid use if diseased plants in your compost heap.
Soft rot	Infected tubers break down partially or completely producing a clear slimy foul smelling liquid	Harvest when skin of tuber are hardened, avoid cuts, wounds and bruises on tuber at harvest, harvest n dry weather to promote rapid drying and healing of wounds, and avoid heat damage

Source:Republic of Kenya, Ministry of Agriculture, Crops extension pocket handbook Vol 1, revised edition 2013

In summary for any plant disease to occur three factors must be present and conducive for the plant i.e. there must be a susceptible host (e.g. potato plant), a conducive environment and a pathogen must be present. These three elements, pathogen, host, and environmental conditions, make up the disease triangle. (Shahbaz & Safdar, 2016) Figure 2.1 illustrates a potato disease triangle.

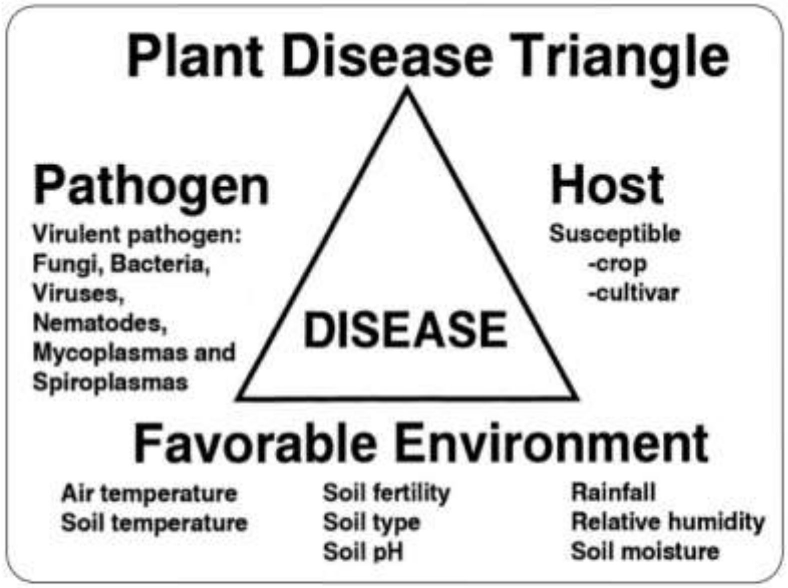


Figure 2.0.1: Plant Disease Triangle (Shahbaz & Safdar, 2016)

2.3.1 Late Blight Disease Development Cycle

Sporangia, coming from cull piles, volunteer potato plants or infected seed, are carried by wind and rain to leaf surfaces. When the moisture and temperature conditions are right, infection occurs. Three to seven days later, the first symptoms appear. The mycelium of the fungus invades plant cells and kills them, causing blight. It continues to grow and eventually emerges from the underside of the leaves. The fungus then grow sporangiophores which release more sporangia to be carried elsewhere in the field by wind and rain. Tubers become infected when the fungus is washed off the leaves and the sporangia and zoospores are carried down into the soil with the water. Exposed tubers are infected more rapidly. Uninfected tubers can be infected during harvesting if they come in contact with sporangia in the soil as they are dug. Blight pathogen overwinters on tubers, not in the soil so cull piles, volunteer plants and diseased seed can be a source of infection the following year. Blight must have living plant tissue on which to live. Figure 2.2 below summarizes the late blight disease cycle in potatoes

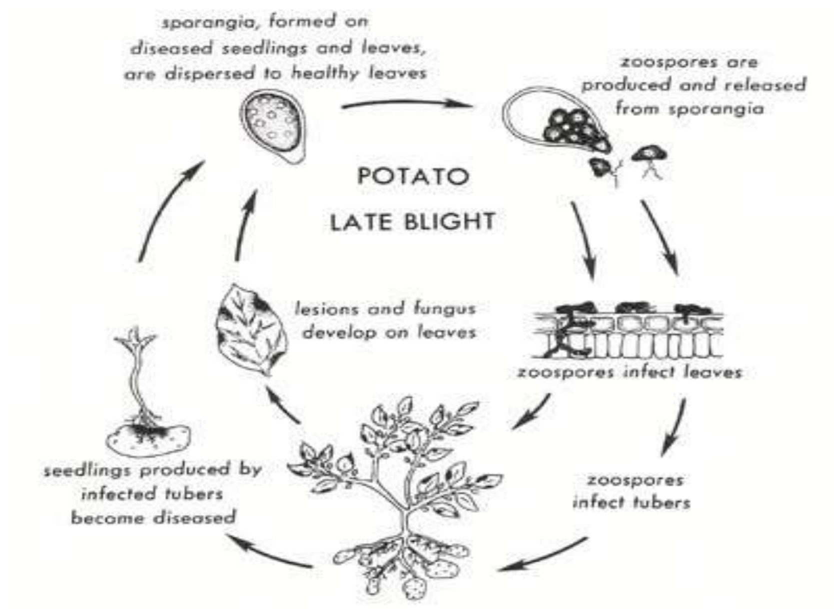


Figure 2.0.2: Late blight disease cycle *(Apple & Fry, 1983)*

2.3.2 Favorable Environment Conditions for Late Blight

Late blight pathogens develops rapidly under moderate temperatures of between 15-24^{o}C, and minimum night temperatures of > 10^{o}C. Free moisture must be present on the plant in order for the sporangia to germinate and infect a new plant. Relative humidity needs to be greater than 90% for sporulation (sporangia) to develop. For infection to take place, cool and cloudy days are required to keep evapotranspiration low coupled with frequent rainfall. (Namanda, et al., 2004)

2.3.3 Monitoring Late Blight of Potato Disease

Late blight of potato disease affects both the leaves, stem and tubers of the plant. Initial symptoms of the disease occurs during the early stages of growth and appear on the leaves as pale green water soaked spots. The spots form around the margins and at tip of the leaves. In humid or damp weather conditions, these spots may sporadically appear on any parts of the leaves. In addition, on the lower side of the leaves whitish ring pattern of sporangia can be visible at close range. On the stem, brown lesions form as illustrated in figure 2.3 below. Stem blight infection is more severe under high temperatures and humid conditions.

Figure 2.0.3: Figure illustrating Blight infection on Leaves and stem (Arora , Sharma , & Singh , 2014)

2.3.4 Organic Fertilizer Effect on Late Blight

(Djeugap et al., 2014) In their study, carried out an experiment to find out the effect of organic amendments and fungicide application on potato late blight, bacterial wilt and yield. The results of the study indicated that potato tuber yield was inversely proportional to late blight severity and bacterial wilt incidence. The study revealed that sub plots that were fertilized with manure recorded high severity values. The study revealed that a possible reason for these high values was directly attributed to high nitrogen content of the manure used. High nitrogen content in the soil has been found to be an accelerator and development of new plant tissues. In addition, the high content of nitrogen creates a favorable microenvironment for disease development. (Talla, Fon, & Fontem , 2011).The formula below used to calculate the standardized area under disease progress curve (SAUDPC).

$$\mathrm{SAUDPC} = \sum_{i-1}^{n-1} \frac{(yi+yi+1)(ti+1-1)}{2(tn-t1)} \qquad \text{Eq. (2.1)}$$

Equation 2.1: Area under disease progress curve

where y1 is the severity at time t1 in days after planting and tn– t1, the duration of the epidemic in days. (Campbell & Madden, 1990). A practical use for this formula is during farm scouting by a farmer or extension officers searching for diseased leaves of a particular crop. The formula is essential for monitoring disease and how it spreads during a planting season. Area under disease progress curve can be used to determine the amount of disease stress faced by a crop.

2.3.5 Review of Models used in Potato Late Blight

According to (Vaibhav , Shailbala*, & Pundhir, 2013) , there are a number of models used in prediction of potato late blight disease. Table 2.2 below highlights a summary of the models reviewed by the authors.

Table 2.3: Summary of Late blight of potato Forecast models

Model Name	**Principles**
Blitecast	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	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:	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.

Similarly, (Fry, Apple, & Bruhn, 1983) developed simcast potato late blight disease model. The model has some unique functionalities. The model gives a recommendation on spraying after reaching a certain threshold. Thresholds are measured as either accumulated blight units or fungicide units. Blight units is an indication of disease severity or diseases pressure on a potato plant. Blight unit accumulation is based on the daily average temperature and hours of relative humidity. Scientist use humidity as a proxy to leaf wetness.

2.3.6 Importance of Late Blight Model

The magnitude of losses of potato caused by late blight in African countries can range from 30 to 75% on susceptible varieties (Olanya, Ojiambo, Ewell, & Hakiza, 2001).According to (Nyankanga, Wien, Olanya, & Ojiambo, 2004) farmers lose up to 30% of potatoes due to late blight, with Meru Central farmers encountering higher loses. The authors estimated that 98% of potato farmers in Meru Central, Mount Elgon and Njabini division, Kenya rely on fungicides to protect their potato crop against late blight, with an average of 5 sprays per season. The most used fungicides as mentioned in the study are Ridomil, Metalaxyl, Mancozeb and different brands of Mancozeb, of which Dithane M45 is mostly used.

2.3.7 Abiotic and Biotic Factors Influencing Potato Plant Growth

Abiotic factors can be defined as non-living physical and chemical properties of an ecosystem that can affect living organisms while biotic factors refer to all living organisms. Examples of biotic factors include plants, animal and fungi. Examples of abiotic factors include rain, wind, temperature soil ph and altitude

2.4 Precision Farming

(Adams, Cook, & Corner, 2000) Define precision agriculture (PA) as an information based and technology driven agricultural system, used to improve agricultural productivity by precisely capturing important parameters and adjusting them adequately. Sowing parameters, modulation of fertilizer doses and site specific application of water, herbicide and pesticides are adjusted to enhance agricultural productivity.

Hyperspectral remote sensing as one of the many techniques of precision agriculture sees beyond the natural limitations of human sight, hence allowing for early detection of crop disease during a crop growth cycle. Early detection enables quick and targeted responses. Hyperspectral remote sensing is less costly than human scouting and frees human labor to address issues rather than spending time looking for them. (AK,Mahlein, 2016)

The Internet of Things, a disruptive technology is predicted to change almost every aspect of our lives. The internet of things gives a ready platform and infrastructure to support precision agriculture by farmers in Kenya. IOT could assist in increasing food production in Kenya whose production is heavily dependent on rainfall. The use of the internet, cloud based services and data sharing platforms are emerging as best avenues that can be used to enhance challenges faced by Kenyan farmers.

IBM Research developed the EZ-Farm, an Internet of Things (IoT) remote monitoring solution that helps small-scale farmers in Kenya to better manage water resources. The project utilizes sensors placed on farms to capture data on water tank levels, soil micro environment variables such as soil moisture and rate of plant photosynthesis. An analytics engine within the Bluemix IBM platform is then used to measure how the plants are growing. It identifies patterns of water usage, then the insights are relayed to the farmer via a mobile app. The intent of the project is to come up with a data centric system that can assist in farm management of resources and monitoring growing conditions.

2.4.1 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)

2.4.2 Remote Sensing

Remote sensing (RS) in agriculture can be described as the art and science of observing and obtaining information on crops and soil characteristics through use of sensors. Near infrared radiation can be used to determine healthy vegetation using a normalized vegetation index technique (NVDI). Remote sensing images and use of geographic information systems can be used to monitor changes in farms and agricultural fields. Wireless soil probes can be used to measure environmental conditions such as soil moisture, soil temperature and soil disturbance rates.

2.4.3 Wireless Sensor Network

(Likhar, Bisen, & Dubey, 2016)Define a wireless sensor network (WSN) as a wireless network consisting of spatially distributed autonomous devices using sensors to monitor physical or environmental conditions. A WSN system incorporates a gateway that provides wireless connectivity back to the wired world and distributed nodes (see Figure 2.4). The wireless protocol you select depends on your application requirements. Some of the available standards include 2.4 GHz radios based on either IEEE 802.15.4 or IEEE 802.11 (Wi-Fi) standards or proprietary radios, which are usually 900 MHz

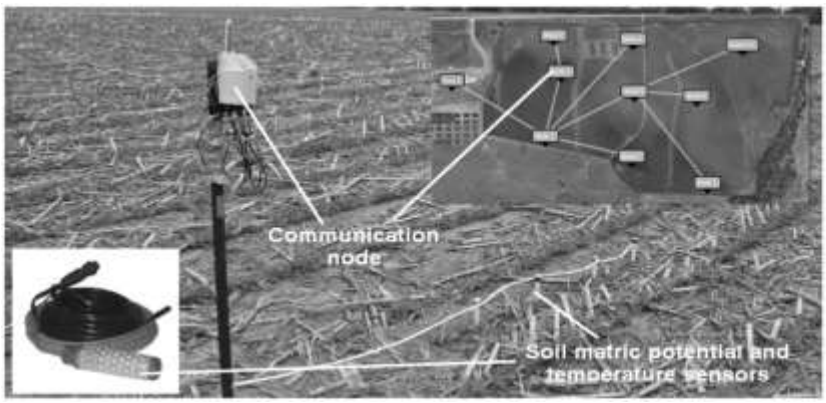


Figure 2.0.4: On farm Wireless Sensor nodes model (Likhar, Bisen, & Dubey, 2016)

2.4.4 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 senor consist 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. Figure 2.5 illustrates a soil hygrometer.

Figure 2.0.5: YL-69 Moisture Sensor (Probes) Source: Guide for Soil Moisture Sensor YL-69 or HL-69 with Arduino, 2016

(Mayur, Mayur, Akshay, & Sachin, 2016) Introduced a concept of smart farming which utilizes wireless sensor technology for soil moisture detection. The authors in the study proposed an automatic plant watering system using Arduino and android. The android application provided information on soil moisture level, type of soil needed, weather forecast, fertilizers, and pesticides to be used. Figure 2.6 below represents a simple hardware representation as stated by the researchers.

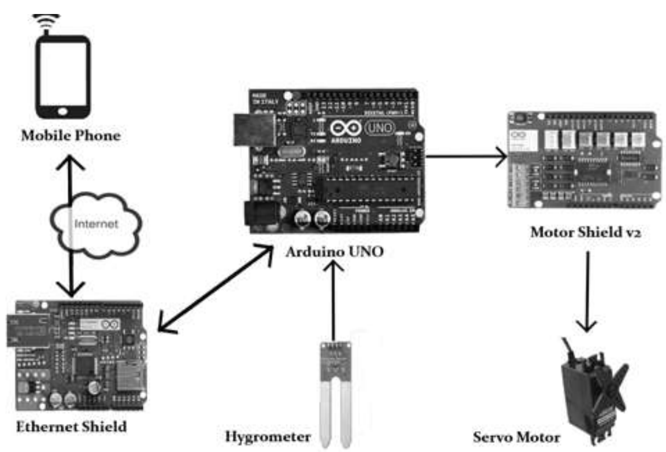


Figure 2.0.6: Hardware Representation (Mayur, Mayur, Akshay, & Sachin, 2016)

2.5 Neural Networks and Machine Learning in Agriculture

Machine learning algorithms have been used in identification, prediction and classification of crop diseases. Machines utilize algorithms to quantify and identify patterns. 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.

Thus one of the aims of this study will emphasize on the full crop life cycle from seeding to harvesting. This study will use environmental data from sensors, time variables from seeding, germination, growth and maturity, biotic properties of the soil. This data will be correlated with common diseases affecting potato growth by small holder farmers. The result of the model will form an advisory knowledge base which can be used to assist farmers in undertaking remedies that will contain the disease and also enhance food production.

2.5.1 Neural Network Architecture

(Koné , 2013) Gives a more understanding to neural network. The author describes a neural network as a system composed of several artificial neurons which have weighted links bounded to them. The artificial neurons are utilized in processing information and are organized into layers which are interlinked with each other. Every neuron in its layer, receives some type of stimuli as input, processes it and sends through its related links an output to neighboring neurons. Artificial neural networks rely on learning algorithms in order to adapt to the environment. A neural network is comprised of four main section:

i. Input which represents a node that receives signals

ii. Interconnections among two or many nodes

iii. An activation function or a transfer function

iv. An optional function whose purpose is used to manage weights of input to output pairs

Figure 2.7 below represents an architecture structure of an artificial neural network.

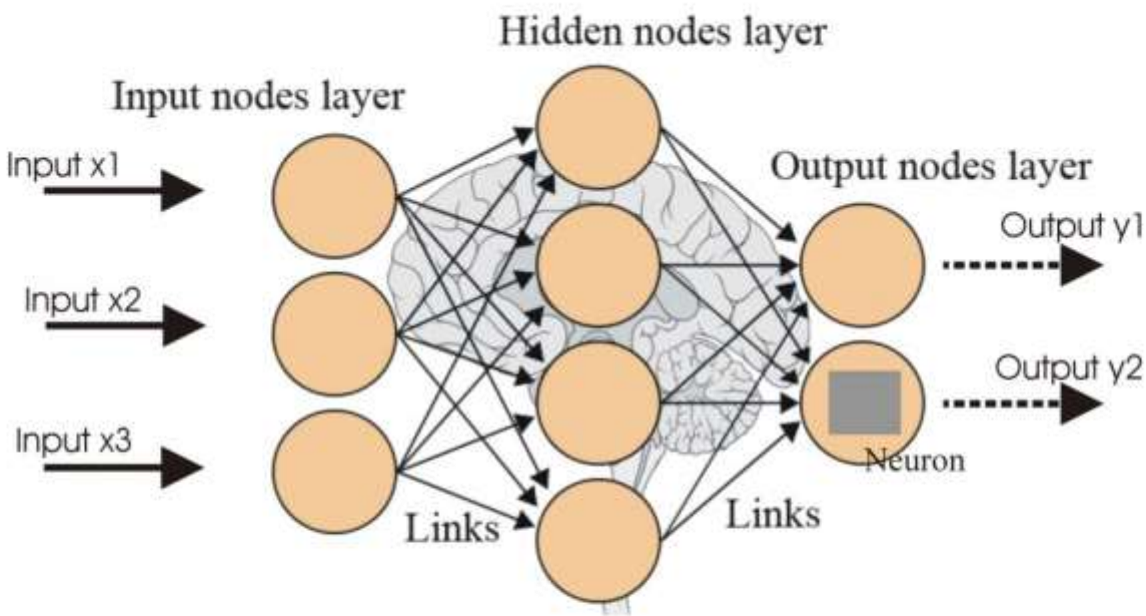


Figure 2.0.7: Artificial Neural Network Architecture

2.5.2 Machine Learning and Image Processing Techniques in Classifying Crop Diseases

Machine learning techniques and image processing techniques form a multidisciplinary approach which can be utilized in the classification of crop diseases through examination of leaf images. Image processing. Image processing relies on the use of computer vision systems in assessing images and quantifying's their characteristics.

(Girish & Priti, 2016) Developed a neural network based detection and classification system for potato leaf diseases. The authors' objective was to develop an automatic and accurate system for disease identification. The system employed image processing techniques to segment diseased regions of the leaf. Consequently, they used a neural network model to classify the queried leaf images as either as healthy or diseased. The results of their study indicated that back propagation algorithm efficiently detected and classified disease spots. The figure below illustrates the working.

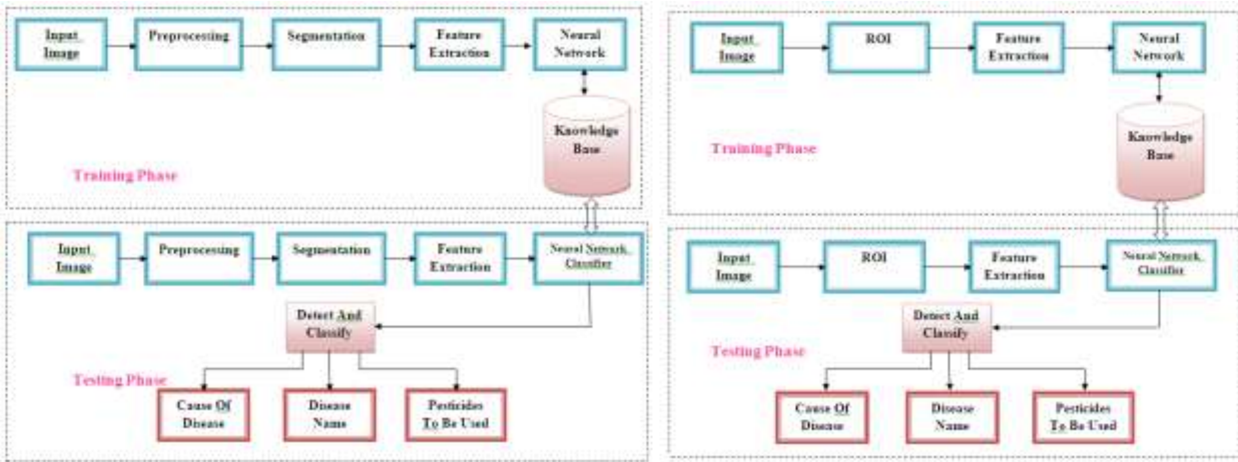


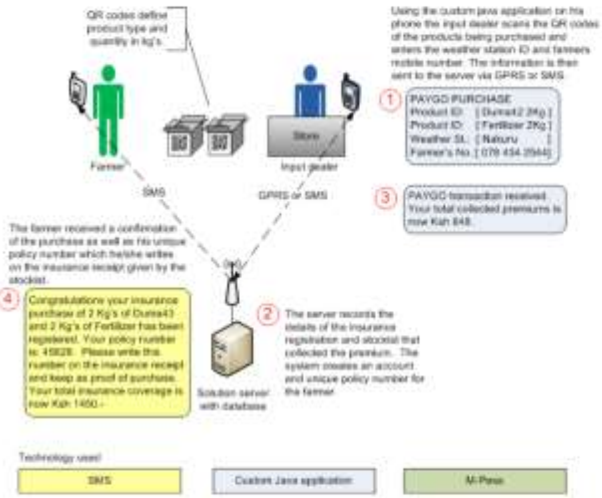
Figure 2.0.8: Potato Leaf Disease identification and Classification Model (Girish & Priti, 2016)

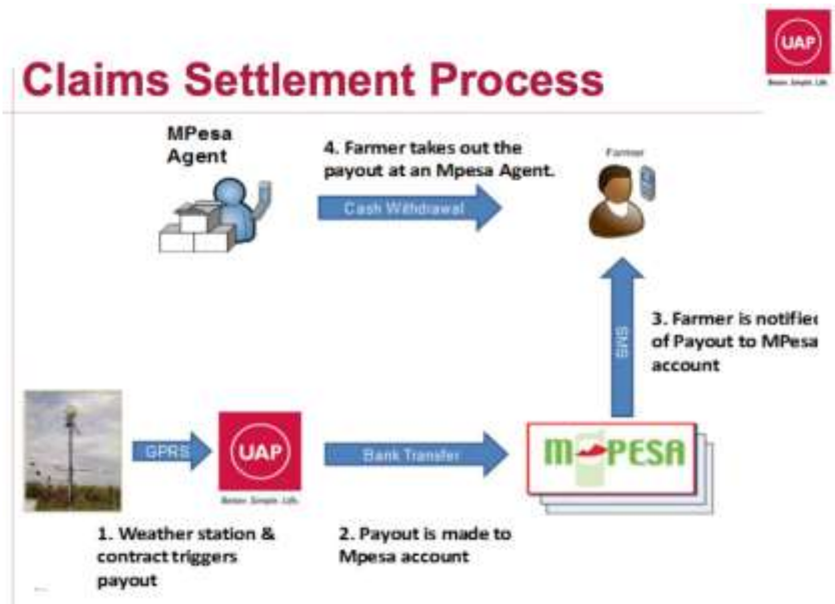
2.6 Mobile Applications Used in Agriculture

This section introduces some of the applications and computing technologies used in agriculture.

2.6.1 Kilimo salama

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)





2.6.2 Wefarm

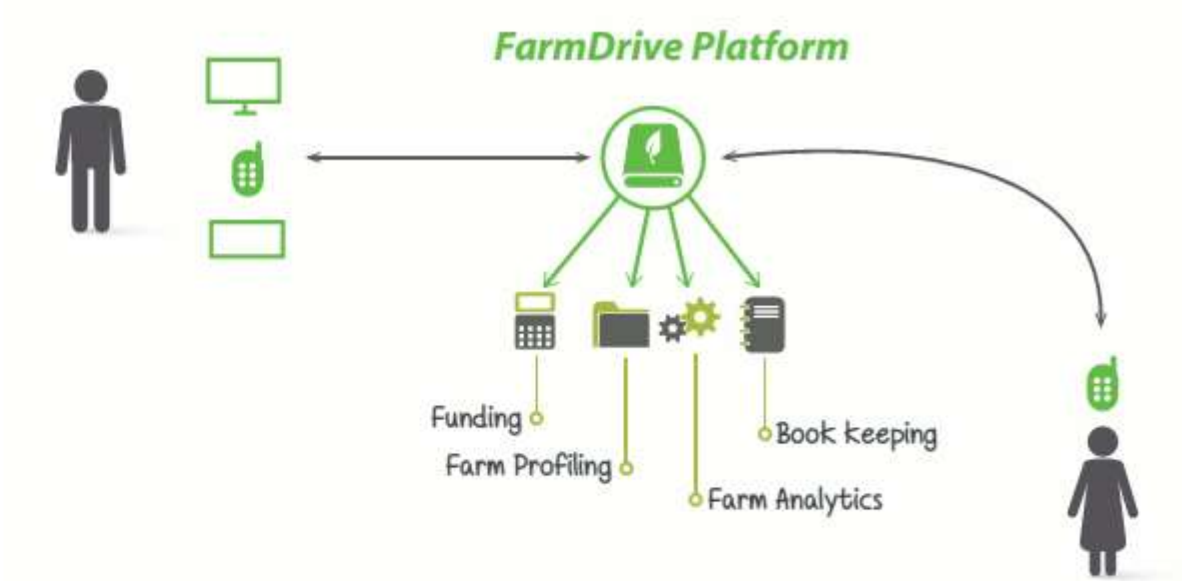
Wefarm is a free peer-to-peer service that enables farmers to share information via SMS, without the internet and without having to leave their farm. Farmers can ask questions on farming and receive crowd-sourced answers from other farmers around the world in minutes. (Temperton, 2016)

2.6.3 Esoko

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)

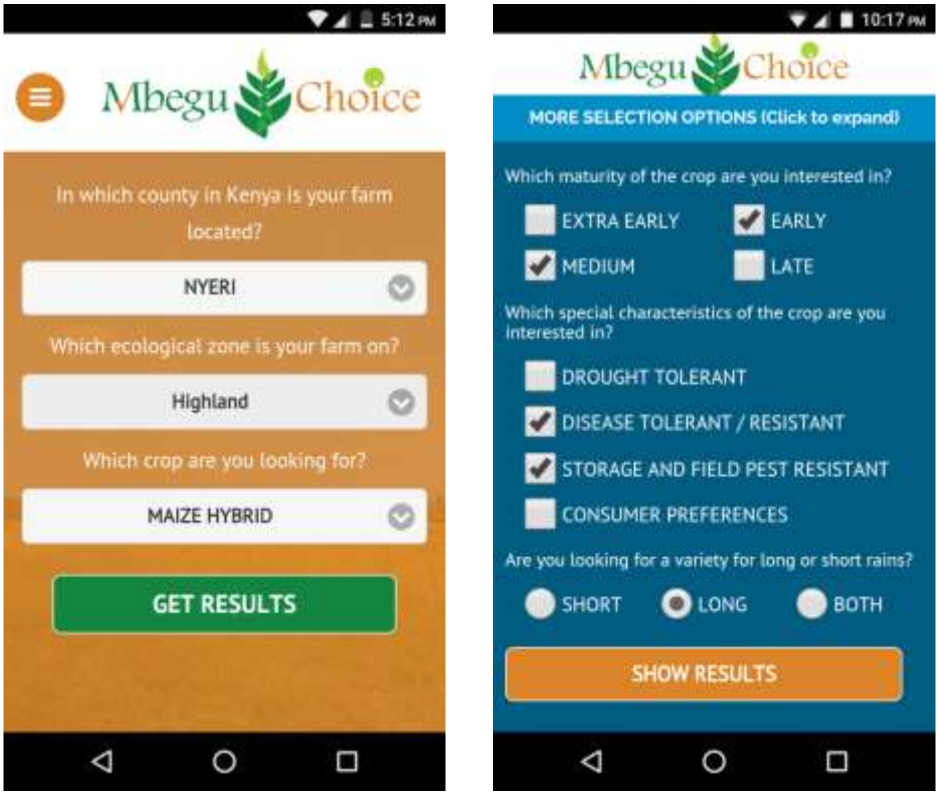
2.6.4 Farmdrive

Farmdrive is a web-based record-keeping application that helps farmers to maintain revenue and expense records. It enables them to access visualized analytical data on farm performance, such as weekly expenses, revenues and an accurate reflection of overall performance. The greatest benefit of Farmdrive is that it bypasses traditional methods of acquiring finance, such as bank loans, and goes directly to individual or group investors for financing farmers. (Mulligan, 2015)



2.6.5 Mbegu choice

Mbegu Choice, the online tool, has mapped seed varieties for all 47 counties and is the first of its kind in Sub Saharan Africa. It allows a smallholder farmer get information on the specific seed variety that would do well in their area while allowing them to choose other attributes that may interest them. These include pest resistance, drought tolerance, and cooking time among others. The online database currently has 237 commercialized crop varieties and is also being used by agro dealers and extension officers. (Arsenault, 2015)



2.6.6 Africa Soil Information Service (AFSIS)

The Africa Soil Information Service project aims to fill a major gap in soil spatial information in Africa. The project has been established to provide spatially referenced soil information to promote agricultural development. The project aims at developing a practical, timely, cost-effective, soil health surveillance service to map soil conditions, providing a foundation for monitoring changes and to provide options for improved soil management. (Shepherd, 2010)

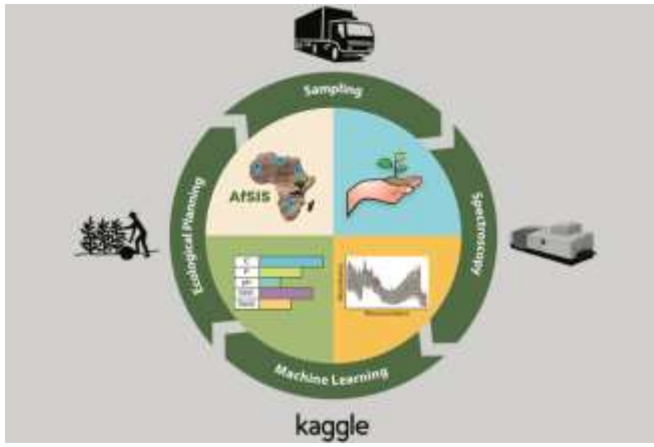


Figure 2.0.9: Africa soil information service model (Shepherd, 2010)

2.6.7 Gaps in Existing Technologies

Innovation in agriculture with a specific context of the Kenyan environment has shown numerous of applications and software being built to support farmers in decision making. However it is important to note that majority of these applications are still on trial phase and some are only focused on market access or produce marketing. Intelligent tools, use of sensors to monitor farms are still lacking in the country. Mobile farm applications used in Kenya such as Farm drive and Mbegu choice do not support use of smart intelligent techniques. They lack in use of machine learning algorithm approaches for decision support.

Taking an instance of Kilimo salama project which aims to protect farmers from adverse climatic condition by offering agricultural insurance to farmers. The project has some major challenges since it lacks capabilities of predicting farmer’s productivity as well as it does not provide for early crop disease recognition services.

The AfSIS project is still in its initial phases and is a much more large scale project and may not address farm specific challenges. There is little literature to support its effectiveness. As such this study will aim to gain more insight on its success and failure through interviews. AFSIS relies on a global soil database which might not reflect current changes at real time on soil properties. Moreover, the soil database are not site specific hence the aim of the study to use soil sensors to gain insight on soil characteristics of different farms.

The difference between the proposed application and the current ones in the market, is that farmers will be able to receive intelligent based information regarding soil, give an indication of likelihood of a certain disease on the farm and recommend control operations. The proposed model will utilize soil sensing devices which will transmit identified key soil characteristics which will be used to train an intelligent machine, there after key decision information can be sent back to the farmers. The model will also offer real time status of their farms and this will go a long way in assisting farmers to be more aware of their farms' ecosystem.

2.7 Crop Disease Determination and Prediction Models

In adding to the body of knowledge, it is constructive to review related literature on disease prediction in crops and agriculture. Some works have been recently carried out in aid to support farmers in crop disease identification and forecasting of crop diseases.

(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. Figure 2.9 below summarizes the researchers' workflow model.

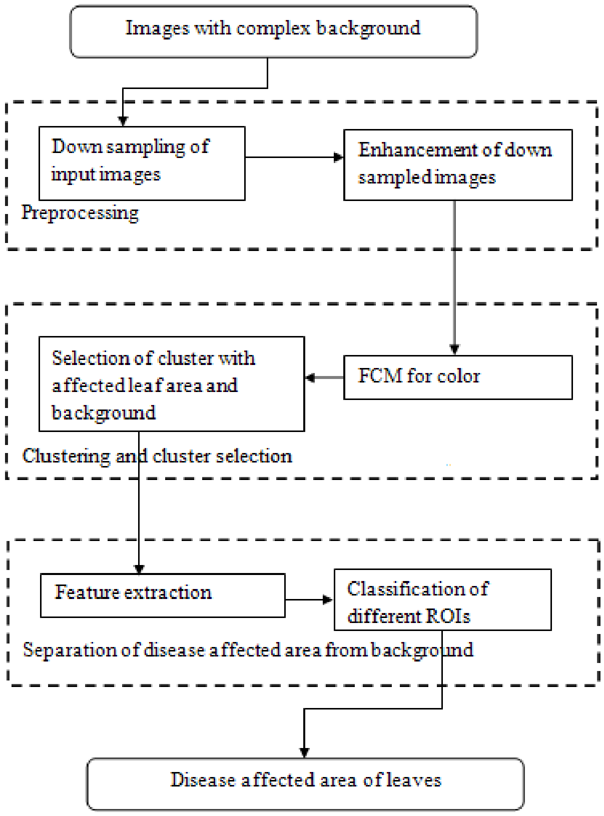


Figure 2.0.10: Overall workflow of the proposed model (Sandika , Bhushan , Bir , & Mehi, 2014)

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. Figure 2.10 below represents the proposed model architecture that was presented in the study.

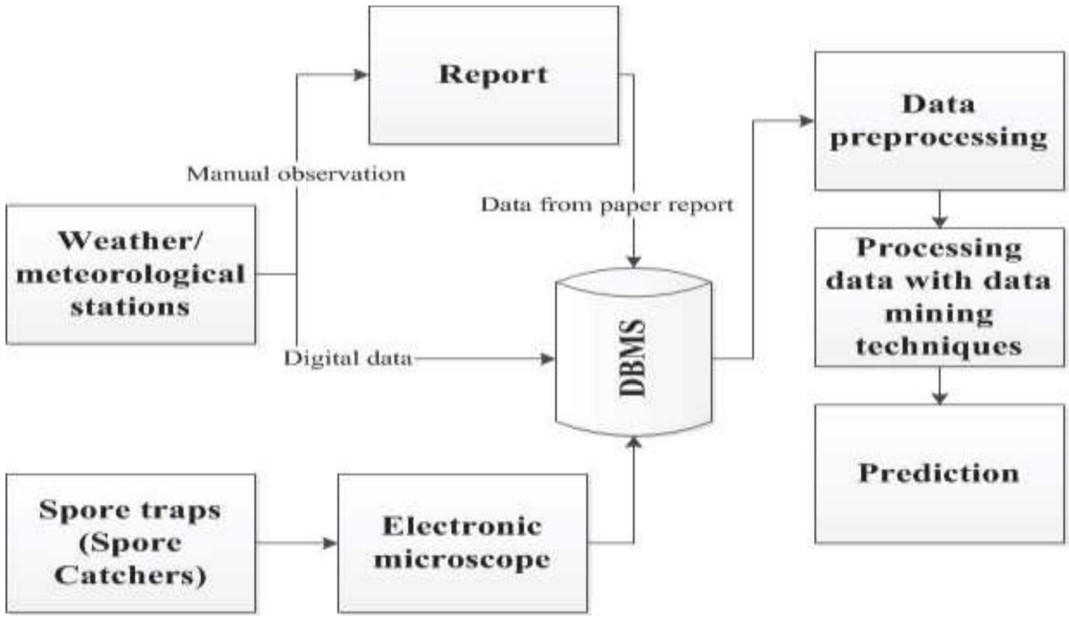


Figure 2.0.11 : Block scheme of the proposed system (Milos, Petar , Mladen , & Abdolkarim, 2015)

(Divya, Manjunath, & Ravindra, 2014) Developed a predictive model to understand the effect of groundnut Thrips under dynamics of crop-weather-pest relations using data mining techniques. The authors worked with micro-level weather data (Temperature, Humidity and Leaf Wetness) which were obtained through wireless mote based AgriSens distributed sensing system and surveillance data were used to understand and quantify hidden correlation between crop-pest/disease weather parameters. The authors indicate that, statistical approach with use of regression mining based correlations assisted in coming up with multivariate regression model that was used to develop an empirical prediction model to issue the forecast for population buildup, initiation and severity of Thrips which will help farmers for crop productivity. Figure 2.11 -below summarizes the process flow model used by the authors.

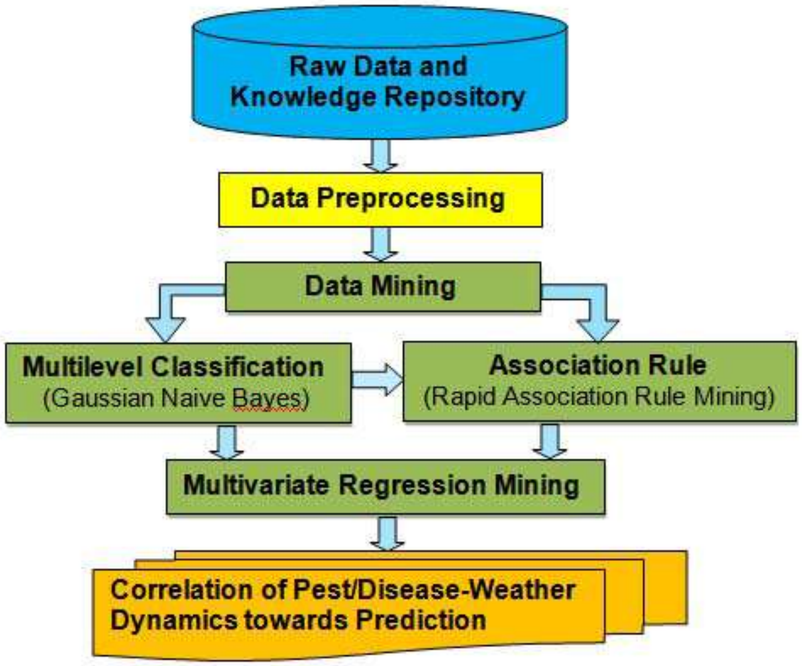


Figure 2.0.12: DM Processing Flow for Pest/Disease Dynamics (Divya, Manjunath, & Ravindra, 2014)

(Hiroyuki, Hiromitsu, & Seiji, 2007) Focused on a prediction model of disease infection for foliar parasite on Welsh onions. The model utilizes temperature and wetness duration to predict the infection of Welsh onions by rust fungus. The authors state that rust fungus disease is the most typical disease on Welsh onions. They further mention that Weibull probability density function is appropriate for approximating the infection rate of the disease.

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. Figure 2.12 below illustrates the authors' design of the system.

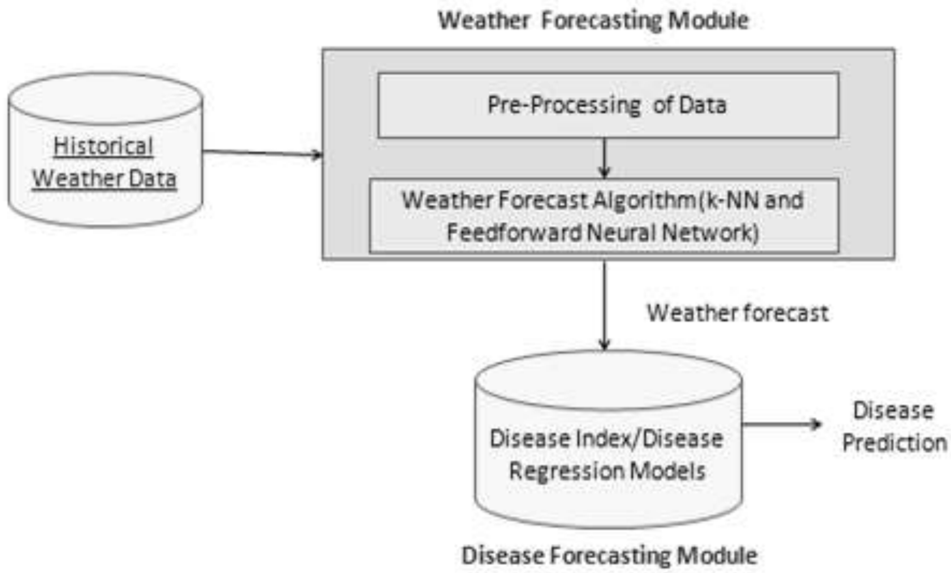


Figure 2.0.13: Crop disease prediction system design (Sannaki, Rajpurohit, Sumira, & Venkatesh, 2013)

(Vidita , Jignesh, & Chetan , 2013) In their study utilized a fuzzy logic approach for plant disease forecasting. The authors developed a weather based plant disease forecasting model using fuzzy logic. The figure below represents the authors' model that they used as part of their study.

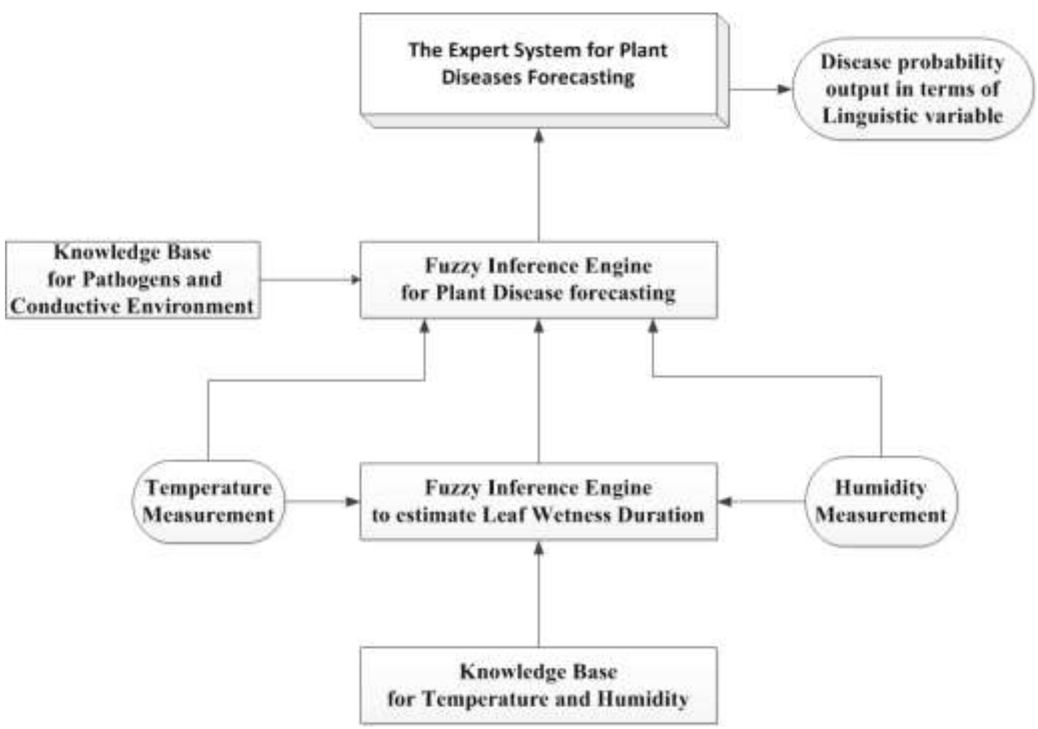


Figure 2.0.14: Conceptual diagram of weather based plant diseases

2.9 Conceptual framework

Figure 2.14 below indicates how the 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.

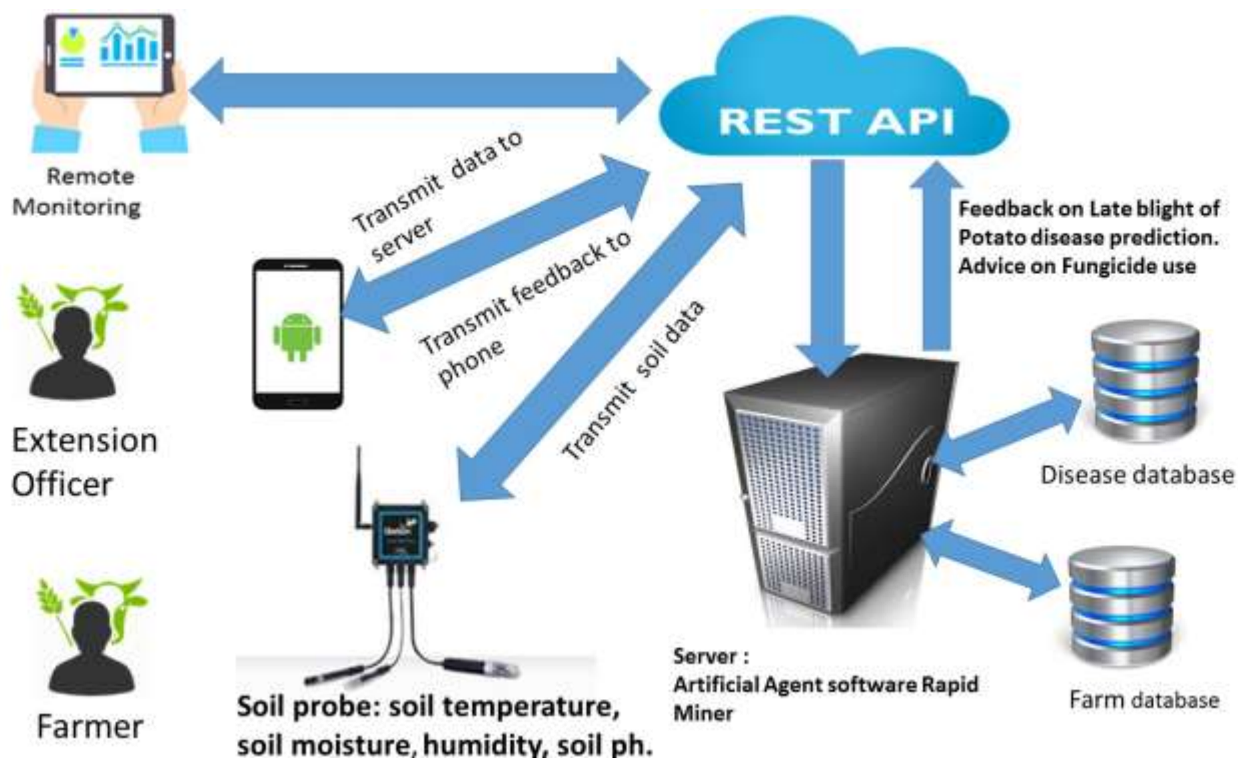


Figure 2.0.15: Conceptual Design of the proposed Model

The sections below elaborate in detail the working principles of the conceptual model.

2.9.1 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 DHT 11 weather prone will then be connected to the set. The 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.

2.9.2 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.

2.9.3 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.

2.9.4 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.

Chapter Three: Research Methodology

3.1 Introduction

This section introduces the research methodology and research approaches used in this study. Research Methodology is defined as the process of systematically solving problems. It can be considered as the science of doing research. (Singh, 2007). This study is anchored on the objectives and key thrusts that aim in solving the problem identified by the research.

This study employed an applied approach taking advantage of internet of things technology, machine learning neural networks and mobile application. Secondary data from both Kenya Metrological department of Kenya and KALRO were used to facilitate this research. This research also utilized expertise and insights from CIP on potato late blight modelling. Fact sheets from KALRO, CIP and Plant wise were utilized. The primary data, hourly humidity and temperature readings that are key factors that influence inoculation of ***Phytophthora infestants*** pathogens were extracted from the metrological data set.

The research also utilized simcast algorithm model to derive blight units and classify them as high, moderate and low. Blight units indicate disease severity stress to the potato plant. The main motivation for the study was to equip farmers with sound decision tools for Late blight of potato forecasting and treatment through use of smart phones, use of ground sensors and machine learning techniques.

3.2 Research Design

Research design is a process used to integrate the studied modules in a logical, understandable manner. Throughout this process the system analyst is able to measure, analyze and systematically arrange the research outcomes.

Applied research design was utilized for this study. Applied research is original investigation undertaken in order to acquire new knowledge. It is, however, directed primarily towards a specific practical aim or objective. Applied research is proposed for this study as it aims find a solution to problems faced by potato farmers. The research attempted to tackle practical questions with regards to early detection of crop diseases problems faced by farmers. Critical insights and findings of the research was used in the design and development of the real time early detection of crop disease model.

A two phased approach research was used for this study. The first phase of the research consisted of qualitative and quantitative studies aimed at investigating the challenges associated with early crop disease detection, investigating the problems associated with the current methods applied in the prediction of potato diseases, to review the existing models, mobile applications, techniques and architectures designs for disease identification and prediction in crops and to establish data and information requirements that will be used by the proposed model. The second phase of the research consisted of the design and development of the potato late blight prediction model and system.

3.2.1 Population and Sampling

Purposeful sampling strategy will be adopted. Purposive or purposeful is a non-probability sampling strategy in which the researcher sampling selects participants who are considered to be typical of the wider population (Singh, 2007).

The population proposed for use for this study will constitute historical for the period 2010-2015 weather data of Nakuru region collected from the Kenya metrological department. Further the study used late blight tolerance level for three potato varieties as published in (Onditi, et al., 2012).In addition, published information on potato blight by the International Center of Potato (CIP) from previous studies was also be used.

3.2.2 Data Acquisition Methods

A non-experimental data collection method was used for this study. Secondary data from the internet and publications will be used in developing the model. The study aimed at using temperature and humidity. These values were used for training, testing and validating the model.

Artificial neural network through back propagation was utilized for the prediction of late blight disease. The advantage of this method is that it is data driven given numerical parameters.

3.2.3 Sample Split

The process involved identifying the data that would be used in training model, the testing data and the validation data used to measure the output error. The sample taken was representative of the identified population. A total of 1461 instances from secondary data collected from the Kenya metrological department. The data was split 50:50 for both testing and training the model.

3.2.4 Model Training

This process involved providing inputs to the model for processing in order to train the model on the type of input data and the expected output of the training session. Training data was fed into the ANN via identified model neurons. A number of iterations will be carried out during the training phase with the aim of reducing the error rate and adjusting of input weights.

3.2.5 Model Testing and Validation

This process entailed the use of test data to cross check whether the system is properly trained by observing the actual model output versus expected output. The use of validation data set assisted in fine tuning the model.

3.2.6 Presentation of Output

Graphical representations such as tables and graphs were used to illustrate the model outputs. Tables was used to present accuracy, precision as well the recall ratio that was obtained during the predictions.

3.3 System Design and Development Methodology

In the development of the application, a rapid prototyping software development approach was used. Rapid application development (RAD) is an object-oriented approach to systems development that includes a method of development as well as software tools. Figure 3.1below illustrates a typical prototyping process.

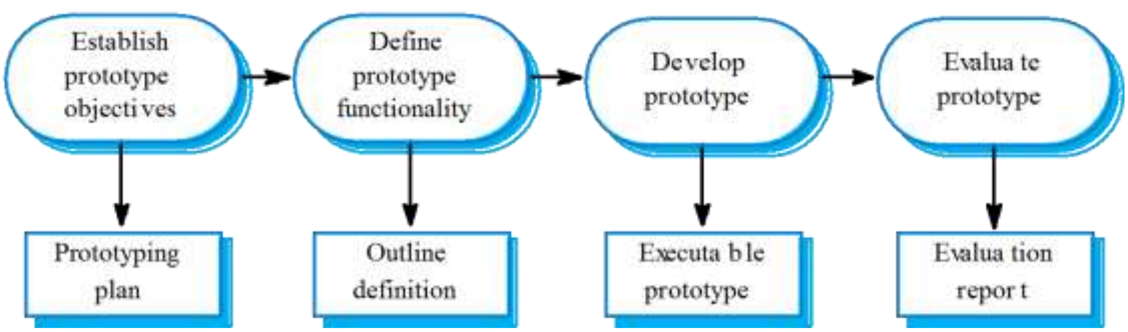


Figure 3.0.1: Rapid Prototyping Life Cycle

Figure 3.1: Rapid Prototyping Life cycle

Rapid prototyping methodology was instrumental for this study and more specifically an evolutionary prototyping model as it enabled the researcher to understand the farmers' requirements at an early stage of development. Further, the methodology was helpful in getting valuable feedback from farmers which was useful for the development of the proposed model. The evolutionary prototyping model aimed at producing a well refined and robust prototype.

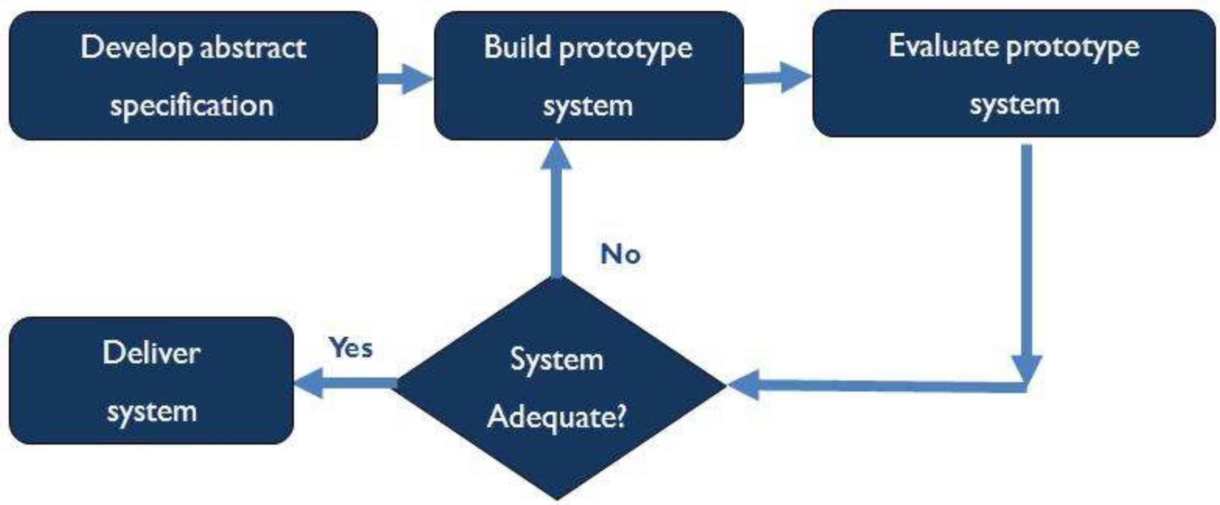


Figure 3.0.2: Evolutionary Prototyping model

The sections below outline in detail the key process above on what will be carried out to achieve the final model for real time early detection of crop disease.

3.3.1 Developing an Abstract Specification

This phase consisted of gathering of basic requirements for the proposed model for early detection of crop disease. The phase addressed at minimum required software and hardware specifications that enabled the successful building of the prototype. Use case diagrams, mock up interfaces were useful at this stage in drawing up basic simple system requirements. Functional and non-functional requirements at this stage were specified to form part of the outline definitions of the working system.

3.3.2 Building a Prototype

This step consisted of creating the real time early crop disease detection prototype. Actual coding and programming of the key components took place during this step. The final output of this stage will have a working system that emulates the intended objectives of this study. A working prototype of the model consisting of temperature and humidity connected to an Arduino board, an android application and web interface that will enable the end users to interact with the system. A fuzzy logic approach will be used to create the prediction module based on data received from the sensors and sent to a centralized server to determine early detection of late blight in potato.

3.3.3 Evaluating Prototype System

An experimentation using participants is proposed to be carried out to assess its viability and performance with regards to giving accurate information for early detection of crop diseases. In particular the prototype will be evaluated to see if it can accurately determine the likelihood of late blight disease of potato. The experiment was aimed at showing how the temperature and humidity probes connected to an Arduino board microcontroller can collect temperature and humidity values. This information together with the site specific geographical coordinates was transmitted to a central server.

3.3.4 Delivery of the System

After successful verification and evaluation of the system, the system will be implemented to support the proof of concept for this study.

3.4 Research Quality

In their study on classification of plant diseases using artificial neural networks, (Muthukannan, Latha, Pon , & Nisha, 2015) applied accuracy, Precision, Recall ratio and F-Measure to carry out their performance evaluations.

3.4.1 Accuracy

The accuracy (AC) is defined as the proportion of the total number of predictions that will be correct. tp represents the true positive, tn represents the true negative, fp represents the false positive and fn represents the false negative in the equations that will be used to measure performance. Equation 3.1 below was used to determine accuracy.

$$Accuracy\ (AC) = \frac{tp+tn}{tp+tn+fp+fn}$$

Eq. (3.1)

Equation 3.1: Accuracy formula

3.4.2 Recall Ratio

The recall or true positive rate (TP) is the proportion of positive cases that will correctly identified as shown in the Equation 3.2 below.

Eq. (3.2)

$$Recall\ ratio = \frac{tp}{tp+fp}$$

Equation 3.2: Recall Ratio

3.4.3 Precision

Precision (P) is the proportion of the predicted positive cases that will be correct, as computed using the Equation 3.3 below.

$$Precision\ (P) = \frac{fp}{fn+fp}$$

Eq. (3.3)

Equation 3.3: Precision

3.4.4 F_Measure

The F-measure computes some average of the information retrieval precision and recall metrics.

3.5 Research Site

For the purpose of this research, Nakuru region will be used as a pilot site for this study. The location is suitable for the study as both large scale and small scale farmers are located in the region. For purposes of this experimental study, small scale farmers will form the target audience for this study. Maize, wheat and potato cultivation is carried out by farmers in the region. The cultivation of these crops usually demands frequent monitoring and more often than not, heavy use of fertilizer and agro-chemicals are used.

3.7 Ethical consideration

In abiding with the principles of ethical research, the researcher will oblige to the standards and rules of conducting a true and just research. The purpose of the study will be explained to all participants during the interview process. This study will reference all materials used for the research The researcher will seek permission before using names and data of individuals during data collection.

Chapter Four: System Design and Architecture

4.1 Introduction

This section illustrates in detail the design and structure of the proposed prototype. The section will focus on key requirements addressed in the previous chapter. A more object oriented approach through unified modelling language is used to describe and map out functional and non-functional requirements of the proposed prototype. Use case diagram, System sequence diagrams and Activity diagrams are used to elaborate the scope and functionality of the system. Context data flow diagram and level Zero diagram will illustrate the key entities and processes interact.

4.2 Requirement Analysis

4.2.1 Functional Requirements

i. The application should allow a farmer to register and login
ii. The application should log humidity and temperature readings from the Arduino SHT10 sensors to the server
iii. The application should be able to transmit humidity and temperature readings to an android client device
iv. The server application should extract data on humidity and temperature in csv format
v. The application should compute risk levels of late blight infection
vi. The application should return a correct prediction based on collected farm conditions
vii. The application should notify extension officers and farmers on late blight occurrence
viii. The application should map farmer location and field of collection points
ix. The application should provide recommendation based on KALRO and International center of potato expertise and knowledge

4.2.2 Usability Requirements

The study identified farmers, late blight potato disease expert and extension agricultural officer as the main users of the system. In designing the system, ease of use by all actors should be adhered to. Navigation and user interfaces should be well labeled to enhance on usability principles.

4.2.3 Supportability Requirements

The study identified an android application and web interface as main communication medium and how users can interact with the system, hence the system should be easily accessible from an android application and web browser.

4.2.4 Reliability Requirements

i. A regular back up of the entire system should be carried periodically
ii. In the event of a failure, the administrator should be able to restore the system to a working state
iii. The android client application should be able to communicate via a restful api.
iv. DHT11 soil sensors should be able to seamlessly interface with the Arduino board and gsm shield module.
v. The application should have the capability of providing warning messages to users

4.3 System Architecture

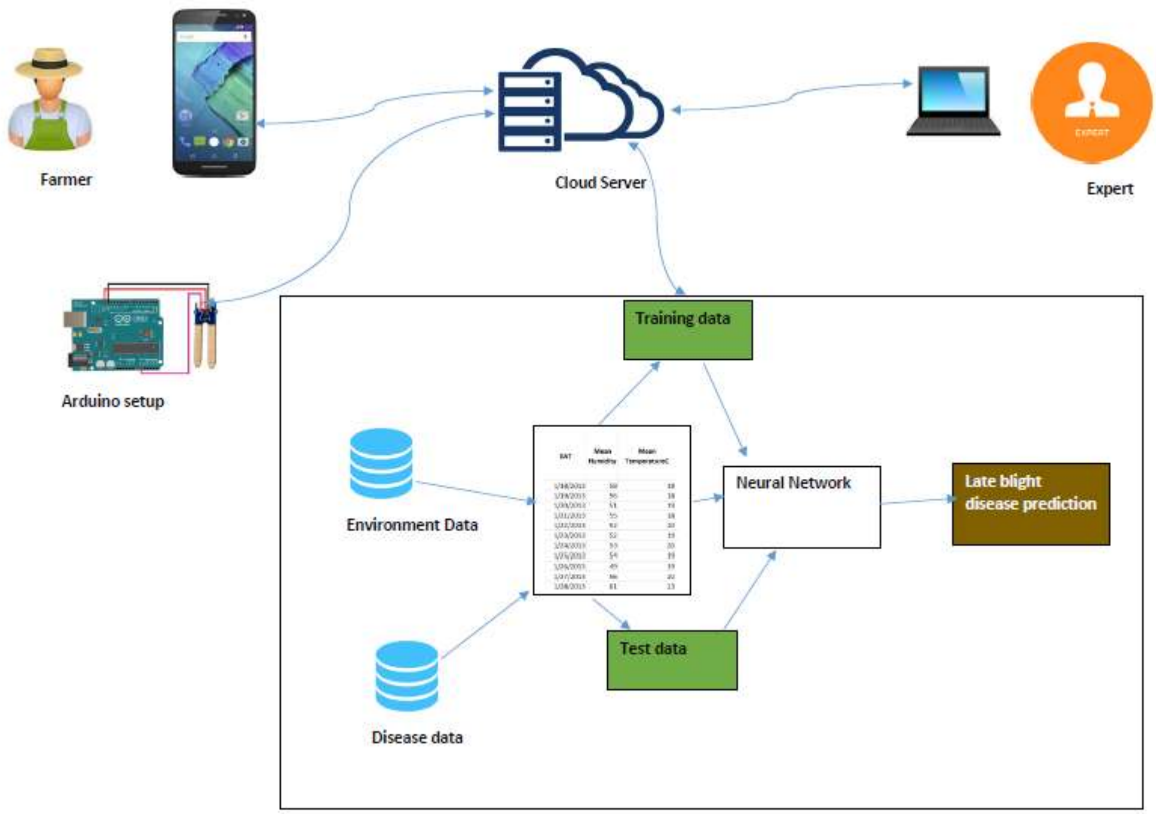


Figure 4.0.1: System Architecture

4.4 Uml Representation of the Model

In software model design and development, models are used to depict user requirements, activities, information structures, components and component interactions of a system. These models provide a framework in software system development to a program code. UML is an industry standard for describing software requirement specifications and design models. (Douglass , 2004)

4.4.1 Use Case Diagram

Use case diagrams are essential in describing how users interact with the information system. (Irwin & Turk, 2005) mention that software developers utilize use case models to primarily capture functional requirements of an information system by focusing on how users interact with the information system and the tasks that users want to accomplish with an information system Figure 4.2 below describes how users interact with the system. The boundary depicts the scope for this study that concentrates on crop disease prediction and more specifically on potato late blight prediction. The diagram elaborates the actors and their associative tasks.

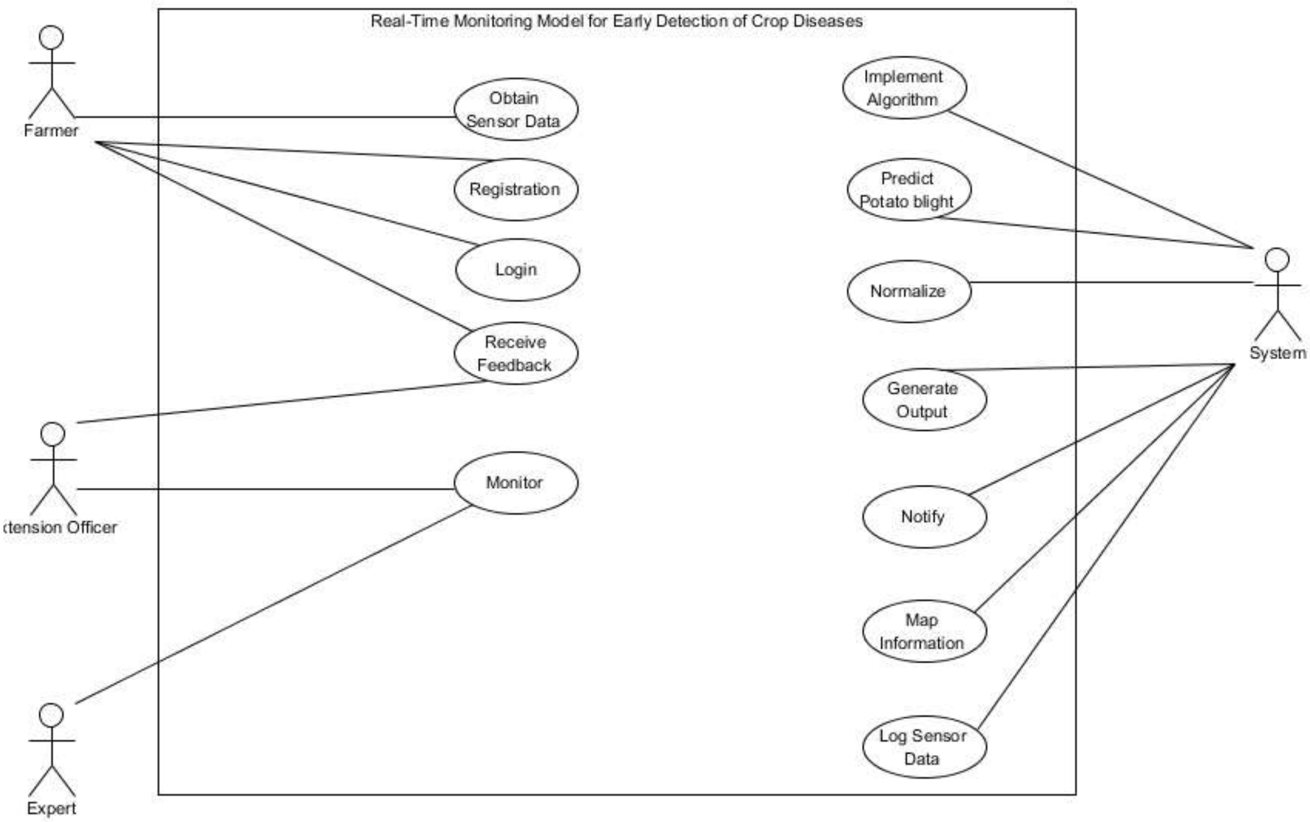


Figure 4.0.2: Use Case Diagram

Use Case Description

Use Case Obtain Sensor Data

Primary Actor

Farmer

Precondition

Arduino Uno is fitted with gsm module with a sim card and Dht11 sensor.

Post Condition

The sensor should be able to transmit readings on temperature and humidity to a cloud server

Main Success Scenario

Actor Intention **System Responsibility**

Actor Intention	System Responsibility
1. Farmer sets up Arduino board with gsm module and sensors	
	2. Log and store sensor data
	3. Relay sensor data to android phone
4.View sensor information	
5.Exit android application	

Extension

At any time the sensor fails to capture temperature readings

Restart the android application

Check if sensor pins have been well linked on Arduino board

Check if remote server is accessible

Check to see if sim card for the gsm module has enough credit

Use Case Description

Use Case Receive Feedback

Primary Actor

Farmer

Precondition

Successful prediction of potato late blight disease by the system.

Post Condition

System to give recommended course of action

If disease risk is high the system to notify extension officer

Main Success Scenario

Actor Intention	**System Responsibility**
1. Farmer requests for potato late blight disease risk information	
	2.Return the output of the prediction
3. View feedback	
	4. Display level of prediction accuracy and time taken

Extension

At any time the request feedback function fails

Retry to request the feedback again

Contact system administrator

Use Case Description

Use Case Description: Log sensor data, Implement Algorithm, Normalize, Predict Potato Blight, Generate Output, Notify

Primary Actor

System

Precondition

Temperature and humidity were accurately logged and stored on the server

Post Condition

Risk of potato blight was accurately calculated and predicted

Main Success Scenario

In determining the risk of late blight and prediction of potato late blight disease.

Log sensor data- The model utilizes sensor information on temperature and humidity to predict the risk of potato late blight disease. Through a restful protocol the system stores this information in a database.

Normalize data- The data use to predict late blight included temperature, humidity, and potato variety resistance score and blight units. In establishing the potato late [illegible] nodel based on a neuron network, meteorological data are normalized by montl Eq. (4.1) e and humidity values are normalized using the below formula.

$$xij = \frac{X0ij - X0jmin}{X0jmax - X0jmin}$$

Equation 4.1: Minmax Scaler

Where Xij refers to the value of either temperature or humidity factor after normalized, between 0 and 1; X0ij refers to the original value of either temperature or humidity; and X0jmin and X0jmax refer to the minimum and maximum values of temperature and humidity factor respectively.

Implement Algorithm- For the purposes of this study, back propagation algorithm was used. Back propagation was identified due to its proven ability in giving accurate classification.

Predict Potato blight- The process was implemented and actualized by providing the system with the training data set from which the system could learn by implementing the algorithm. The system is then provided with a test data set to validate and ensure that the prediction was carried out correctly.

Generate output- The system presented prediction results based on new temperature and humidity readings logged from the sensors.

Use Case Description

Use Case Monitor

Primary Actor

Extension officer, Expert

Precondition

Monitoring interface and server are working optimally

Post Condition

Warnings and alert should be well determined by the system and transmitted to extension officer.

Main Success Scenario

Actor Intention	**System Responsibility**
1. Extension officer enters username and password	
	2.System checks for authentication details
	3. System grants access to the identified user
4.Requests report and monitor interface	
5.Extension officer downloads and print summary report	

Extension

At any time the system fails

Check if you entered valid username and password

Contact system administrator

4.4.2 Domain Model

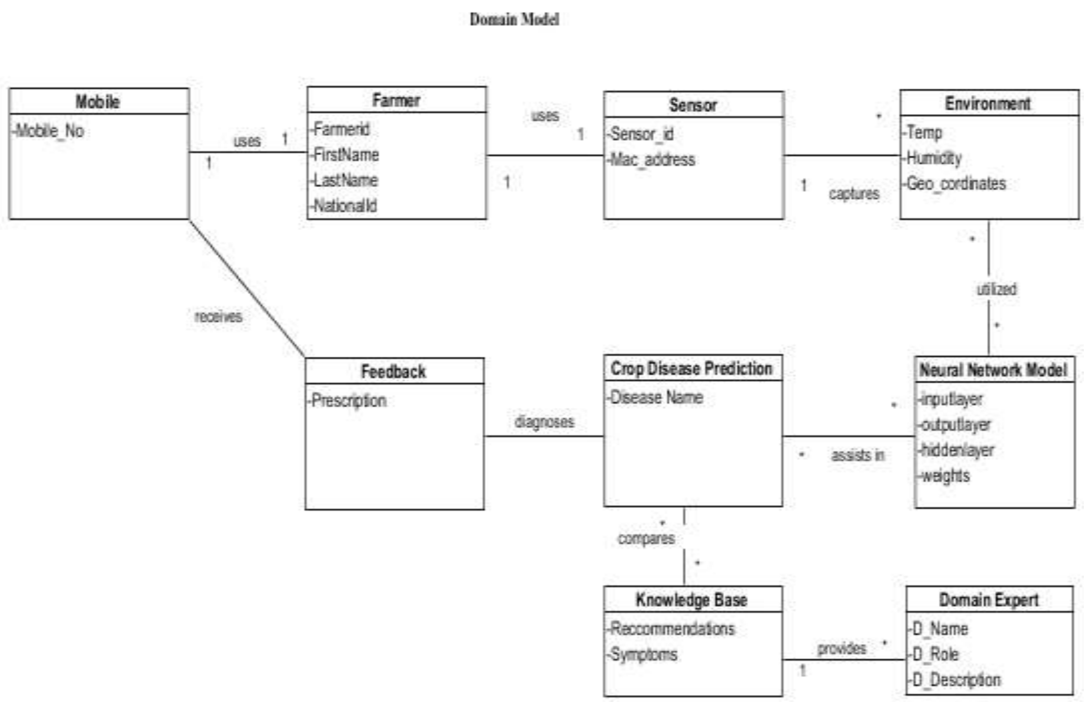


Figure 4.0.3: Domain Model

4.4.3 Sequence Diagram

Sequence diagrams are used to illustrate how objects in a system interact with each other. Figure 4.4 below illustrates the sequence diagram for this study. The farmer captures farm conditions information via an Arduino based sensor setup. The information once uploaded to the server, humidity and temperature readings are extracted and normalized. The normalized data sets as input weights to be used by the crop disease prediction module. The normalized data is further used for test and validation purposes. The crop disease prediction module utilizes an artificial neural network algorithm to determine the risk of late blight potato disease. The level of risk blight infection and corrective measures are then passed as feedback to the farmer.

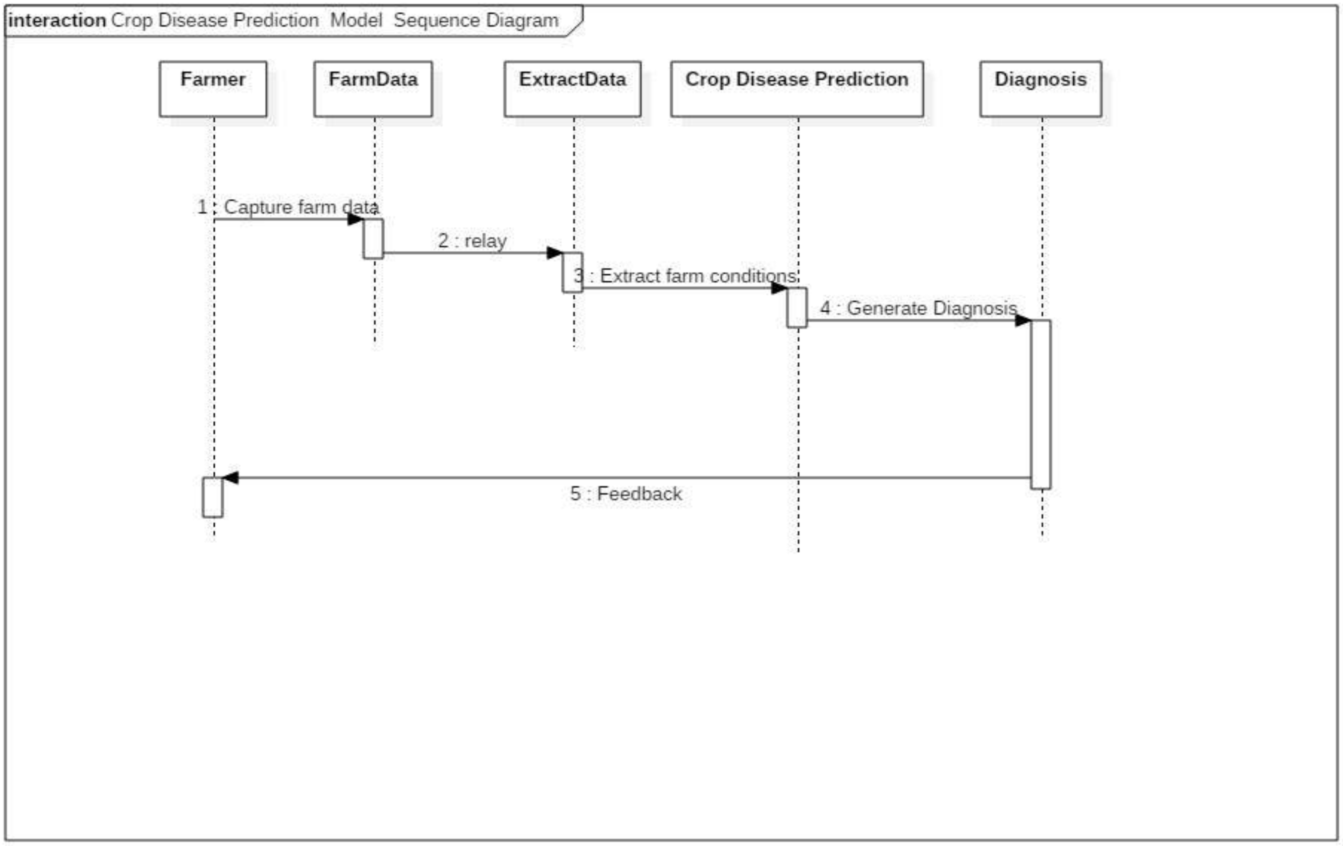


Figure 4.0.4: Sequence Diagram

4.4.4 Neural Network Activity Diagram

An important modeling artifact used in UML, is the Activity Diagram. Activity Diagrams are used to indicate sequence of actions as part of a process flow. Activity diagrams are used to model sequence of actions that capture process flow actions and their results. Activity diagrams concentrates on the work carried out in the implementation of an operation or a method and activities in a use case instance of an object. (Bhattacharjee & Shyamasundar, 2009)

Figure 4.5 below illustrates the neural network activity diagram for the study. In meeting the objective of the study, activities that were involved in giving an accurate and timely prediction of potato late blight disease involved capturing specific measurable potato farm conditions. Temperature and humidity values that represent environmental stress factors were obtained through use of temperature and humidity probes. The following explain further key workings of the components identified in figure 4.6 below.

a) Input temperature and humidity data values- This data represent measurable farm conditions realized at a farm. Blight units are estimated based on monthly mean values of temperature and humidity readings

b) A correct prediction of potato late blight disease and the degree of risk of its risks formed the expected output. Based on the resistance of the potato variety identified in section x above. The risk of infection is then classified as high, moderate, and low

c) The sigmoid activation function was used as the activation function in the neural network. Adjustments to weights was carried out to enable the machine to have a better environment to learn. This is very important in increasing the accuracy of the machine learning algorithm. Testing data set was supplied to the back propagation algorithm to verify and validate if the algorithm was producing the expected output.

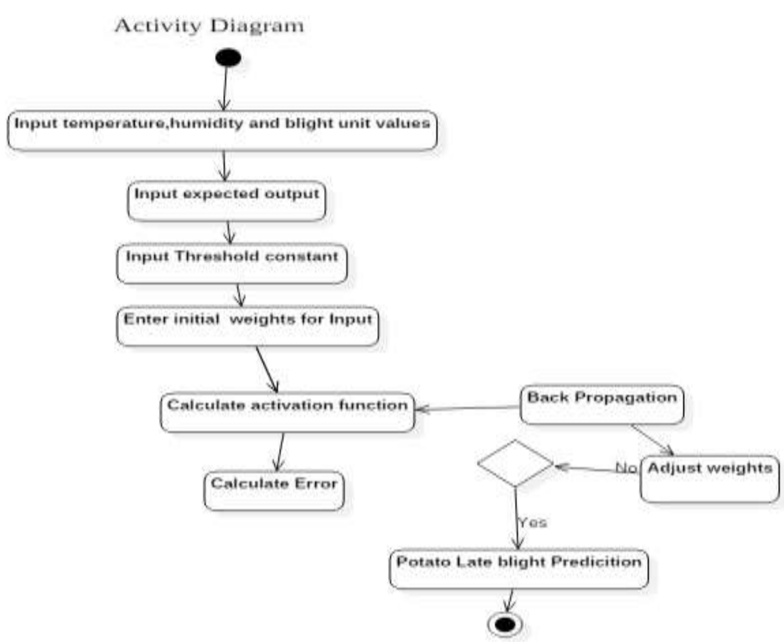


Figure 4.0.5: Activity Diagram

4.5 Model Design

4.5.1 Context Diagram

Figure 4.6 below illustrates the data flow among the entities and the system. The main entities (users) identified for this study are the farmer, domain expert and agricultural extension officer. The farmer will use temperature and humidity probes to log humidity and temperature parameters to the system. The domain expert provides expert knowledge on potato late blight disease and recommended treatment. The agricultural extension officer can use the system to monitor and locate farmers to offer on farm extension services and advice on potato late blight disease treatment and management.

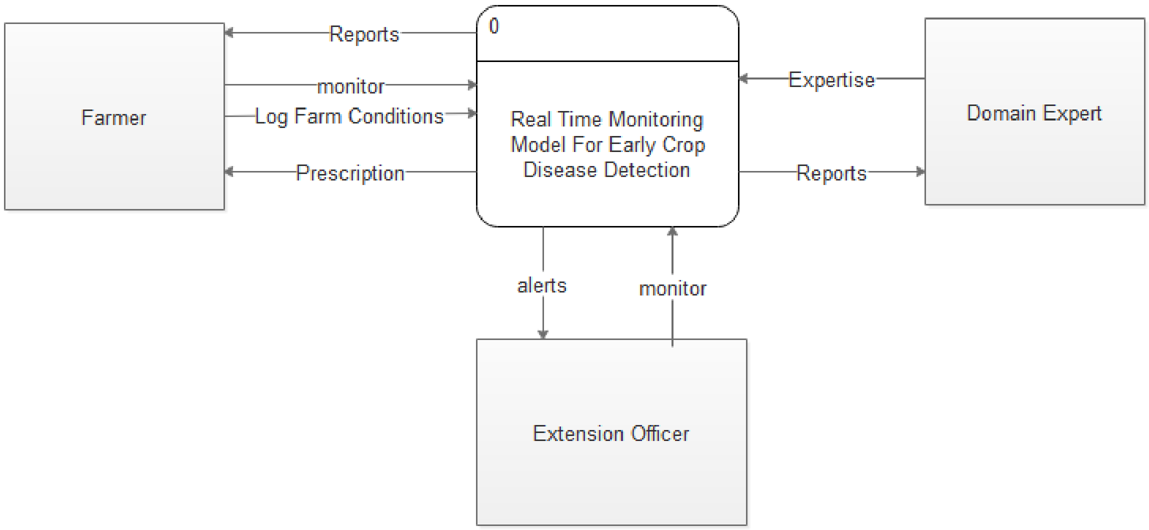


Figure 4.0.6: Context Diagram

4.5.2 Level 0 Data Flow Diagram

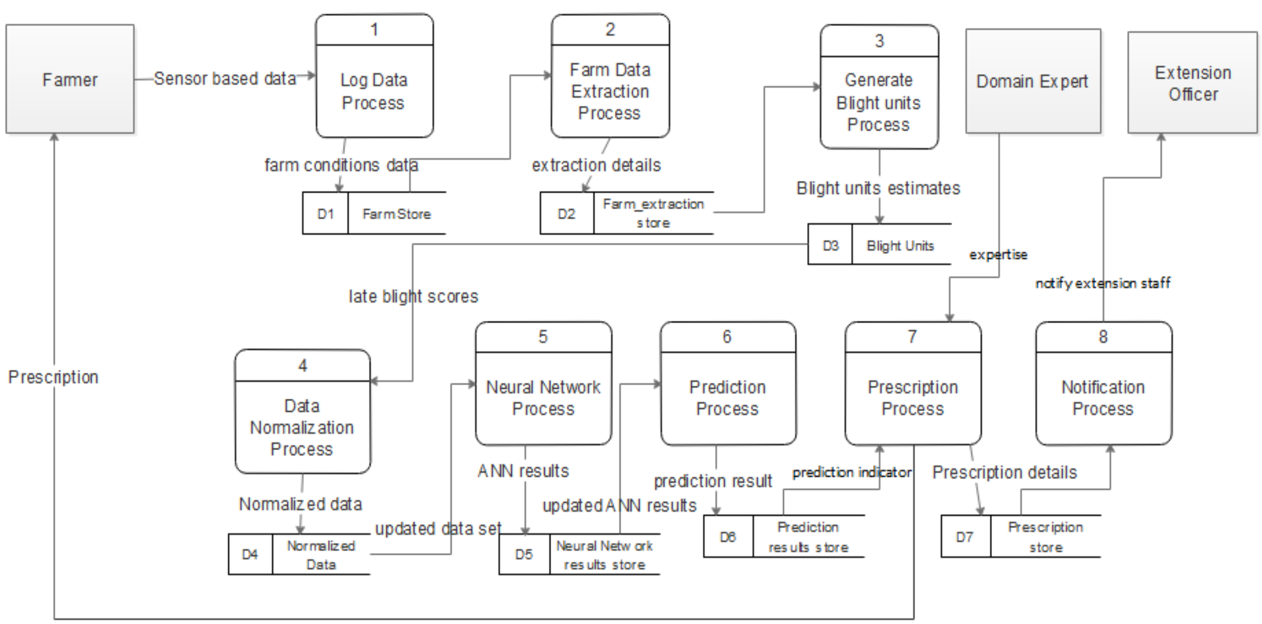


Figure 4.0.7: Level 0 Diagram

Figure 4.7 above represents the level 0 data flow diagram elaborating the key processes of the system and how they interact with the various entities. The arrows depict direction of messages from entities to processes and to respective data stores. The first process of the real time model for early detection of crop diseases is the Log Data process. Sensor data from the farm acts as the source of data for this process. The data is then stored in a farm data store. The data is the extracted by Farm data extraction process. This data is then utilized to generate blight unit scores. Both farm data and blight unit scores are then normalized. The neural network process uses a training data set consisting of normalized data from the normalized data store. Prediction of potato late blight disease is then computed by the prediction process. The resultant information is then stored and used by the recommendation process which will advise the farmer on the correct course of action. The notification process utilizes prediction results and prescription process to alert an agricultural extension officer based on certain threshold.

4.6 Arduino Schematic Design

Figure 4.8 below illustrates the schematic design of the sensor probe set up to an Arduino microcontroller. An Arduino uno microcontroller board was used for this study. DHT11 temperature and humidity probe was identified for the study to facilitate developing of the prediction model. The sensor captures temperature data from the environment in degrees Celsius while humidity is recorded as a percentage. Sim 900 was used to transmit data via a rest api to the server.

Figure 4.0.8: Arduino Schematic Design

The following are the ranges and accuracy levels provided by the sensor:

i. Humidity Range: 20-90% RH
ii. Humidity Accuracy: ±5% RH
iii. Temperature Range: 0-50 °C
iv. Temperature Accuracy: ±2% °C
v. Operating Voltage: 3V to 5.5V

This study chose the DHT11 as it is lab calibrated, accurate and stable. In addition the signal output is digital which can easily be transmitted to a server. Further, the sensor is relatively inexpensive given the performance. A sample code below show how the device will transmit data.

```
#include <dht.h>
dht DHT;
#define DHT11_PIN 12
#include <Event.h>
#include <Timer.h>
#include "SIM900.h"
#include <SoftwareSerial.h>
#include "inetGSM.h"
//#include "sms.h"
//#include "call.h"
Timer ti;
//To change pins for Software Serial, use the two lines in GSM.cpp.

//GSM Shield for Arduino
//www.open-electronics.org
//this code is based on the example of Arduino Labs.

//Simple sketch to start a connection as client.

InetGSM inet;
//CallGSM call;
//SMSGSM sms;

char msg[50];
int numdata;
char inSerial[50];
int i=0;

boolean started=false;
static char postUrl[150];
char *api_key="ee13504e4b14163cb32331dfd013f3af";

void setup()
{
     //Serial connection.
     Serial.begin(9600);
     Serial.println("GSM Shield testing.");
     //Start configuration of shield with baudrate.
```

```
     //For http uses is raccomanded to use 4800 or slower.
     if (gsm.begin(2400)) {
          Serial.println("\nstatus=READY");
          started=true;
     } else Serial.println("\nstatus=IDLE");

     if(started) {
          //GPRS attach, put in order APN, username and password.
          //If no needed auth let them blank.
          if (inet.attachGPRS("internet.wind", "", ""))
               Serial.println("status=ATTACHED");
          else Serial.println("status=ERROR");
          delay(1000);

          //Read IP address.
          gsm.SimpleWriteln("AT+CIFSR");
          delay(5000);
          //Read until serial buffer is empty.
          gsm.WhileSimpleRead();

          //TCP Client GET, send a GET request to the server and
          //save the reply.


     }
};

void loop()
{

    //Read for new byte on serial hardware,
     //and write them on NewSoftSerial.
     serialhwread();

     //Read for new byte on NewSoftSerial.
     serialswread();


     Serial.println(" Humidity " );
//float h = DHT.humidity;
Serial.println(DHT.humidity, 1);

Serial.println(" Temparature ");
//float t = DHT.temperature;
Serial.println(DHT.temperature, 1);
 int chk = DHT.read11(DHT11_PIN);

                  int h = DHT.humidity;
                  int t = DHT.temperature;
                  char tempStr[15];
                  char humidStr[15];
                  dtostrf(t, 4, 2, tempStr);
                  dtostrf(h, 4, 2, humidStr);
```

```
          sprintf(postUrl,
"/api/data/push?id=4rps21&t=%s&h=%s&api_key=%s",tempStr,humidStr,api_key);
          numdata=inet.httpGET("41.215.34.154", 80,postUrl, msg, 50);
delay(2000);


          //Print the results.
          Serial.println("\nNumber of data received:");
          Serial.println(numdata);
          delay(5000);
          Serial.println("\nData received:");
          Serial.println(msg);
          delay(15000);

};


void serialhwread()
{
     i=0;
     if (Serial.available() > 0) {
          while (Serial.available() > 0) {
               inSerial[i]=(Serial.read());
               delay(10);
               i++;
          }

          inSerial[i]='\0';
          if(!strcmp(inSerial,"/END")) {
               Serial.println("_");
               inSerial[0]=0x1a;
               inSerial[1]='\0';
               gsm.SimpleWriteln(inSerial);
          }
          //Send a saved AT command using serial port.
          if(!strcmp(inSerial,"TEST")) {
               Serial.println("SIGNAL QUALITY");
               gsm.SimpleWriteln("AT+CSQ");
          }
          //Read last message saved.
          if(!strcmp(inSerial,"MSG")) {
               Serial.println(msg);
          } else {
               Serial.println(inSerial);
               gsm.SimpleWriteln(inSerial);
          }
          inSerial[0]='\0';
     }
}

void serialswread()
{
     gsm.SimpleRead();
}
```

Chapter 5: Implementation and Testing

5.1 Introduction

The implementation was carried out in four parts. The first initial set up was assembling DHT11 sensors with the microcontroller fitted with a gsm module. The sensor was placed 0.1 m above the soil in a potato farm to capture humidity readings. The data readings from the sensor were sent to the server via an MQTT protocol to a cloud based server. Daily data extracts were then used to come up with disease severity values known as blight units. Back propagation neural network algorithm was then used to predict occurrence of blight diseases. A training set was then presented to the neural network to activate the multilayer perceptron. Prediction of late blight potato disease was then carried out. Prediction results were then produced and farmer received notification message via an android application and email.

5.2 Model components

5.2.1 Weather Data Logging Components

i. DHT 11 sensor – The model requires the use of a probe to capture humidity and temperature readings. The probe is a physical device placed on the farm to capture these conditions.

ii. Arduino uno micro controller- The device act as a micro controller that can be programmed to capture probe data.

iii. Sim 900 GSM module- The model requires that the data parameters are sent to the server via an api route. The gsm module facilitates transmission of data from the farm to remote server via TCP or MQTT protocol.

5.2.2 Simcast Model Algorithm

Simcast model was used to come up with daily blight units or daily severity values. Simcast takes into account humidity and temperature readings from 12 to 12.The model also used resistant factor of three potato varieties in the form of severe, fairly resistant and resistant. The SimCast algorithm estimates the risk of damaging late blight levels, expressed as ‘blight units’, based on the temperature during the consecutive hours in a day when relative humidity is above 90%. SimCast thus uses hourly weather data as input.

5.2.3 Neural Network Components

Back propagation algorithm was used in predicting potato late blight. The components of the algorithm are mentioned below.

5.2.3.1 Input Layer

In an artificial neural network, the input layer acts as the first layer of the network. The layer consisted of one neuron for each specific attribute used by the network to predict potato late blight disease. The number of neurons elaborate how the input layer is structured. The input layer interacts with the external environment and consists of an independent variable interacting with the environment.

5.2.3.2 Output Layer

The final layer is the output layer, where there is one node for each class. The output is provided after all the inputs have been processed and presents a pattern to the external environment. A single sweep forward through the network results in the assignment of a value to each output node. The record is assigned to the class node with the highest value.

5.2.3.3 Hidden Layer

The second layer is the hidden layer, in which the processing units are interconnected to the layers below and above it. The hidden layer functions to ensure better results of the output are achieved given a set of input.

5.3 Model Implementation

5.3.1 Sensor Data Capture

Information of farm conditions on temperature and humidity were captured using farm sensor probes. The sensor used for this research included a DHT11 temperature and humidity sensor. The user placed the sensor on a potato farm plot at 0.01m above the soil. Figure 5.1 below illustrates a sample output of the sensor readings.

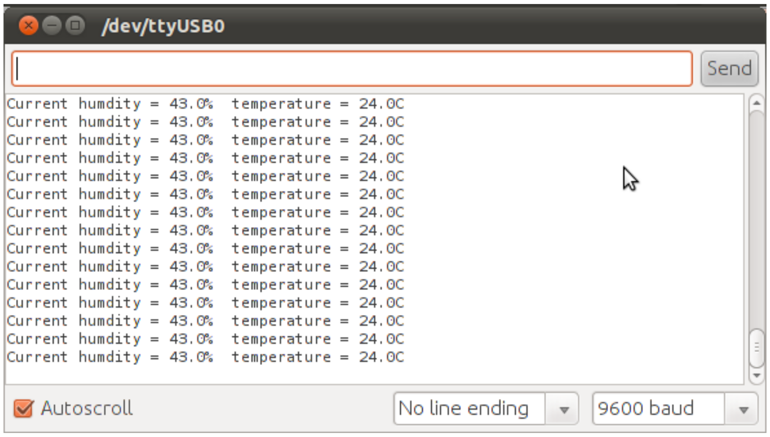


Figure 5.0.1: DHT 11 Serial monitor sample output

5.3.2 Data Preprocessing

Data preprocessing is a requirement for any data mining project. Data preprocessing involves cleaning of data set in order to support the data mining and modelling process. Several steps were used in in data cleaning as explained in sections below. The activities include,

5.3.3 Data for Use for the Model

The model for this study focused on environmental conditions influencing development of ***Phytophthora Infestants***. Historical hourly data set from the metrological department of Kenya for Nakuru county was used.

5.3.4 Simcast Algorithm

Figure 5.2 below illustrates the output of the simcast algorithm comprising of daily mean temperature and humidity and accumulated blight units. Simcast blight units represent the favorability of the prevailing weather for late blight progress and are also calculated based on the relationships between duration of relative humidity periods >= 90% and average temperature during those periods. The output of the algorithm constitutes blight units, daily average temperature and daily average humidity readings that form part of the input for the neural network. The simcast algorithm has been referenced in Appendix G.

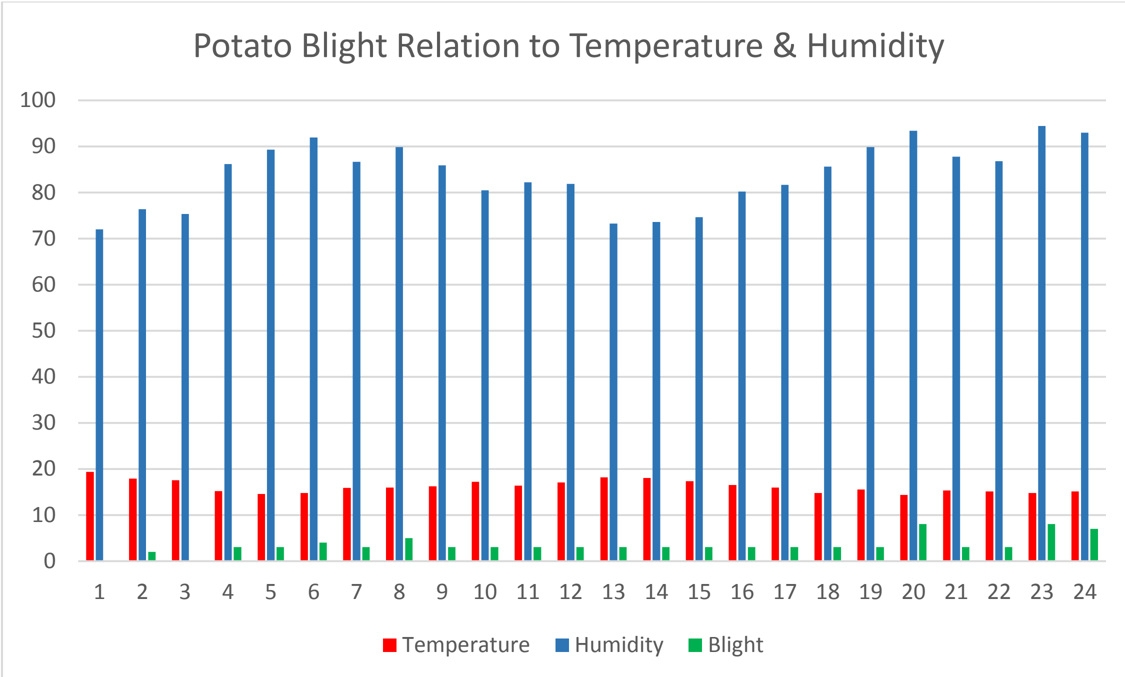
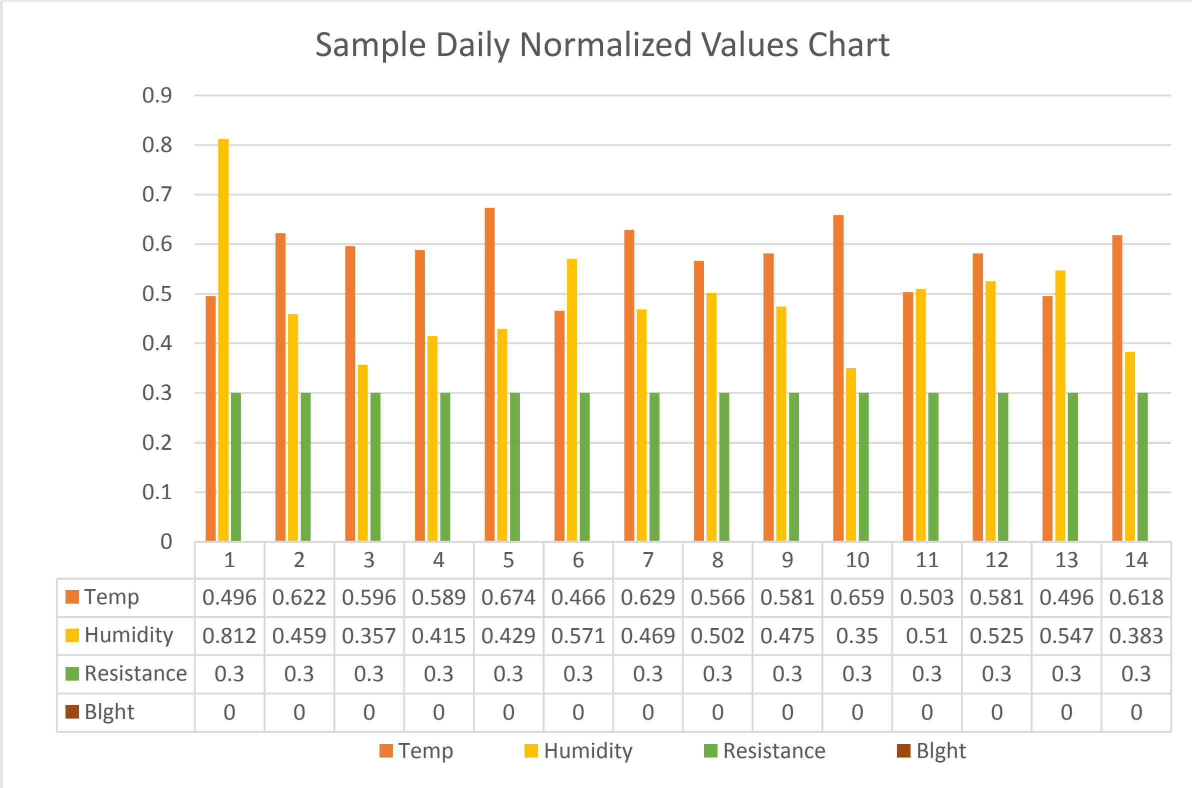


Figure 5.0.2: Sample Blight Units, Temperature and Humidity Trend from Simcast Algorithm

5.3.5 Data Normalization

Data normalization is a key step for data preparation in a data mining project. Data normalization is the process where data of attributes are fit in a specific range. Neural networks are designed to accept float values as their inputs. Since the model made use of a sigmoid function, the range of values used for the study ranged between 0 and +1. The min max scaler algorithm was used to fit values.

In developing the late blight forecasting model based on neuron network both farm condition data and blight units are normalized. Equation 4.1 above was utilized to carry out this function. Figure 5.3 below illustrates a chart showing daily normalized values for temperature and humidity as well as a classification of accumulation of blight units -1 indicating low, 1 indicating moderate and 2 indicating high.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Temp	0.496	0.622	0.596	0.589	0.674	0.466	0.629	0.566	0.581	0.659	0.503	0.581	0.496	0.618
Humidity	0.812	0.459	0.357	0.415	0.429	0.571	0.469	0.502	0.475	0.35	0.51	0.525	0.547	0.383
Resistance	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
Blght	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 5.0.3: Sample chart of showing normalized values

5.3.6 Neural Network and Model Implementation

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.

5.4. Training and Testing the Model

Training data set was constructed from daily means of temperature and humidity readings from historical metrological data for Nakuru County for the 2015 year. Adjustment of weights was carried out to produce more accurate results. Training was conducted routinely until the error rate reached an appropriate accepted. Epoch of 2500 iterations was set. Accurate results were obtained using a higher epoch rate compared to using a low epoch rate. A split ratio of 50:50 was used to segregate the normalized data into training data set and validation set. The test was utilized to validate if the model produced desired output. Back propagation algorithm was used to train the neural network. The algorithm tasks constituted a feed forward approach and a back pass. A sigmoid activation function was used.

5.5 Software Flow

The key deliverable of the study was to come up with a real time monitoring model for early detection of crop diseases. A proof of concept application was developed. The application utilized sensors to record temperature and humidity readings which are key environmental conditions for monitoring potato late blight disease. This information was transmitted via a normal gsm subscriber line to the server using MQTT protocol. A web and android application were developed to allow farmers and an expert to visualize and monitor trends of these variables.

5.6 Model Architecture

Figure 5.4 below describes the architecture of the model used for this study. Farm conditions were captured through use of sensor probes placed on the farm which transmitted the information to a central server. Hourly data extracts were utilized in estimation of accumulation of blight units. Scaled daily blight units, temperature, humidity readings and potato late blight resistance factor were then fed to the neural network. The output of the model consisted of target classes of the intensity of blight as high, medium or low. The output classes inform on the severity of potato blight infection. A low blight prediction requires monitoring, a moderate and high blight intensity prediction indicates that corrective measures such as use of fungicide or thorough farm inspection is required to be carried out.

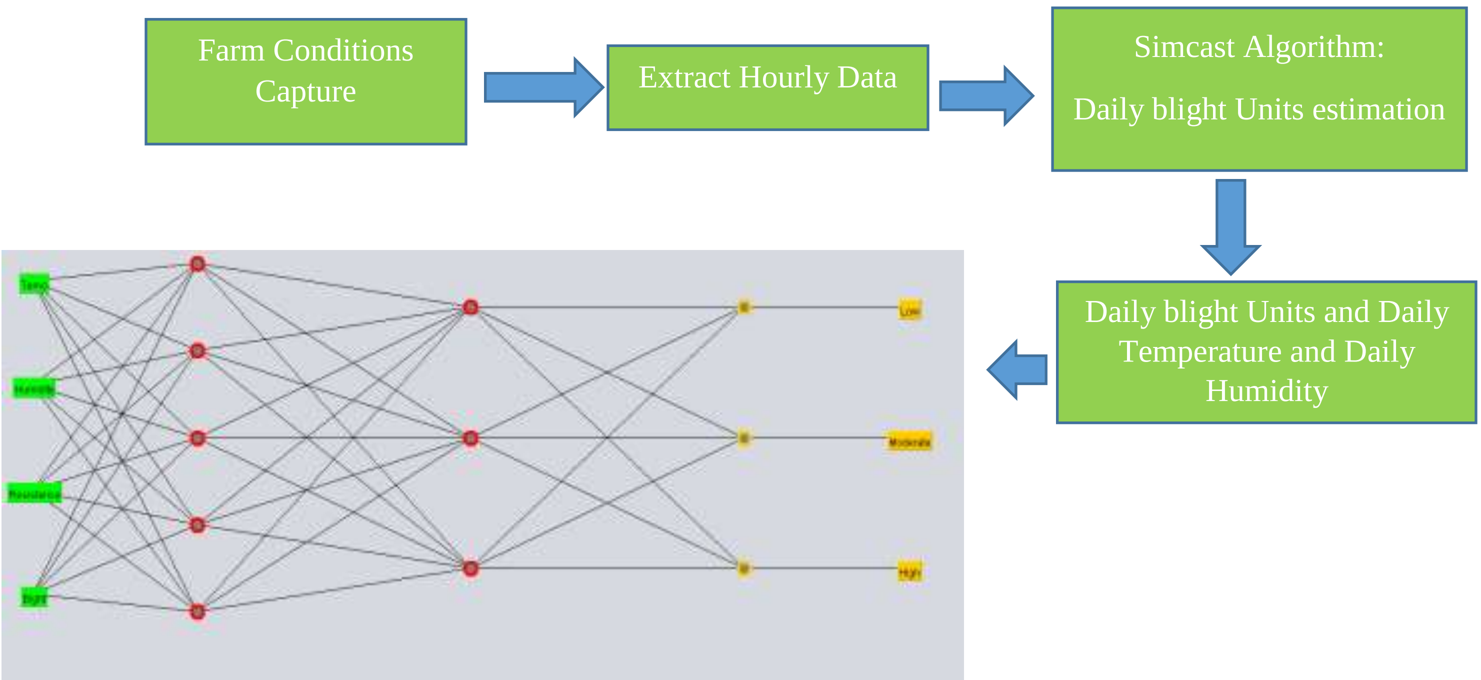


Figure 5.0.1: Model Architecture

5.6.1 Model Data Description

This study utilized historical weather data in developing the model. This study relied on 4 input attributes and three output classes. A total of 836 instances were used during the development of the model. The output class variable for the study included High, Low and Medium. Blight units represents disease pressure or disease severity on a potato plant.

5.7 Model Testing

Table 5.1 below illustrates the test cases used in developing the model. This activity involved testing reliability and functionality parameters.

Table 5.1: Model Test Case

Test class	Inspection Check	Priority
Functional	Does the application allow a farmer to register and log in	High
Functional	Does application allow for transmission of Temperature and Humidity readings via an api.	High
Functional	Does the application abide by KALRO and CIP symptoms of potato late blight disease	High
Reliability	The applications should be able to provide warning messages to users	Medium

5.7.1 Model Testing Results

The application successfully validated user information and valid readings from dht11 sensors in the required format via an api. A unique api key and unique id was created for a user to allow only registered users to utilize the service.

Table 5.2: Model Test Case Results

F.ID	Test Results	Comment
Functional	Pass	Temperature , humidity, Blight units and resistance factor used to identify disease
Reliability	Pass	Potato late blight treatment based on recommendations by KALRO and CIP

5.8 System Testing

The research tested the model to evaluate how the model performed in relation to the limited agricultural extension officers and potato late blight experts visiting various farms. Table below illustrates the test case for performance.

Table 5.3: System Test Case

Test class	Inspection Check	Priority
Performance	Does the prediction of late and advisory process take a short duration	High

5.9 Component Testing

5.9.1 Android Module Testing

Table 5.4: Android Module Test Case

Test class	Inspection Check	Priority
Usability	Does	Moderate
Usability	Does the farmer navigate easily via menus	Moderate
Performance	Does the application crash on running functions	Moderate

5.9.2 API Module Testing

Table 5.5: Api Module Test Case

Test class	Inspection Check	Priority
Functionality	Do the sensors transmit data via the api	High
Performance	Does the api reduce time taken to transmit and receive information	Moderate

5.9.3 Arduino and Sensor Module Testing

Table 5.6: Arduino Sensor Module Test Case

Test class	Inspection Check	Priority
Functionality	Does the sensor capture temperature and humidity readings in correct format	High
Performance	Does the sensor communicate seamlessly with the Arduino microcontroller	High

5.9.4 Sim5320 Data Transmission

Table 5.7: 3G Module Test Case

Test class	Inspection Check	Priority
Functionality	Does the module connect to a mobile subscriber network via a sim card	High
Performance	Does the module facilitate faster data transfer to server via MQTTT or TCP protocol	Moderate
Functionality	Does the module support AT commands	Moderate

5.10 Acceptance Testing

Acceptance testing illustrated by Table below was carried out to verify whether usability aspects of the application were achieved.

Table 5.8: Acceptance Test Case

Test Class	Inspection Check	Priority
Usability	Does the application meet user requirements	High
Usability	Are end users of the application satisfied by the output	Moderate

Chapter Six: Discussions

6.1 Introduction

Internet of things and machine learning techniques are being adopted in smart farming. The real time monitoring model for early detection of crop diseases utilized these two principles. Temperature and humidity probes connected to an arduino microcontroller were used in recording the favorable conditions for potato late blight disease. Simcast model was useful in determining disease pressure or accumulation of blight units factoring susceptibility to potato late blight. In forecasting occurrence of potato blight, back propagation neural network was used to classify presence of blight as high, medium and low. The model was evaluated for accuracy using root mean square technique.

Farmers in Nakuru county largely depend on intuition and human vision for monitoring crops and occurrences of diseases, additionally cultural and routine practices of fungicide use without proper guidance are some of the farming practices are prone to inaccuracies and are costly. Farmers in the county heavily depend on agricultural extension officers and experts from government and non-governmental organizations to assist in identifying potato late blight disease and giving treatment recommendations. These officers are too few and usually take time to reach the farm. This problem is also compounded by the fact that majority of the farmers are served by a poor road network. This delay would lead to late diagnosis of potato late blight disease hence this significantly affected control and treatment of the disease.

This study through use of a structured interview gained insight and input from potato late blight disease expert and an agricultural extension officer at the county. The results of the interview identified use of farm weather probes and a mobile application would be instrumental in monitoring and assisting in early detection of potato late blight disease. Farmers in the region lack tools to assist in monitoring and identifying potato late disease.

In addition, queries on the spread of the disease and its uniqueness in Kenya revealed that for potato late blight disease in Kenya is ubiquitous. This can be alluded by the fact that the ecological landscape of the country provide for a temperate setting that favors development of ***Phytophthora Infestants.***

A key deliverable of the study was to come up with a model for potato late blight disease prediction. The model utilized an Arduino based sensor set up to capture farm conditions.The prediction model developed for this study gives a more accurate prediction and classification of potato late blight disease since it was developed using a neural network algorithm. Machine learning algorithm facilitated the model to be faster, efficient and robust, thus a farmer could receive feedback and notifications within a short period. In utilizing the benefits and strengths of internet of things technology, mobile computing and machine learning techniques enabled the potato late blight prediction model developed by this research to provide accurate results for early detection of potato late blight disease and recommendation amendments for treatment of the disease.

6.2 Validation of the Model

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 93.8995%.

Table 6.1: 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**	

6.2.1 Validation Class

Table 6.2: Validation Class Results

	Tp Rate	Fp Rate	Precision	Recall	F-measure	ROC Area	Class
	0.980	0.262	0.949	0.968	0.964	0.951	Low
	0.731	0.041	0.853	0.878	0.876	0.949	Moderate
	0.000	0.000	0.000	0.000	0.876	0.935	High
Weight Avg	0.936	0.222	0.930	0.936	0.932	0.959	

The performance evaluation for this study was based on the precision, recall ratio and F-Measure per class as stated in Chapter 3 above. Table 6.2 above indicates the results of the various evaluation parameters. TP represents the true positive, FP represents the false positives. The Receiver Operator Characteristics (ROC) refers to the classification the false positives and false negatives.

6.2.2 Confusion Matrix

Table 6.3 below illustrates the classification matrix that was obtained by the potato late blight prediction model. The confusion matrix describes information on the actual and predicted classifications. A total of 1461 instances were used to train and test the network. 1193 instances were correctly classified as Low and 174 instances were correctly classified as having a moderate intensity of blight accumulation. 64 instances were incorrectly classified as low and 6 instances were incorrectly classified as moderate.

Table 6.3: Confusion Matrix

Low	**Moderate**	**High**	
1193	24	0	**Low**
64	174	0	**Moderate**
0	6	0	**High**

6.3 Contribution of the Model to Research

Taking account of the various challenges farmers are faced with in identifying, monitoring and controlling potato late blight disease, the real time monitoring model for early detection of crop diseases developed by this research provided a more improved solution. Current human visual based techniques are poor in early identification and estimating of potato late blight disease. This research utilized novel principles of internet of things through use of sensors to monitor farm environmental conditions. The model utilized back propagation to give a more accurate classification of the intensity of potato blight disease.

The model presented by this 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 human vison 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 abiotic environmental stress variables that promote growth, spread and development of crop diseases. This study utilized temperature and humidity readings as some of the inputs in the prediction model.

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 be accorded on farm services on control measures that can curb spread of the crop disease.

6.4 Research shortfalls

The developed model had the following shortfalls:

i.) The model did not consider all the biotic and abiotic stress factors that influence potato late blight disease development.

ii.) The model was limited to potato crops

iii.) The model did not consider farmer cultural practices which influence the transfer of the disease from one plot to another.

iv.) For the purposes of this study, a moderate susceptible cultivator was used in developing the model.

Chapter 7: Conclusion and Discussions

7.1 Conclusions

As revealed by the extension officer through a structured interview potato farmers in Nakuru county face several challenges in early detection, identification of potato late blight disease and more often lack sound crop management skills. Farmers and extension staff lack intelligent tools to monitor and identify potato diseases. Nakuru County is one of the leading producers of potato for both domestic and commercial use. Farmers in the region rely heavily on human vision in identifying of potato diseases. The drawback of such technique can lead to misdiagnosis hence contributing to low yield. Potato late blight disease unlike other potato diseases is costly to manage since farmers more often rely on fungicide use to manage and control the spread of the disease. Early and accurate detection and diagnosis of potato late blight disease.

In addition farmers do routinely use uncertified seeds for planting. Recycling of seeds and use of uncertified seeds have been noted to be susceptible to potato late blight disease. Poor farm and crop management coupled with cultural routine practices have been observed as other factors that promote the spread and development of potato late blight disease.

This study utilized internet of things techniques and machine learning for the prediction of potato late blight disease. The research on monitoring environmental conditions through use of sensor probes. The research focused on temperature and humidity which represent favorable condition for sporangia development. For the purposes of monitoring change in the phenotypic characteristics once an alert was notified to a user, an android application form was created to capture image of leaves for further analysis by an expert at a remote office.

The per capita consumption in Kenya is around 25 kg annually. Potato occupies a prime position in terms of the contributions to food security, poverty eradication and economic development in the region. Consequently, there is growing attraction to the production of potato. The area under which it is grown increases steadily over the years, principally in smallholder production systems Potato is the second most consumed crop after maize in Kenya and has an economic importance to the growers, hence it is therefore necessary to ensure that potato late blight disease stress on potato is reduced to enhance on high yields and returns. Information technology provides necessary tools and infrastructure that can be used at real time for early detection of crop diseases.

The Kenyan government through the ministry of agriculture has embarked on distribution of disease free seeds. Such initiatives are welcomed as they reduce the likelihood of infection of potato farms by the disease. Consequently, the Kenya Plant Health Inspectorate recognizes the need to adequately use and apply information technology in early detection of crop diseases. Early detection of crop diseases leads to better control and management. This results in higher yields and hence a more food secure nation. (Otipa, et al., 2015)

Effective potato late blight disease management should focus on disease prevention, farm sanitation, farmer cultural practices, farm field monitoring and a precise fungicide spray program hence farmers should closely check on conditions favorable for potato blight disease development.

7.2 Recommendations

i.) The researcher recommends use of leaf wetness information should be considered when modelling for potato late blight disease. Leaf wetness will provide a clear indication on the probability of leaf infection by blight. Leaf wetness is the presence of free water on the surface of a crop canopy. This research utilized humidity to act as a proxy variable as a measure of environmental stress factor.

ii.) The model could be further expanded to include other disease affecting diseases affecting potato such as early blight. Symptoms of potato early blight and potato blight share similar symptoms.

iii.) The use of unmanned aerial vehicles or on farm image sensors would be useful in monitoring potato late blight disease development in large farms.

iv.) *Phytophthora infestants* being largely transmitted by wind, the researcher recommends factoring wind direction and use of spore traps to assist in determining potato late blight disease transfer

7.3 Suggestions for Future Research

i. The researcher recommends use of fungicide amendments schedule to be part of the model to ensure accuracy of early detection of potato late blight disease.

ii. 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.
iii. 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 potato late blight disease
iv. The researcher recommends factoring biotic factors that influence spread of potato blight disease
v. The researcher recommends utilization of remote sensing technique for monitoring potato late blight disease

References

Adams, M. L., Cook, S., & Corner, R. (2000). Managing Uncertainty in Site-Specific Management: What is the Best Model? *Precision Agriculture*, 39–54. doi:10.1023/A:1009984516714

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

AK,Mahlein. (2016). Plant Disease Detection by Imaging Sensors – Parallels and Specific Demands for Precision Agriculture and Plant Phenotyping. *APS Journals*, 241-251.

Apple, A., & Fry, W. E. (1983). *Potato Late Blight Fact Sheet Page*. Retrieved 11 15, 2016, from Cornell University: http://vegetablemdonline.ppath.cornell.edu/factsheets/Potato_LateBlt1983.htm

Arora , R., Sharma , S., & Singh , B. (2014). Late blight disease of potato and its management. *Potato Journal*, 14-60.

Arsenault, C. (2015, June 10). *New app to connect Kenyan farmers with climate-smart seeds*. Retrieved from Reuters: http://www.reuters.com/article/food-africa-climate-idUSL5N0YW1W220150610

Bhattacharjee, A. K., & Shyamasundar, R. K. (2009, January-February). Activity Diagrams : A Formal Framework to Model Business Processes and Code Generation. *Journal of Object Technology*, *8*(1).

Campbell, C., & Madden, L. (1990). *Introduction to plant disease epidemiology*. New York: John Wiley & Sons.

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.

Djeugap, J. F., Eko, D., Nguefack, J., Columbus, T. N., & Fontem, D. A. (2014, October 9). Effect of organic amendments and fungicide application on potato late blight, bacterial wilt and yield in Cameroon. *International Journal of Agronomy and Agricultural Research (IJAAR), 5*(4), 12-19. Retrieved from http://www.innspub.net

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.

Douglass , B. P. (2004). *Real Time UML Advances in the UML for Real-Time Systems.* Pearson Edition.

Elliott, M. (2015, January 31). *Sustainable and scalable business models. ICT Update Bulletin, Issue no. 80.* Retrieved from CTA website: http://ictupdate.cta.int/Regulars/Q-A/Sustainable-and-scalable-business-models/%2880%29/1439915106

Fry, W., Apple, A., & Bruhn, J. (1983). Evaluation of potato late blight forecasts modified to incorporate host resistance and fungicide weathering. *Phytopathology*, 1054-9.

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. Retrieved from http://www.irdindia.in/journal_ijraet/pdf/vol2_iss3/1.pdf

Gichamba, A., & Lukandu, I. A. (2012). A model for designing M-agriculture applications for dairy farming. *The African Journal of Information Systems, 4*(4).

Girish, A., & Priti, B. (2016). Potato Leaf Diseases Detection and Classification System. *International Journal of Computer Science and Mobile Computing, 5*, 76-88.

Government of Kenya (GOK). (2010). Agricultural Sector Development Strategy 2010- 2020. Retrieved November 9, 2016, from http://faolex.fao.org/docs/pdf/ken140935.pdf

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

Government of Kenya. (2012 a). *Kenya National Climate Change Action Plan 2013 - 2017.Nairobi.*

Guide for Soil Moisture Sensor YL-69 or HL-69 with Arduino. (n.d.). Retrieved November 20, 2016, from randomnerdtutorials.com: https://randomnerdtutorials.com/guide-for-soil-moisture-sensor-yl-69-or-hl-69-with-the-arduino/

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 (CCAFS), Copenhagen. Retrieved from www.ccafs.cgiar.org

Hiroyuki, T., Hiromitsu, F., & Seiji, C. (2007). Prediction of Disease Infection of Welsh Onions by Rust Fungus Based on Temperature and Wetness Duration. *2007 IEEE International Conference on Control Applications*, (pp. 325-330). Singapore. Retrieved from http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4389251&isnumber=4389191

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

Irwin, G., & Turk, D. (2005, January). An ontological analysis of use case modeling grammar. *ournal of the Association for Information Systems, 6*(1), 1-36. Retrieved from https://pdfs.semanticscholar.org/db93/b312f15f2ea41c363f6be64e08d0d09df91d.pdf

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

Kaguongo, W. (., Gildemacher, P., Demo, P., Wagoire, W. (., Kinyae, P., Andrade-Piedra, J. L., . . . Fuglie, K. (. (2008). *Farmer practices and adoption of improved potato varieties in*

Kenya and Uganda. International Potato Center. Retrieved November 15, 2016, from http://cipotato.org/wp-content/uploads/2014/08/004365.pdf

Kenya, G. o. (2012 a). *Kenya National Climate Change Action Plan 2013 -2017.* Nairobi.

Koné , M. T. (2013). *The Future of Human Evolution*. Retrieved from http://futurehumanevolution.com/artificial-intelligence-future-human-evolution/artificial-neural-networks

Kothari, C. R. (2004). *Research Methodology: Methods and Techniques.*

Likhar, K., Bisen, N., & Dubey, S. (2016, August). A Coverage and Latency Aware Route. *International Journal of Science, Engineering and Technology Research (IJSETR), 5*(8).

Lipper, L., Thornton, P., Campbell, B. M., Baedeker, T., Braimoh, ., Bwalya, M., . . . Torquebiau, E. F. (2014). Climate-smart agriculture for food security. *Nature and climate change*. Retrieved from http://dx.doi.org/10.1038/nclimate2437

Maina, C. N. (2016). Vision-Based Model for Maize Leaf Disease Identification: A Case Study in Nyeri County.

Mayur, B., Mayur, B., Akshay, N., & Sachin, D. (2016). Smart farming using Arduino and data mining. *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 913-1917). New Delhi. India: IEEE. Retrieved November 20, 2016, from http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7724599&isnumber=7724213

Milos, I., Petar , S., Mladen , V., & Abdolkarim, A. (2015, November 24-26). Data mining model for early fruit diseases. *23rd Telecommunications forum TELFOR 2015*.

Mulligan, G. (2015, July 31). *Farmers, techies, entrepreneurs: the story of the FarmDrive girls*. Retrieved from disrupt-africa.com: http://disrupt-africa.com/2015/07/farmers-techies-entrepreneurs-the-story-of-the-farmdrive-girls/

Muthoni, J. J., Shimelis, H., & Melis, R. (2013). Potato Production in Kenya: Farming Systems and Production Constraints. *Journal of Agricultural Science, 5*, 182. doi:http://dx.doi.org/10.5539/jas.v5n5p182

Muthukannan, K., Latha, P., Pon , R. S., & Nisha, P. (2015, March). CLASSIFICATION OF DISEASED PLANT LEAVES USING NEURAL NETWORK ALGORITHMS. *ARPN Journal of Engineering and Applied Sciences, 10*(4).

Namanda, S., Olanya, O. M., Adipala, E., Hakiza, J. J., El-Bedewy, R., Baghsari, A. S., & Ewell, P. (2004, November 23). Fungicide application and host-resistance for potato late blight management: benefits assessment from on-farm. *Crop Protection*, 1075-1083.

Nyankanga, R., Wien, H., Olanya, O., & Ojiambo, P. (2004). Farmers' cultural practicesand management of potato late blight in Kenya highlands: implications for development of integrated disease management. *International Journal of Pest Management*(50), 135-144.

Olanya, O. E., Ojiambo, R., Ewell, P., & Hakiza, J. (2001). *Relationships of fungicide application to late blight development and potato growth parameters in the tropical highlands of Uganda and Kenya.* Lima,Peru: International Potato Centre (CIP).

Onditi, J. O., Nderitu, S. W., Landeo, J. A., Abong', G. O., Sikinyi, E. O., & Kabira, J. N. (2012). Release of three improved varieties for the expanded potato market in Kenya. *AGRICULTURE AND BIOLOGY JOURNAL OF NORTH AMERICA*, 192-197.

Otipa, 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. Retrieved November 20, 2016, from https://www.plantwise.org/Uploads/Plantwise/Kenya%20Poster%20Web.pdf

Ousmane, , B. M., & Collins, T. J. (2016). *Regional Developments; Africa. Chapter 9. 2016 Global Food Policy Report.* IFPRI.

Republic of Kenya. (2005). *Economic Survey 2005.* Nairobi: Government Printer.

Riungu, C. (2011). No easy walk for potatoes. Horticultural News. *The East African Fresh Produce Journal*, 19,16-17.

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

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

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

Shahbaz, T. S., & Safdar, A. (2016, February 22). Introduction, History and Importance of Plant Pathology. Retrieved November 17, 2016, from http://www.slideshare.net/NiamatAzam/introduction-to-plantpathology

Shepherd, K. D. (2010). Soil spectral diagnostics – infrared, x-ray and laser diffraction spectroscopy for rapid soil characterization in the Africa Soil Information Service. *19th World Congress of Soil Science, Soil Solutions for a Changing World.* Brisbane, Australia.

Singh, K. (2007). *Quantitative Social Research Methods.* New Delhi: Sage Publications India Pvt Ltd.

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.

Talla, D., Fon, D., & Fontem , D. (2011). Economic analysis of fungicide and fertilizer applications on huckleberry (Solanum scabrum Mill.) fresh shoot yield. *Journal of Tropical Agriculture*, 58-63.

Temperton, J. (2016, January). *The peer-to-peer SMS network helping farmers connect*. Retrieved from http://www.wired.co.uk: http://www.wired.co.uk/article/wefarm-sms-farmers-peer-to-peer-network

Thorpe, W. M., Omare, A., Owango, M., & Staal, S. (2000). *Development of smallholder dairying in Eastern Africa with particular reference to Kenya. Paper prepared for the UZ/RVAU/DIAS/DANIDA -ENRECA PROJECT REVIEW WORKSHOP.* Harare,Zimbabwe.

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.arccjournals.com/uploads/articles/R3421.pdf

Vidita , T., Jignesh, P., & Chetan , B. (2013). Weather based plant diseases forecasting using fuzzy logic. *Nirma University International Conference on Engineering (NUiCONE)* (pp. 1-5). IEEE.

Appendix

Appendix A: Originality Report

Turnitin Originality Report

Final submission by Patrick Toroitich Kiplimo

From 2016 Plagiarism Check (GS) (Library Services Plagiarism Checker (2016+))

Processed on 06-Apr-2017 2:15 PM EAT
ID: 791292983
Word Count: 18257

Similarity Index	Similarity by Source	
23%	Internet Sources:	17%
	Publications:	11%
	Student Papers:	12%

sources:

1 1% match (student papers from 27-May-2015)
Submitted to Strathmore University on 2015-05-27

2 1% match (Internet from 25-Mar-2016)
http://www.innspub.net/wp-content/uploads/2014/10/IJAAR-V5No4-p12-19.pdf

3 < 1% match (Internet from 29-May-2016)
http://web2.utc.edu/~fgg366/4270/4270Notes/XLMiner/XLMinerUserGuide.pdf

4 < 1% match (Internet from 29-Mar-2017)
http://www.arccjournals.com/uploads/articles/R3421.pdf

Appendix B: Interview Guide

Introduction

Dear Respondent,

This interview questionnaire is part of study conducted by Toroitich Patrick Kiplimo as part of the requirements for the award of a degree of Master of Science in Information Technology at Strathmore University. The main objective of this research is to come up with a model for early detection of potato blight disease. The knowledge that we will gain from your responses will aid in the development of the application. The information requested will be used for academic purposes only and will be treated in strict confidence.

Kind Regards,

Patrick Kiplimo

Real Time Monitoring model for early detection of Crop Detection

i. What are the common diseases that affecting potato?

ii. What are the methods used to identify potato diseases?

iii. What are the factors that influence potato blight development?

iv. What are the symptoms used in monitoring potato late blight?

v. Do extension services provided by extension officers efficient in identifying crop diseases?

vi. Which potato varieties are grown in Nakuru County?

vii. What challenges are faced by the extension workers and farmers in identifying the diseases affecting various crops?

viii. Would use of sensors mobile application make the work of identifying diseases easier for the extension worker?

Your assistance will be highly appreciate

APPENDIX C: Interview Feedback

Appendix C: Interview Feedback

Interview Guide

Real-Time Monitoring Model for Early Detection of Crop Diseases

i. **What are the common diseases that affecting potato?**
   - Bacterial wilt
   - Potato late blight
   - Soft rot
   - Early blight

ii. **What are the methods used to identify potato diseases?**
   Check leaf and stem for lesion formation and discoloration. Check tuber for pale brown Rings

iii. **What are the factors that influence potato blight development?**
   - Weather conditions heavily contribute to spread and development of late blight. High humidity and warm conditions create a good environment for sporangia development.
   - Poor farming practices.
   - Recycling of tuber seeds

iv. **What are the symptoms used in monitoring potato late blight?**
   One should inspect the farm for dark lesions form on leaf and whitish substance underneath the leaf during the early morning hours.

v. **Do extension services provided by extension officers efficient in identifying crop diseases?**
   Yes they do assist as officers are able to assist farmers in identifying disease and giving treatment recommendations.

vi. **Which potato varieties are grown in Nakuru County?**
   - Kenya Mpya
   - Asante
   - Tigoni

vii. **What challenges are faced by the extension workers and farmers in identifying the diseases affecting various crops?**
   - Reaching farmers in remote locations is difficult due to poor road network
   - Lack of technology to assist in early detection of crop diseases

- The extension workers are too few to serve farmers

viii. **Would use of sensors mobile application make the work of identifying diseases easier for the extension worker?**

Yes the technology would greatly assist in monitoring crops and also assist farmers with knowledge of crop diseases.

Appendix D: Arduino Test Case

Title	GSM FUNCTIONALITY TEST
Pre-Conditions	
5 Volts, Antennae, GSM Module, Arduino-UNO board, Working SIM-Card	
Test Scenario	
1.	Start power
2.	Observe Power LED
3.	Observe GSM LED
Expected Results	
1.	Power LED should light immediately
2.	GSM LED should blink fast
3.	In maximum of 2 minutes, the rate of blinking should slow down

Appendix E: API Route

Rest API – Push data from sensor to server
Sensor ID: 4rps21

API Key: ee13504e4b14163cb32331dfd013f3af

Method: GET

URL: http://41.215.34.154/api/data/push

Parameters

`id` (sensor ID)

`t` (temperature reading in °C)

`h` (humidity in %)

`api_key` (unique API key)

`response_type` (optional) (empty or *basic*)

Example request

http://41.215.34.154/api/data/push?id=4rps21&t=30.15&h=41.25&api_key=ee13504e4b14163cb32331dfd013f3af

Example responses

{"status":"success","read_every":15}

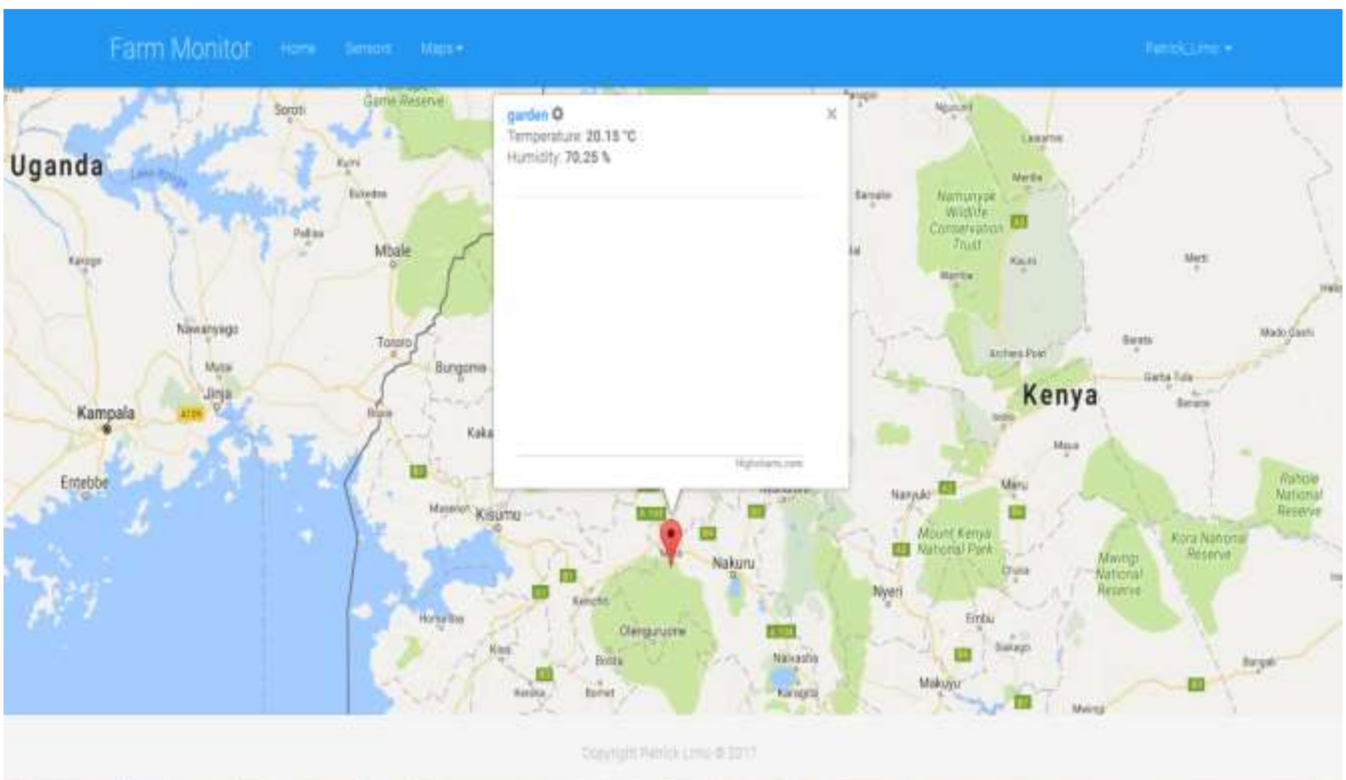
{"status":"error","reason":"Invalid temperature or humidity readings. Must be numeric value."}

{"status":"error","reason":"Invalid API key."}

{"status":"error","reason":"Invalid sensor unique ID."}

With &response_type=basic server returns plain text with either success or fail.

Appendix F: Sample Screen shot



Appendix G: Simcast Algorithm

```
###########################################################################
####
# title          : 01 - SimcCast_Blight_Units.R;
# purpose        : recreate SimCast model as implemented by:
#                  Grünwald, N. J., Montes, G. R., Saldana, H. L.,
Covarrubias,
#                  O. A. R., & Fry, W. E. (2002).
#                  Potato Late Blight Management in the Toluca Valley:
Field
#                  Validation of SimCast Modified for
#                  Cultivars with High Field Resistance. Plant Disease, 86,
#                  1163–1168.;
# producer       : prepared by A. Sparks and K. Garrett;
# last update    : in Toowoomba, Qld, Australia, Jun 2016;
# inputs         : HUSWO weather data to calculate blight units for each
weather
#                   station in the dataset;
# outputs        : blight unit values, averaged hourly weather data to
daily for
#                  each weather station;
# remarks 1      : cultivar resistance values are changed on lines 28-30;
# remarks 2      : this model does not attempt to recreate the entire
SimCast
#                  model, only the blight unit portion
#                  the fungicide unit portion is not a part of this script
and
#                  was not written in R;
# remarks 3      : blight units calculated using this script are used in
the
#                  creation of the SimCastMeta model;
# license        : GPL2;
###########################################################################
####

# Select the resistance level that should be run

#resistance = "S"
#resistance = "MS"
resistance = "R"

# Load libraries ----------------------------------------------------------
------
require("readr")

ConsR <- NULL
DayR <- NULL
blightR <- NULL

# Run this function to generate blight unit calculations for the HUSWO
data set
DailyBlightUnitFiles <- function() {
  files <- list.files("Data/HUSWO", pattern = ".txt$", full.names =  TRUE)
```

```
  for (i in files) {
    weather_data <- as.data.frame(read_tsv(i))
    weather_data[, 6:7] <- round(weather_data[, 6:7], 1)
    colnames(weather_data) <-
      c("stationID",
        "year",
        "month",
        "day",
        "hour",
        "temperature",
        "relativeHumidity")
    blight_calcs <-
      DayR(weather_data = weather_data,
           max_year = 2016,
           min_year = 2000)
    Date <-
      paste(blight_calcs$oYear,
            blight_calcs$oMonth,
            blight_calcs$oDay,
            sep = "-")
    weather_data <- cbind(Date, blight_calcs)
    weather_data <- subset(weather_data, oYear >= 1)
    if (resistance == "S") {
      resistance <- "susceptible"
    } else if (resistance == "MS") {
      resistance <- "moderate"
    } else
      resistance <- "resistant"
    filename <- paste0(basename(i), resistance, "_dayR.txt")
    write_tsv(
      weather_data,
      path = paste0("Cache/Blight Units/", filename),
      col_names = FALSE,
      append = FALSE
    )
  }
}

DayR <- function(weather_data, min_year, max_year) {
  ##Take all weather data and output blight units for a selected range of
months.
  ##This assumes that functions blightR and ConsR are available
  nYear <- max_year - min_year + 1
  oStation <-
    0 * (1:(366 * nYear)) ##longer than 365 days to account for leap years
  oYear <- oStation
  oMonth <- oStation
  oDay <- oStation
  oBlight <- oStation
  oC <- oStation
  oRH <- oStation
  uStation <- unique(weather_data$stationID)
  tC <- 0 * (1:24)
  tRH <- 0 * (1:24)
```

```
globalIndex  <-  1
for (iStation in (uStation)) {
  tStation = subset(weather_data, stationID == iStation)
  for (iYear in (min_year:max_year)) {
    tYear = subset(tStation, year == iYear)
    umonth  =  unique(tYear$month)
    for (i_month in (umonth)) {
      t_month <- subset(tYear, month == i_month)
      rh_test <- sum(t_month$relativeHumidity == 999)
      t_test <- sum(t_month$temperature == 999.9)

      if (rh_test + t_test == 0) {
        if (i_month  ==  4 |
            i_month  ==  6 | i_month  ==  9 | i_month == 11)
        {
          maxDay = 30
        } else
          if (i_month  ==  2) {
            if (iYear / 4 == round(iYear / 4)) {
              maxDay = 29
            } else {
              maxDay = 28
            }
          } else {
            maxDay = 31
          }
        for (iDay in (1:maxDay)) {
          if (iYear < max_year | i_month < 12 | iDay < 31) {
            # Hours are 13:00 - 23:00, 0:00 - 12:00
            # Note that the last day of the last year is excluded
            tDay <- subset(t_month, day == iDay)
            tC[1:12] <- tDay$temperature[13:24]
            tRH[1:12] <- tDay$relativeHumidity[13:24]
            if (iDay < maxDay) {
              tDay2 <- subset(t_month, day == iDay + 1)
            }
            # 100201 - Was "else if"????  Incorrect???
            if (i_month < 12) {
              t_month2 <- subset(tYear, month == i_month + 1)
              tDay2 <- subset(t_month2, day == 1)
            }

            tC[13:24] <- tDay2$temperature[1:12]
            tRH[13:24] <- tDay2$relativeHumidity[1:12]
            out1 <- ConsR(tC  =  tC, tRH  =  tRH)
            out2 <-
              blightR(consmc = out1$consmc,
                      tcons = out1$tcons,
                      resistance = resistance)
            blight <- sum(out2)
            oStation[globalIndex] <- iStation
            oYear[globalIndex] <- iYear
            oMonth[globalIndex] <- i_month
            oDay[globalIndex] <- iDay
```

```
              oBlight[globalIndex] <- blight
              oC[globalIndex] <- mean(tC)
              oRH[globalIndex] <- mean(tRH)
              globalIndex <- globalIndex + 1
            }
          }
        }
      }
    }
  }
  outDaily <-
    data.frame(oStation, oYear, oMonth, oDay, oC, oRH, oBlight)
  return(outDaily)
}


ConsR <- function(tRH, tC) {
  # This outputs tcons and consmc as a list.
  # To get each assign out1 = ConsR(tRH, tC) then out1$tcons and
out1$consmc
  # tRH  =  Temporary Relative Humidity
  # tC  =  Temporary *C
  consmc <- 0 * (1:12) - 99
  first <- TRUE
  tcons <- 0 * (1:12)
  cons_index <- 1
  tttemp <- (-99)

  for (j in (1:24)) {
    if (!is.na(tRH[j]) & tRH[j] >= 90) {
      tcons[cons_index] <- tcons[cons_index] + 1
      if (first) {
        tttemp <- tC[j]
      }

      else{
        tttemp <- c(tttemp, tC[j])
      }
      first <- FALSE

      if ((tRH[j + 1] < 90 & j < 24) | j  ==  24) {
        consmc[cons_index] <- mean(tttemp)
        cons_index <- cons_index + 1
        tttemp <- (-99)
        first <- TRUE
      }
    }
  }
  cons_out <- data.frame(tcons, consmc)
  return(cons_out)
}

## Blight Unit Calculation
blightR <- function(consmc, tcons, resistance) {
```

```
blight_unit = 0 * (1:12)
if (resistance == "S") {
  for (k in (1:12)) {
    if (consmc[k] <= 27 & consmc[k] >= 3) {
      if (consmc[k] >= 23 & consmc[k] <= 27) {
        if (tcons[k] >= 7 & tcons[k] <= 9) {
          blight_unit[k] = 1
        } else
          if (tcons[k] >= 10 &
              tcons[k] <= 12) {
            blight_unit[k] = 2
          } else
            if (tcons[k] >= 13 &
                tcons[k] <= 15) {
              blight_unit[k] = 3
            } else
              if (tcons[k] >= 19 &
                  tcons[k] <= 24) {
                blight_unit[k] = 5
              }
      }
      if (consmc[k] >= 13 & consmc[k] <= 22) {
        if (tcons[k] >= 7 & tcons[k] <= 9) {
          blight_unit[k] = 5
        } else
          if (tcons[k] >= 10 &
              tcons[k] <= 12) {
            blight_unit[k] = 6
          } else
            if (tcons[k] >= 13 &
                tcons[k] <= 24) {
              blight_unit[k] = 7
            }
      }

      if (consmc[k] >= 8 & consmc[k] <= 12) {
        if (tcons[k] == 7) {
          blight_unit[k] = 1
        } else
          if (tcons[k] >= 8 &
              tcons[k] <= 9) {
            blight_unit[k] = 2
          } else
            if (tcons[k] == 10) {
              blight_unit[k] = 3
            } else
              if (tcons[k] >= 11 &
                  tcons[k] <= 12) {
                blight_unit[k] = 4
              } else
                if (tcons[k] >= 13 &
                    tcons[k] <= 15) {
                  blight_unit[k] = 5
                } else
```

```
                  if (tcons[k] >= 16 &
                      tcons[k] <= 24) {
                    blight_unit[k] = 6
                  }
      }

      if (consmc[k] >= 3 & consmc[k] <= 7) {
        if (tcons[k] >= 10 & tcons[k] <= 12) {
          blight_unit[k] = 1
        } else
          if (tcons[k] >= 13 &
              tcons[k] <= 15) {
            blight_unit[k] = 2
          } else
            if (tcons[k] >= 16 &
                tcons[k] <= 18) {
              blight_unit[k] = 3
            } else
              if (tcons[k] >= 19 &
                  tcons[k] <= 24) {
                blight_unit[k] = 4
              }
      }
    }
  }

}

if (resistance  ==  "MS") {
  for (k in (1:12)) {
    if (consmc[k] <= 27 & consmc[k] >= 3) {
      if (consmc[k] >= 23 & consmc[k] <= 27) {
        if (tcons[k] >= 10 & tcons[k] <= 18) {
          blight_unit[k] = 1
        } else
          if (tcons[k] >= 19 &
              tcons[k] <= 24) {
            blight_unit[k] = 2
          }
      }

      if (consmc[k] >= 13 & consmc[k] <= 22) {
        if (tcons[k] == 7) {
          blight_unit[k] = 1
        } else
          if (tcons[k] == 8) {
            blight_unit[k] = 2
          } else
            if (tcons[k] == 9) {
              blight_unit[k] = 3
            } else
              if (tcons[k] == 10) {
                blight_unit[k] = 4
              } else
```

```
              if (tcons[k] >= 11 &
                    tcons[k] <= 12) {
                blight_unit[k] = 5
              } else
                if (tcons[k] >= 13 &
                      tcons[k] <= 24) {
                  blight_unit[k] = 6
                }
      }

      if (consmc[k] >= 8 & consmc[k] <= 12) {
        if (tcons[k] >= 7 & tcons[k] <= 9) {
          blight_unit[k] = 1
        } else
          if (tcons[k] >= 10 &
                tcons[k] <= 12) {
            blight_unit[k] = 2
          } else
            if (tcons[k] == 13 &
                  tcons[k] <= 15) {
              blight_unit[k] = 3
            } else
              if (tcons[k] >= 16 &
                    tcons[k] <= 18) {
                blight_unit[k] = 4
              } else
                if (tcons[k] >= 19 &
                      tcons[k] <= 24) {
                  blight_unit[k] = 5
                }
      }

      if (consmc[k] >= 3 & consmc[k] <= 7) {
        if (tcons[k] >= 13 & tcons[k] <= 24) {
          blight_unit[k] = 1
        }
      }
    }
  }
} else if (resistance  ==  "R") {
  for (k in (1:12)) {
    if (consmc[k] <= 27 & consmc[k] >= 3) {
      if (consmc[k] >= 23 & consmc[k] <= 27) {
        if (tcons[k] >= 14 & tcons[k] <= 16) {
          blight_unit[k] = 1
        }
      }

      if (consmc[k] >= 13 & consmc[k] <= 22) {
        if (tcons[k] == 7) {
          blight_unit[k] = 1
        } else
          if (tcons[k] == 8) {
            blight_unit[k] = 2
```

```
          } else
            if (tcons[k] == 9) {
              blight_unit[k] = 3
            } else
              if (tcons[k] >= 10 &
                  tcons[k] <= 12) {
                blight_unit[k] = 4
              } else
                if (tcons[k] >= 13 &
                    tcons[k] <= 24) {
                  blight_unit[k] = 5
                }
      }

      if (consmc[k] >= 8 & consmc[k] <= 12) {
        if (tcons[k] >= 10 & tcons[k] <= 12) {
          blight_unit[k] = 1
        } else
          if (tcons[k] >= 13 &
              tcons[k] <= 15) {
            blight_unit[k] = 2
          } else
            if (tcons[k] >= 16 &
                tcons[k] <= 24) {
              blight_unit[k] = 3
            }
      }

      if (consmc[k] >= 3 & consmc[k] <= 7) {
        if (tcons[k] >= 19 & tcons[k] <= 24) {
          blight_unit[k] = 1
        }
      }
    }
  }
 }
 return(blight_unit)
}

# eos
```