University of Nairobi

School of Mathematics

ISSN: 2410-1397

Master Project in Mathematics

Predicting Potato Late Blight Disease Pressure using Artificial Neural Network

Research Report in Mathematics, Number 41, 2019

Benson Kisinga

July 2019

Submitted to the School of Mathematics in partial fulfilment for a degree in Master of Science in Biometry

Master Project in Mathematics **University of Nairobi**
July 2019

Predicting Potato Late Blight Disease Pressure using Artificial Neural Network

Research Report in Mathematics, Number 41, 2019

Benson Kisinga

School of Mathematics
College of Biological and Physical sciences
Chiromo, off Riverside Drive
30197-00100 Nairobi, Kenya

Master Project

Submitted to the School of Mathematics in partial fulfilment for a degree in Master of Science in Biometry

Submitted to: The Graduate School, University of Nairobi, Kenya

Abstract

Late blight causing *Phytopthora infestans* (Mont.) de Bary largely depends on weather parameters such as temperature and relative humidity for survival, spread and its ability to attach and infect new plants. The variations in weather across different agro-ecological zones can be used to explain the different levels of disease severity experienced across these regions.

Potato late blight disease evaluation data from five locations was coupled with GIS-linked weather data. A series of neural network models were developed and validated with 10-fold cross validation and the optimal model selected based on accuracy achieved on validation set. The selected model had 1 hidden layer with 14 nodes achieving an accuracy of 88% in the validation set. The final model was used to predict disease severity with 89% accuracy on new data. It was also found that the number of precipitation days and number of days with temperature and relative humidity favorable to disease development were amongst the top significant variables in the model hence a target for monitoring.

This model can be used to estimate the expected late blight severity in a target region hence support the decisions on the appropriate varieties and management regimes to be used, reducing yield loss and excessive use of fungicides.

Master Project in Biometry at the University of Nairobi, Kenya.
ISSN 2410-1397: Research Report in Biometry

DISTRIBUTOR: School of Mathematics, University of Nairobi, Kenya

Declaration and Approval

I the undersigned declare that this dissertation is my original work and to the best of my knowledge, it has not been submitted in support of an award of a degree in any other university or institution of learning.

Signature Date

Benson Kisinga
Reg No. I56/8489/2017

In my capacity as a supervisor of the candidate's dissertation, I certify that this dissertation has my approval for submission.

Signature Date

Dr John Ndiritu
School of Mathematics,
University of Nairobi,
Box 30197, 00100 Nairobi, Kenya.
E-mail: jndiritu@uonbi.ac.ke

Dedication

I dedicate this project to my wife Freddah, son Aiden, parents Sarah and Charles for their endless support and patience.

May they all live to love, cherish and appreciate the power of education.

Contents

Figures and Tables

Figures

Tables

Acknowledgments

I wish to thank the International Potato Center (CIP) for the opportunity and support granted to me to pursue my masters degree. My sincere appreciation to Dr. Monica Parker for the love, guidance and patience through out my studies.

Special thanks to my very supporting supervisor Dr. Ndiritu for the assistance, ideas and guidance offered all through.

Thanks to my colleagues Elly, Thiago, Kalpana and Nyawade. I appreciate your inputs and support and time you gave me during my studies.

Thanks to my family for the encouragement, inspiration , prayers and patience all through. Thanks for making it possible to balance my work, study and family life.

Benson Kisinga

Nairobi, 2019.

1 Introduction

1.1 Background

Potato diseases have been identified as the major constraint affecting production leading to huge yield losses and high management cost. In Kenya, potato late blight is the second major potato disease with over 67% prevalence after bacterial wilt (Kaguongo et al. 2010), causing 30 to 75% yield losses and even up to 100% loss in susceptible varieties (Mariita, Nyangeri, and Makatiani 2016). Potato is an important crop in Kenya with the total area estimated at 192,341 hectares and an average yield of 7.9 tonnes per hectare in 2017 according to FAOSTAT, 2017. Most of the production is done by small-scale farmers concentrated on the highland areas (1500 - 3000 masl) under rain fed conditions (Muthoni, Nyamongo, and Mbiyu 2017). The weather patterns vary in different areas influencing the crop productivity and disease incidences. Figure (1.1) and Table (1.1) show the agro-climatic zones of Kenya with their respective characteristics. These zones differ in amounts received for precipitation, temperature and relative humidity, hence introducing variations in the disease risk.

Crop disease prediction models are based on the interaction between a susceptible host plant and a pathogen under favorable weather conditions; constituting the disease triangle. Potato late blight disease is caused by *Phytophthora infestans* (Mont.) de Bary and leads to serious production losses through premature defoliation of the potato plants and further during storage by infected tubers. Late blight is a polycyclic disease whose development is depend on the host plant, environment, the pathogen and the management measures applied to the exposed fields. The pathogen depends on weather factors for survival, spread and ability to attach to and infect more plants. Shifts in weather conditions may lead to favourable conditions for the disease development hence greatly affect production. The continuous cultivation of the crop in the growing areas ensures abundant pathogen to cause late blight through out the year (Muchiri et al. 2009), posing a challenge in the disease management which is mainly by use of fungicides (Nyankanga et al. 2004).

A system is required to link the weather data with the disease pressure and severity for particular region, so that decisions can be made on the selection of cultivars and application of management technologies to ensure optimal crop performance with reduced cost and negative environmental impact due to excessive use of fungicides. A disease model that provides a magnitude of the disease pressure for a given agro-ecological zone under the existing weather conditions would be a great tool in deciding on which varieties to expose to various disease levels for resistance evaluation and also for selection of the best suited varieties for a site. Therefore, a successful disease forecasting scheme should be reliable, simple for the target users and should lower production cost compared to the existing crop management practices.

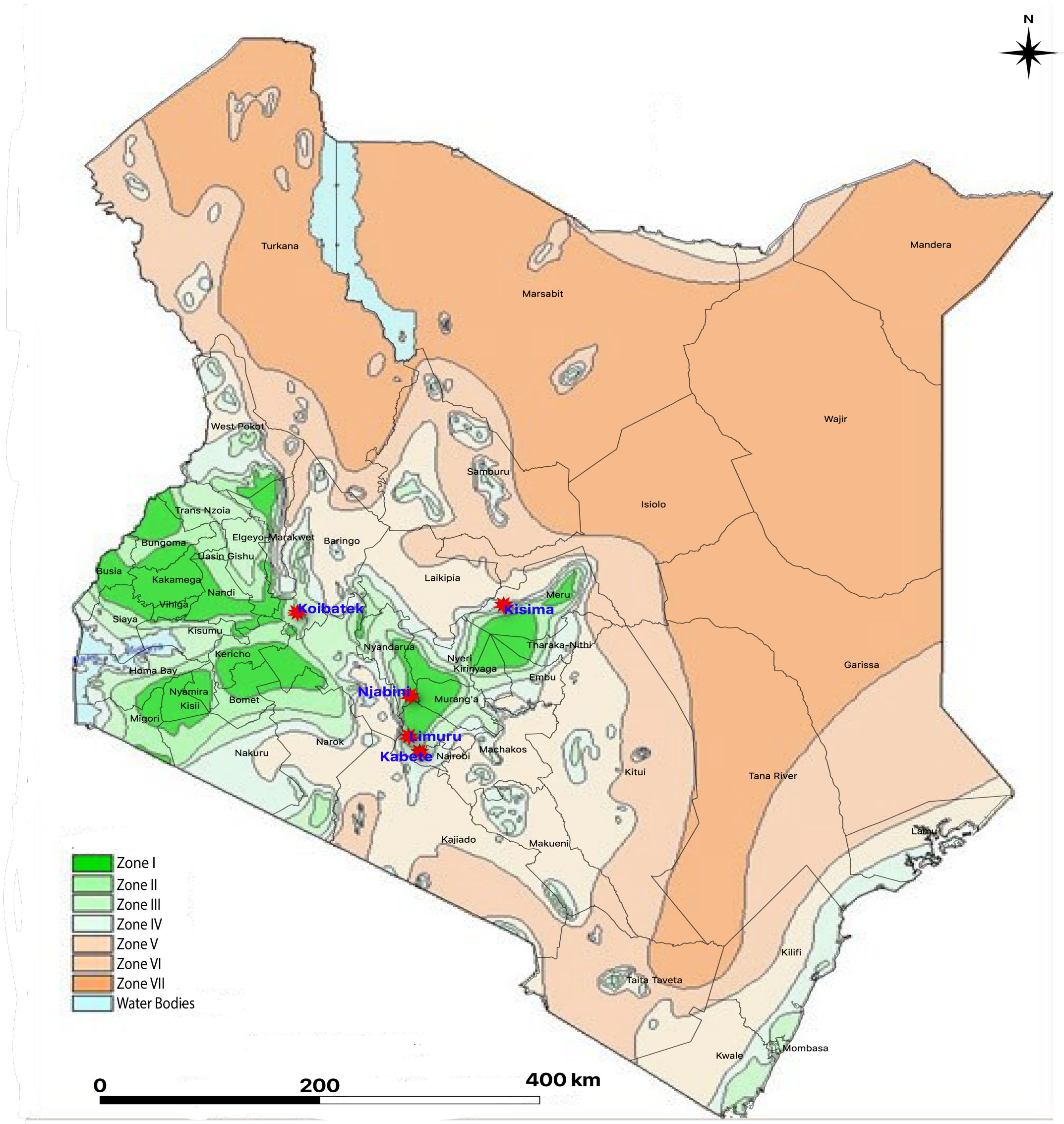

Figure 1.1. Agro-climatic zones in Kenya and the selected study sites

Table 1.1. Characteristics of agro-climatic zones in Kenya

Agro-Climatic Zones	Classification	Moisture Index (%)	Annual rainfall (mm)	Land Area (%)
Zone I	Humid	>80	1100-2700	2
Zone II	Sub-humid	65 - 80	1000-1600	5
Zone III	Semi humid	50 - 65	800-1400	5
Zone IV	Semi humid-Semi Arid	40 - 50	600-1100	5
Zone V	Semi Arid	25 - 40	450-900	15
Zone VI	Arid	15 - 25	300-550	22
Zone VII	Very Arid		150-350	46

1.1.1 Potato late blight disease development cycle

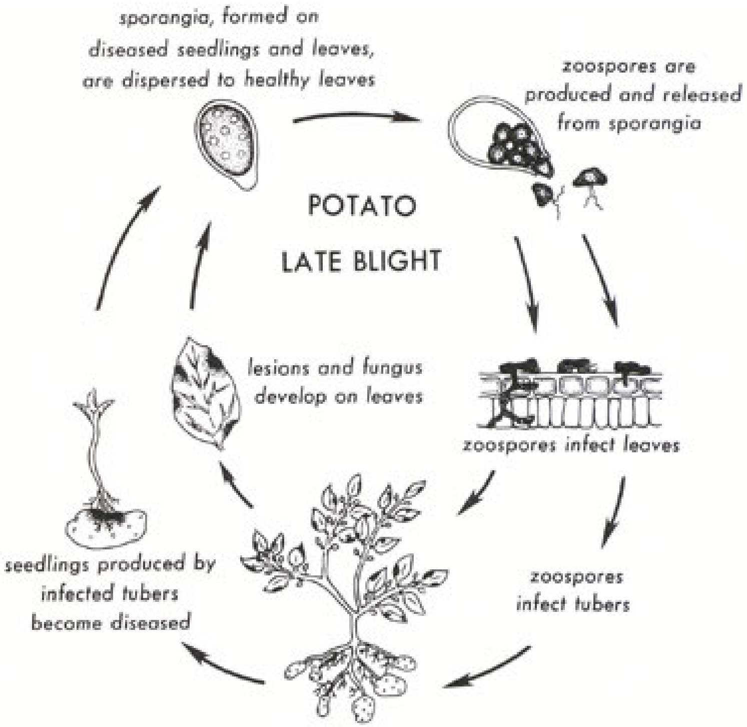
sporangia, formed on diseased seedlings and leaves, are dispersed to healthy leaves
zoospores are produced and released from sporangia
POTATO LATE BLIGHT
lesions and fungus develop on leaves
zoospores infect leaves
seedlings produced by infected tubers become diseased
zoospores infect tubers

Figure 1.2. Late blight development cycle

Sporangia from volunteer plants, infected seed or piles of discarded potatoes are carried by wind to the leaf surfaces. The mycelium starts to grow when temperature and moisture are favorable, invading and killing the plant cells. Symptoms appear three to seven days post infection as lesions. The fungus grow sporangiophores which produce more sporangia starting a new cycle of infection. Tubers can be infected when rain water washes down the sporangia and zoospores from the leafs into the soil (Figure 1.2). Weather parameters have been found to influence the pathogen's dispersal capability and the ability to attach to new host plants.

With the emergence of smart farming concept, there is a continued development of high precision algorithms that help make agriculture more efficient and more effective. Several algorithms have been employed in the area of crop management assisting in yield prediction, crop quality control, weed detection and disease management (Yang and Guo 2017). Artificial neural networks have found a wide scope of applications in precision agriculture ranging from acquisition of meteorological data, disease forecasting, and yield prediction - tasks that involve huge calculations and require quick execution. Neural networks are inspired by the human brain functionality and can detect complex patterns and relationships between different data structures and extract high dimensional interactions. They can be used to solve both classification and regression problems with a high degree of accuracy and performance.

A multilayer perceptron neural network consists of neurons arranged in successive layers from input to output layer through the hidden layer(s). It can therefore be interpreted as a simple input-output model, with weights and biases as the free parameters in the model. During the training phase, the model compares its own output to the target output and tries to minimize the difference through a learning algorithm.

1.2 Problem statement

Late blight leads to huge losses in terms of yield and cost of managing the disease in fields. The spatial variations in climatic conditions can be related to the disease risk to provide an insight on the right choice of varieties to expose to the different risk areas. This will reduce losses associated with the disease and also the environmental degradation caused by overuse of fungicides. The proposed study focuses on the use of artificial neural networks which is a generalization of logistic regression able to detect any possible interactions between the predictor variables to predict late blight disease response in unobserved environments.

1.3 Objectives

1.3.1 General Objectives

The main purpose of this study is to use multi-environment late blight data alongside weather data to develop a model for predicting late blight disease pressure in unobserved environments characterized by their weather patterns .The study will use artificial neural network model architecture to model both agronomic and weather data across seasons an environments.

1.3.2 Specific Objectives

1. To aggregate data on existing multi-environment trial data with past geo-linked weather data.
2. To develop a neural network model using crop and weather data.
3. To use the model to predict disease response in unobserved conditions.

1.4 Significance/Justification

Due to time, resources and financial constraints, it makes it hard to evaluate performance of varieties for selection in different environments. This model will enable users to simulate disease risk in new environments and help in deciding the best suited varieties for a particular zone, thus reducing on time and cost of management.

1.5 Limitation

This study assumes that late blight disease is present in all the selected sites and may not apply to cases where the pathogen is not present.

1.6 Outline

The following sections include a detailed review of machine learning models used in disease forecasting and models that have been specifically applied to forecasting of potato late blight disease. The methodology describes procedures used for data acquisition and preparation, a detailed description on artificial neural network model creation and evaluation. The results on data preparation, model selection and evaluation are given and interpreted in the results section. There is a detailed discussion on the obtained results and a conclusion are given in the discussion section. Included is also an appendix on the code used to develop and test the model.

2 Literature Review

The chapter will focus on providing details on the potato late blight disease, machine learning and precision agriculture models that have been used to monitor potato productivity.

2.1 General overview

2.1.1 Review of models used in late blight

Several predictive models have been developed and used to explain the the effects of crop diseases under different climatic conditions (Hijmans, Forbes, and Walker 2000; Crane-Droesch 2018; Rizzo, Conklin, and Dougherty 2003; Sannakki et al. 2013; Gu et al. 2016). Weather parameters such as temperature, intensity and duration of solar radiation, precipitation and relative humidity are common in most of these models as they largely affect the host-pathogen interaction hence determine whether a disease occurs or not.

The pathogen causing potato late blight depends a lot on the prevailing weather conditions as these determine the pathogen's spread and survival. Several models have tried to explain how and which of these parameters should be monitored in order to predict the existence or magnitude of late blight infection.

Initial models were based on a combination of weather parameters which included night time dew, night temperature, average cloudiness and temperature developed by van Everdingen. The model was used to predict late blight disease within 14 days of favorable conditions in Holland (Henderson, Williams, and Miller 2007). The model was later modified by Beaumont (1947) to include temperature-humidity rule in defining the favorable days. Prediction was made 2 days when minimum temperature was above 10^oC and relative humidity above 75% in the UK.

Cook (1949) developed a simulation based on daily mean temperature and rainfall starting at the onset of planting season. Farmers were advised to apply fungicides following a build-up of rainfall and temperature beyond a certain threshold. Hyre classified Cook's days as either favoring or not favoring the disease initiation. Favorable days had the five day moving average temperature below 25.6^oC and for the previous 10 days with a total precipitation of $>$3.0 cm. Smith considers favorable conditions to be 2 consecutive days with 10^oC minimum temperature and over 11hr exposure to relative humidity above 90%.

Wallin (1953) introduced severity values for forecasting initial occurrence and subsequent dispersion of late blight. The severity values are numbers assigned to specific combinations between periods with relative humidity greater than 90% and mean temperature during these periods. Initial occurrence of the disease was predicted within a fortnight after 18-20 severity value accumulation from the day of emergence. BLITECAST (Krause 1976) was developed to combine the Wallin and Hyre models achieving success in application in North America. Farmers could send weather data from potato farms to the forecasting center and receive recommendations. The regions practical to BLITECAST were humid, with high rainfall and frequent yearly blight incidences and using such humid based models may not be applicable to Kenya where there is increased number of unfavorable conditions.

2.1.2 Artificial Neural Network Methods

Shastry, Sanjay, and Deshmukh (2016) has compared the prediction efficiency for ANN and Multiple linear regression (MLR) to predict wheat yield using weather parameters and nature of the soil. ANN was found to have a higher performance than a regular MLR on the test dataset. The study shows that the ANN model was improved by varying the number of perceptrons and hidden layers which increased R^2 at a lower prediction error.

Alves et al. (2017) investigated the prediction capacity of ANN to obtain AUDPC values for tomato late blight using fewer evaluations. Different combinations of the evaluation data points representing the percent leaf area damaged by the pathogen at three days intervals. The ANN architecture was Multilayer perceptron (MLP) with one to three neuron for the input layer, two to sixteen neurons for the hidden layer established iteratively and a single output node. The model used logistic sigmoid and hyperbolic tangent activation functions. The studied model predicted AUDPC values with correlation of 0.94 with two evaluations and 0.97 with three evaluations between the observed and the predicted AUDPC values. This could translate to reduced investment in human, time and economic resources dedicated to screening the fields multiple times.

Vianna, Cunha, and Oliveira (2017) has evaluated a computational approach to early detection of tomato late blight by using a pair of two multi perceptron artificial neural networks to analyze digital images from fields and classify the image pixel. One ANN identifies the healthy leaf areas while the other ANN identifies the damaged areas. The output from the two is combined to produce a final classification with the injured areas highlighted with 97% accuracy. This can be used as an early detection system for large fields and assist in timely application of control measures employed.

Yang and Guo (2017) has reviewed the use of machine learning algorithms in plant disease prediction. The study highlights the use of Naïve Bayes, SVM and ANN in detection, identification and prediction of crop disease and the need to focus more on prediction and quantification models than identification and classification problems as implied by precision agriculture.

Sharma, Singh, and Singh (2018) studied potato late blight prediction from weather variables using ANN while comparing three activation functions for the model. The results shows that a maximum prediction accuracy of 90.9% when using the logistic activation function, outperforming ReLU and hyperbolic tangent activation functions on the test dataset. Taylor et al. (2003) also compared the efficiency of five predictive models for potato late blight with the goal to improve the accuracy of disease alert system. Among the evaluated models, NegFry (Hansen 1995), which combines a negative prognosis model with a weather element based model was the most ideal predictive model for a 10 day alert system.

Maina (2016) and Toroitich (2017) have studied disease forecasting by ANN in Kenya. Maina explored the use of a vision based model to classify maize diseases from images taken from maize fields

3 Methods

The model to predict disease pressure based on weather variables uses data from past years coupled with variety performance data under different environments and treatments.

3.1 Data sources and Data preparation

3.1.1 Agronomic data

Late blight evaluation data was obtained from field experiments performed between years 2009 and 2015 in Kisima, Koibatek, Limuru, Njabini and Kabete regions (Kromann et al. 2012). Two-way factorial strip-plot design was used in each site with late blight treatment and potato genotypes as the factors. Late blight disease occurred naturally in all the experiments. Percentage leaf area damaged by late blight was assessed and recorded from one month after planting until there was 100% infection on control experiments on susceptible genotypes. The disease assessments were used to calculate area under disease progress curve (AUDPC) according to Shaner (1977) :

$$AUDPC = \left[\sum_{i=1}^{n-1} [(y_i + y_{i+1})/2] \times (t_{i+1} - t_i) \right], \tag{3.1}$$

where y_i and y_{i+1} are the percentages of damaged leaf area observed between time t_i and t_{i+1} and n is the total number of evaluations.

Table 3.1. Sample field experiment data showing disease scores for different treatments and varieties across sites and seasons

site	season	treatment	variety	audpc	raudpc	year
UoN	SR	Control	Arka	4998.00	0.89	2010
Kabete	SR	Control	Tigoni	581.00	0.09	2011
Njabini	LR	Phosphonate	Kenya Karibu	91.00	0.02	2010
Limuru	LR	Agrifos	Shangi	1168.85	0.25	2014
Limuru	LR	Fosphite	Asante	1285.09	0.27	2014
UoN	LR	Control	Asante	1074.50	0.16	2011
UoN	SR	Fosphite	Mavuno	2352.00	0.42	2010
Koibatek	SR	Ridomil	Kenya Karibu	857.50	0.27	2010

In order to compare disease incidences across multiple environments, relative Area Under Disease Progress Curve (rAUDPC) was calculated by dividing the AUDPC with the highest value expected, assuming 100% disease incidence (Table 3.1). Mean group clustering was used to define the varieties using the rAUDPC estimates of the control experiments accordin to a hierarchical clustering method proposed by Scott and Knott (1974), where the varieties with the lowest means were taken to be the representatives of the most resistant varieties . Mean grouping was also done on the treatments where the groups with the lowest means were considered the most effective treatments. Disease incidence was split to generate three categories according to the relative AUDPC scores.

3.1.2 Weather data

Past daily weather data on temperature, precipitation, relative humidity and windspeed was obtained from NASA POWER using geo-points for all the sites. The GIS-linked weather data for all the sites was combined and model variables extracted according to the growing seasons. These are factors known to favor disease development during the growing season.

Table 3.2. Weather variables calculated from 2009 to 2015 for 5 sites for construction of a neural network model

Variable (Units)	Description
FDTP (days)	Number of days with 5 day average temperature below 25^oC and a 10 day precipitation total of 30mm or above
FDTRh (days)	Number of days with temperature between 10 and 25^oC and relative humidity of above 80%
AP10d (mm)	Average precipitation for 10 consecutive days during the growing season
PD (days)	Number of days with precipitation of 0.25 mm or higher
PT (mm)	Total Precipitation over the growing season
MRH	Mean relative humidity over growing season
Tmax5	Average Maximum temperature for 5 consecutive days
Tmin5	Average Minimum temperature for 5 consecutive days
Tav5	Average temperature for 5 consecutive days
WIND	Average Wind Speed at 2 meters above the surface

Both datasets were combined to make the working data for model development.

3.2 Model description

3.2.1 Introduction: The logistic function

The logistic/sigmoid function is a probability soft threshold function that maps output to a range between 0 and 1.

$$\sigma(z) = \frac{1}{1+e^{-z}} \tag{3.2}$$

Large negative z values makes the denominator to grow exponentially and consequently $\sigma(z)$ approaches 0. Large positive z values shrink the e^{-z} term to zero hence $\sigma(z)$ approaches 1.

$$\text{Weighted input sum, } z = \sum_{i=0}^{d} w_i x_i \tag{3.3}$$

The hypothesis $h(x) = \sigma(z)$ is interpreted as a probability of the output given x and the signal $z = w^\top x$ referred to as the **risk score**. In order to calculate the error gradient, the derivative of the sigmoid function, $\frac{\partial}{\partial z}\sigma(z)$ is required. This can be obtained as:

$$\begin{aligned}
\sigma'(z) &= \frac{\partial}{\partial z}\left(\frac{1}{1+e^{-z}}\right) \\
&= \frac{e^{-z}}{(1+e^{-z})^2} \\
&= \frac{1+e^{-z}-1}{(1+e^{-z})^2} \\
&= \frac{1+e^{-z}}{(1+e^{-z})^2} - \left(\frac{1}{1+e^{-z}}\right)^2 \\
&= \sigma(z) - \sigma(z)^2 \\
\sigma'(z) &= \sigma(z)(1-\sigma(z))
\end{aligned} \tag{3.4}$$

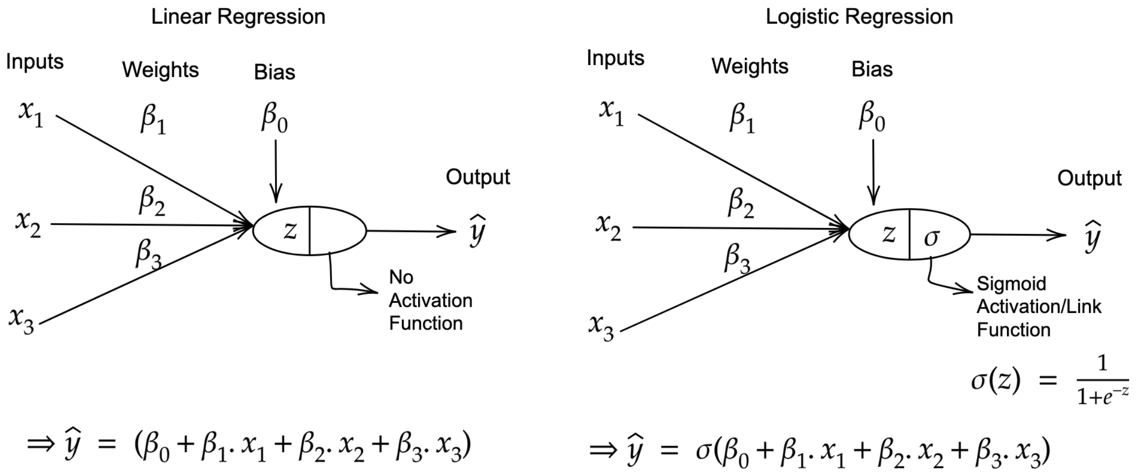

Figure 3.1. A representation of Linear/Logistic regression as a computation node

A logistic regression employs a link function that maps the output to a probability range of 0 and 1. The goal is to minimize the negative log-likelihood of the Bernoulli distribution.

3.2.2 Artificial Neural network

Artificial neural network (ANN) resembles the computational model of the brain with a characteristic node (artificial neuron) and node connectivity network. A neural network can be viewed as a series of many stacked logistic regressions that generate features from the input data and an output layer. The model structure contains a series of highly connected nodes (neurons) with each connection having a different weight.A basic node consists of three elements. A set of inputs $(X_1, X_2 \ldots, X_n)$ characterized by weights which define their contribution to the nodes computation, a summation of all the input signals weighted by their respective signal strengths and an activation function which defines the output of the node.

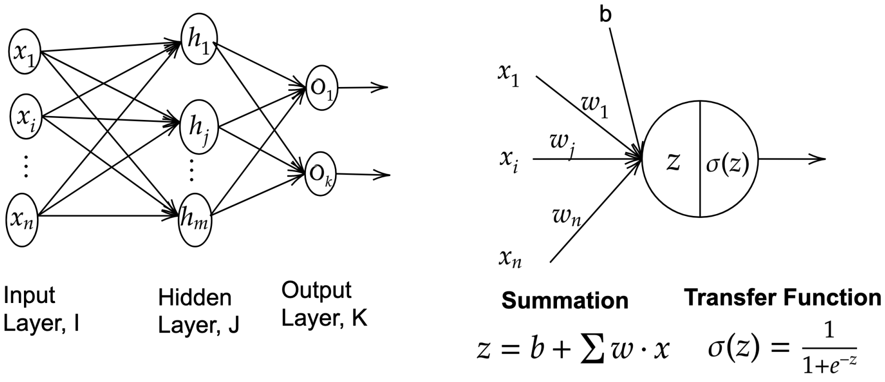

Figure 3.2. A feedforward-back propagation multilayer perceptron neural network

The activation function takes the weighted sum of all the inputs to a particular node, transforms it and generates an output. Given a set of inputs, x_1, x_2, the weighted sum is a linear discriminator given by:

$$f(x_1, x_2) = b + w_1x_1 + w_2x_2 \tag{3.5}$$

$$f(X) = b + \sum_i w_ix_i \tag{3.6}$$

In regression, the predicted output is given while in classification, the predicted class is given by

$$Class = \begin{cases} 1, & \text{if } f(X) > 0 \\ 0 & \text{if } f(X) \leq 0 \end{cases}, \tag{3.7}$$

where w_i is the signal weight for each neuron and b term is the bias which controls how the neuron outputs close to 0 or 1, irrespective of the weights. If n variables are added to equation (3.5), the linear function for the weighted sum becomes

$$f(x) = b + w_1x_1 + w_2x_2 + \cdots + w_nx_n \tag{3.8}$$

In matrix notation,

$$f(x) = b + W^\top X \quad where \quad W = \begin{bmatrix} w_1 \\ w_j \\ \vdots \\ w_m \end{bmatrix} \quad and \quad X = \begin{bmatrix} x_1 \\ x_i \\ \vdots \\ x_n \end{bmatrix} \tag{3.9}$$

Setting the original weighted sum of inputs to a variable z, we have

$$z = b + \sum_i w_ix_i \tag{3.10}$$

and transforming the sum z using the sigmoid activation function $\sigma(z)$ condenses the output to a range between 0 and 1.

A neural network may consist of several hidden layers between the input and the output layers. The hidden layers prevent the model from mapping inputs directly to the output. The effect of the inputs on the model output is highly interdependent and hidden layers enable the model to capture the the fine interactions among the input variables which affect the final model output. Each of the neurons in the hidden layers combine the inputs differently to learn different characteristics of the data and the final model output is a function of these characteristics instead of the raw input values. The neural network model learns by finding a set of approximate weights that can generalize well on new data.

3.2.3 Model Learning Process

Back propagation is a widely used algorithm for training multilayer networks by minimization of the sum of the squared errors using gradient descent methods. The signal is transferred from the input neurons through the hidden layers to the output layer by a feed forward mechanism. Initially, the random weights are assigned to the network to obtain the initial error estimate then the weights are adjusted to obtain the lowest possible error value through gradient descent method which attempts to determine which direction the loss function steeps downwards the most with respect to adjusting the weight parameters. A loss function is defined in order to determine how good the line of best fit is. It is a measure of the distance between the model output $\mathcal{O}$ given a set of inputs X and the actual target value of y. A common loss function is the mean squared error (MSE) which is the average squared difference between the model output and the target y value.

$$E(X, \theta) = MSE = \frac{1}{n}\sum_{i}(y_i - \mathcal{O}_i)^2 \tag{3.11}$$

The gradient of the loss function is then calculated with respect to all the parameters to find the direction in which the function steeps downhill the most using chain rule and product rule in differential calculus. The gradient is a vector of partial derivatives of the loss function with respect to each variable.

3.2.4 Back propagation algorithm

Back propagation is a supervised learning algorithm based on error correction and the minimization of the error function with respect to the weights of the neural network. The error function can be decomposed into a sum of all the error terms for all input-output pairs, thus the derivatives can be calculated for individual terms and then summed at the end - the derivative of a sum of functions is the sum of the derivatives of each function. The derivation begins by applying chain rule to the error function and then steps back into the hidden layers.

Notation

- x_j : Input to node j
- W_{jk} : Connection weight from node j to node k
- $\sigma(z) = \frac{1}{1+e^{-z}}$: The sigmoud activation/transfer function
- $\mathcal{O}_j$: Output of node j
- y_j : Target value of node j in the final layer

The output layer: Given a set of input variables x_1 to x_n, the output of node k the final layer $\mathcal{O}_k$ and the respective observed levels for the dataset, y_k, the error can be expressed as:

$$E = \frac{1}{2}\sum_{k \in K}(\mathcal{O}_k - y_k)^2 \tag{3.12}$$

If E is the error obtained from a single network iteration, we need to calculate the rate of change of E with respect to connectivity weights, $\frac{\partial E}{\partial W_{jk}}$, so that it can be minimized. For the nodes in the final output layer;

$$\frac{\partial E}{\partial W_{jk}} = (\mathcal{O}_k - y_k)\frac{\partial}{\partial W_{jk}}\mathcal{O}_k \tag{3.13}$$

The output at $\mathcal{O}_k$ is obtained by passing the weighted sum of inputs to the output node k to the transfer function $\sigma(z_k)$.

$$\frac{\partial E}{\partial W_{jk}} = (\mathcal{O}_k - y_k)\frac{\partial}{\partial W_{jk}}\sigma(z_k) \tag{3.14}$$

Substituting $\sigma(z_k)$ with its derivative in Equation 3.4 above and applying the chain rule we get

$$\frac{\partial E}{\partial W_{jk}} = (\mathcal{O}_k - y_k)\sigma(z_k)(1-\sigma(z_k))\frac{\partial}{\partial W_{jk}}z_k \tag{3.15}$$

which can be simplified to

$$\frac{\partial E}{\partial W_{jk}} = (\mathcal{O}_k - y_k)\mathcal{O}_k(1-\mathcal{O}_k)\mathcal{O}_j \tag{3.16}$$

since the derivative of the weighted sum z_k with respect to the input weight W_{jk} is the output of the node j, $\mathcal{O}_j$. We can define δ_k to be the expression $(\mathcal{O}_k - y_k)\mathcal{O}_k(1-\mathcal{O}_k)$ involving k and rewrite the error derivative as

$$\frac{\partial E}{\partial W_{jk}} = \mathcal{O}_j\delta_k \tag{3.17}$$

where $\delta_k = (\mathcal{O}_k - y_k)\mathcal{O}_k(1-\mathcal{O}_k)$.

The hidden layer: We compute the derivative of the error with respect to the weights in the hidden layer as follows;-

$$\frac{\partial E}{\partial W_{ij}} = \frac{\partial}{\partial W_{ij}}\frac{1}{2}\sum_{k\in K}(\mathcal{O}_k - y_k)^2 \tag{3.18}$$

$$\frac{\partial E}{\partial W_{ij}} = \sum_{k\in K}(\mathcal{O}_k - y_k)\frac{\partial}{\partial W_{ij}}\mathcal{O}_k \tag{3.19}$$

$$\frac{\partial E}{\partial W_{ij}} = \sum_{k\in K}(\mathcal{O}_k - y_k)\frac{\partial}{\partial W_{ij}}\sigma(z_k) \tag{3.20}$$

$$\frac{\partial E}{\partial W_{ij}} = \sum_{k\in K}(\mathcal{O}_k - y_k)\sigma(z_k)(1-\sigma(z_k))\frac{\partial z_k}{\partial W_{ij}} \tag{3.21}$$

$$\frac{\partial E}{\partial W_{ij}} = \sum_{k\in K}(\mathcal{O}_k - y_k)\mathcal{O}_k(1-\mathcal{O}_k)\frac{\partial z_k}{\partial \mathcal{O}_j}\cdot\frac{\partial \mathcal{O}_j}{\partial W_{ij}} \tag{3.22}$$

$$\frac{\partial E}{\partial W_{ij}} = \frac{\partial \mathcal{O}_j}{\partial W_{ij}}\sum_{k\in K}(\mathcal{O}_k - y_k)\mathcal{O}_k(1-\mathcal{O}_k)W_{jk} \tag{3.23}$$

$$\frac{\partial E}{\partial W_{ij}} = \mathcal{O}_j(1-\mathcal{O}_j)\frac{\partial z_j}{\partial W_{ij}}\sum_{k\in K}(\mathcal{O}_k - y_k)\mathcal{O}_k(1-\mathcal{O}_k)W_{jk} \tag{3.24}$$

$$\frac{\partial E}{\partial W_{ij}} = \mathcal{O}_j(1-\mathcal{O}_j)\mathcal{O}_i\sum_{k\in K}(\mathcal{O}_k - y_k)\mathcal{O}_k(1-\mathcal{O}_k)W_{jk} \tag{3.25}$$

Recalling the expression of δ_k, Equation 3.25 can be rewritten as

$$\frac{\partial E}{\partial W_{ij}} = \mathcal{O}_i\mathcal{O}_j(1-\mathcal{O}_j)\sum_{k\in K}\delta_k W_{jk} \tag{3.26}$$

We can also define all terms except $\mathcal{O}_i$ to be δ_j such that

$$\frac{\partial E}{\partial W_{ij}} = \mathcal{O}_j\delta_j \tag{3.27}$$

where

$$\delta_j = \mathcal{O}_j(1-\mathcal{O}_j)\sum_{k\in K}\delta_k W_{jk}$$

Once all the δ parameters have been calculated, a gradient descent can be performed by adjusting each of the connectivity weights to achieve a lower error rate. The back propagation algorithm can be summarized as follows:

1. With the inputs, run the network forward to obtain the model output
2. Calculate δ_k for each of the output layer nodes where $\delta_k = \mathcal{O}_k(\mathcal{O}_k - y_k)(1-\mathcal{O}_k)$
3. Calculate δ_j for each of the hidden layer nodes where $\delta_j = \mathcal{O}_j(1-\mathcal{O}_j)\sum_{k\in K}\delta_k W_{jk}$
4. Update the network weights as follows
   Given $\Delta W = -\alpha\delta_l\mathcal{O}_{l-1}$ where $l =$ network layer and α is the learning rate.

$$W_{new} \leftarrow W_{old} + \Delta W$$

The hidden layers introduce several local minima and stochastic gradient descent (SGD) is used to overcome this by randomizing the observations and updating the weights after each sample has been propagated through the network. Too low of a learning rate makes the learning process very slow. Too high of a learning rate will result in no weight update at all and the model may fail to converge. A negative is introduced to ensure that the gradient steeps downwards.

The process of training a neural network model iterates trough a series of weight modification cycles known as **epochs**. Once new weights have been set, the inputs are feed into the model resulting into a new error value and then another back propagation process begins. This goes on until the model converges - achieving the best accuracy for the given set of conditions.

3.2.5 Model evaluation

The trained model is tested on new data to asses the quality of predictions. The metrics may be in form of a score, a matrix or a curve. Some of the model metrics can be calculated using the true positive(tp), true negative (tn), false positive (fp) and false negative(fn) values from the confusion matrix. The model metrics used in this study include:

Confusion Matrix

This is a matrix of the model's predicted classes against the actual classes. It shows the extend to which the model is confused between the classes and highlights instances where one class is confused with the other. The leading diagonal shows the models correct classifications.

Classification Accuracy

This is the ratio of correctly predicted classes to the total number of predictions made.It is given by the sum of the diagonal values of the confusion matrix divided by the entire table.

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn} \tag{3.28}$$

Kappa Statistic

This is computed as a measure of agreement between predictions and observed labels. It compares the overall accuracy of the model to the expected random chance accuracy. High Kappa values show shows a better classification model.

$$Kappa = \frac{\textit{Model Accuracy} - \textit{Expected Accuracy}}{1 - \textit{Expected Accuracy}} \tag{3.29}$$

No information Rate

The No information Rate (NIR) metric shows the largest class percentage in the data. It is the best guess when we decide to always pick a member of the majority class. A significant model should perform better than a choice that always predicts the most common class.

Class-wise Precision, Recall and F-1 measure

When the class levels have non-uniform distribution, a single class may have most instances and therefore members of this dominant class may be predicted all the time hence showing a misleading accuracy. Precision and recall (sensitivity) metrics can be calculated for each class to support the accuracy values achieved. **Precision** (Positive Predictive Value) is the fraction of correct predictions for a particular class while **Recall** is the ratio of class members that were correctly predicted. The weighted average of precision and recall is called the **F-1 Measure**.

$$Precision = \frac{tp}{tp + fp} \tag{3.30}$$

$$Recall(sensitivity) = \frac{tp}{tp + fn} \tag{3.31}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{3.32}$$

3.3 Model Development

Neural networks with different combination of hidden layers and number of neurons per hidden layer were developed and compared to choose the best simple model that would best fit to the data. For the model selection, training was done on a subset of the data and another subset was used for testing the predictive capacity of the trained model. A 10-fold cross-validation was also done to ensure that the model does not overfit the training data. This involved splitting the data into 10 equal portions, training on 90% of the observations and testing on the remaining 10%, and repeating the process 3 times. Classification models developed to explain the disease pressure using weather variables. For classification, the late blight disease incidence was divided into five categories (1-3) with '1' indicating cases were low or no disease was observed to '3' with the highest disease recorded, while considering the relative AUDPC value . Prediction results from the model was compared to the observed values to asses the model performance on test data. Model accuracy on the test data was used to select the best model for classification. Accuracy and Kappa metrics were calculated to asses the overall model performance. Class wise performance was assessed by calculating the class precision, recall and F1 measure. All the computations were done using R and Rstudio software.

4 Results

4.1 Data Preparation and overview

The final dataset contained 817 cases and 14 variables. Relative risk of late blight incidence was the response variable. 80% of the data was used to train and validate the model and the rest was used to test the models performance on unobserved data. The cases to train and test model performance were selected at random.

Table 4.1. Sample observations from the final dataset used in model

lb_risk	season	treatment	variety	FDTP	FDTRh	Tmax5	Tmin5	Tav5	AP10d	PD	PT	MRH	WIND
Low	2	c	b	16.00	0.00	26.68	14.18	19.89	17.46	78.00	161.68	62.64	2.54
Medium	1	a	b	45.00	43.00	23.66	13.87	18.12	48.51	82.00	444.28	78.12	1.87
Medium	1	a	b	15.00	0.00	26.21	14.27	19.36	17.38	73.00	150.40	68.35	2.49
High	1	a	b	57.00	30.00	23.35	13.65	18.13	50.18	85.00	427.50	77.47	1.15
Low	2	c	b	16.00	0.00	26.68	14.18	19.89	17.46	78.00	161.68	62.64	2.54
Low	2	c	b	16.00	0.00	26.68	14.18	19.89	17.46	78.00	161.68	62.64	2.54
Low	2	a	b	50.00	11.00	25.37	14.31	19.07	34.75	87.00	319.13	71.63	2.71
Medium	1	d	b	15.00	0.00	26.21	14.27	19.36	17.38	73.00	150.40	68.35	2.49
Medium	2	a	a	50.00	11.00	25.37	14.31	19.07	34.75	87.00	319.13	71.63	2.71
Medium	1	b	b	15.00	0.00	26.21	14.27	19.36	17.38	73.00	150.40	68.35	2.49

Varieties in the control experiments were put into 3 clusters using Scott-Knott mean clustering method at $p < 0.05$ (Table 4.2 and Figure 4.1).

Table 4.2. Variety mean groups on rAUDPC of the control experiments

Variety	Means	Scott-Knott
		($p < 0.05$)
Arka	0.81	a
Nyayo	0.64	a
Desiree	0.60	a
Asante	0.47	b
Purple Gold	0.46	b
Dutch Robjin	0.45	b
Shangi	0.44	b
Tigoni	0.42	b
Mavuno	0.40	b
Dutch Robijn	0.36	b
Kihoro	0.27	c
Kenya Karibu	0.15	c
Kenya Sifa	0.09	c
Kenya Mpya	0.04	c

The control represents the varieties genetic resistant characteristics as there is no masking by treatment. This grouping agrees with previous studies on resistance evaluation done by Kamuyu 2017. Similarly, the treatments were put into five clusters (Figure 4.2).

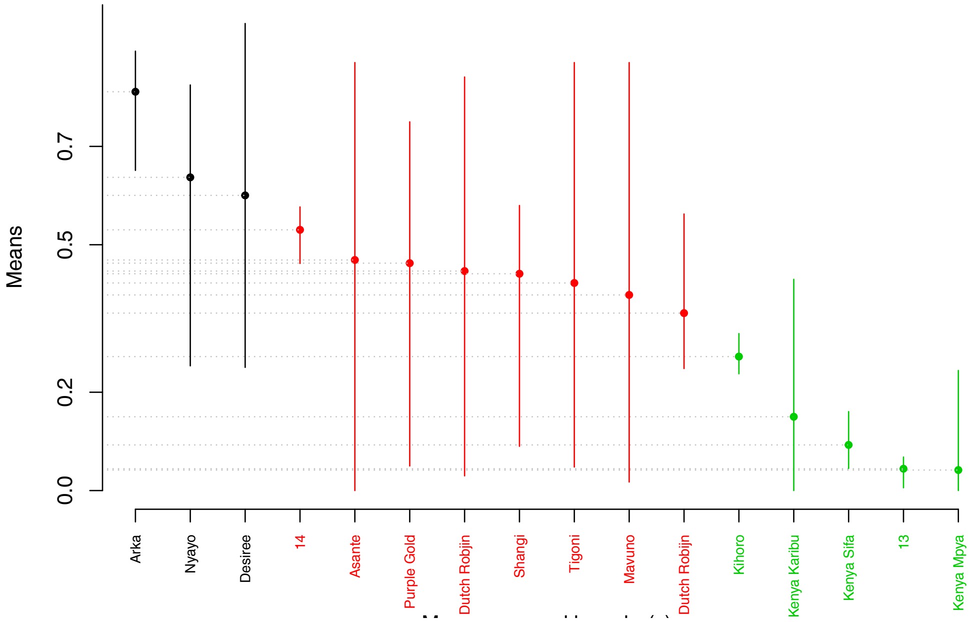

Figure 4.1. Variety Mean Groups and Range

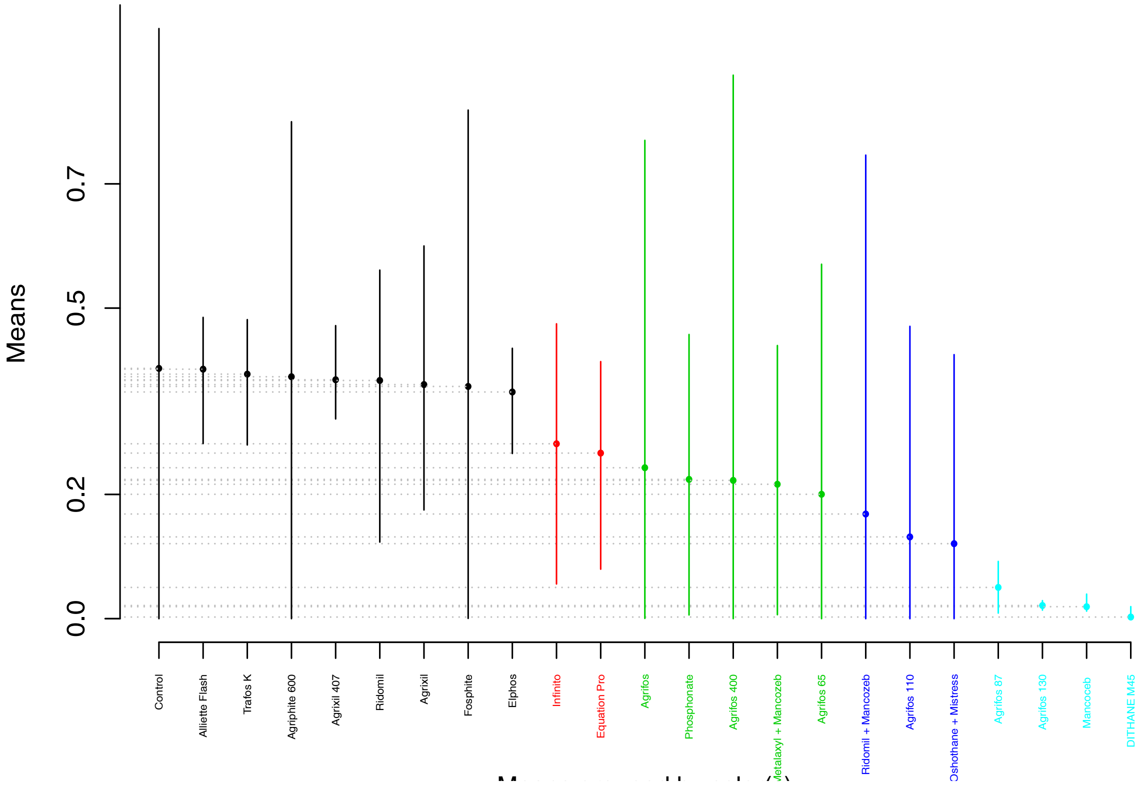

Figure 4.2. Treatment Mean Groups and Range

4.2 Model Results

4.2.1 Neural Network Model

The model was trained on 651 observations, with 10 fold cross validation repeated 3 times. The numeric variables were preprocessed by scaling between 0 and 1 and centered. 7 different model combinations were run and the optimal model chosen based on the largest accuracy value.

Table 4.3. Resampling results across tuning parameters to select the best tune model

Model	Size	Decay	Accuracy	Kappa
Model 1	2	0.10	0.87	0.74
Model 2	6	0.13	0.87	0.76
Model 3	10	0.16	0.87	0.76
Model 4	14	0.19	0.88	0.76
Model 5	18	0.22	0.87	0.75
Model 6	22	0.25	0.87	0.75
Model 7	26	0.28	0.87	0.75

Table 4.3 and Figure 4.3 show the that the best performing model by accuracy has 14 nodes in the hidden layer, achieving an accuracy of 88%.

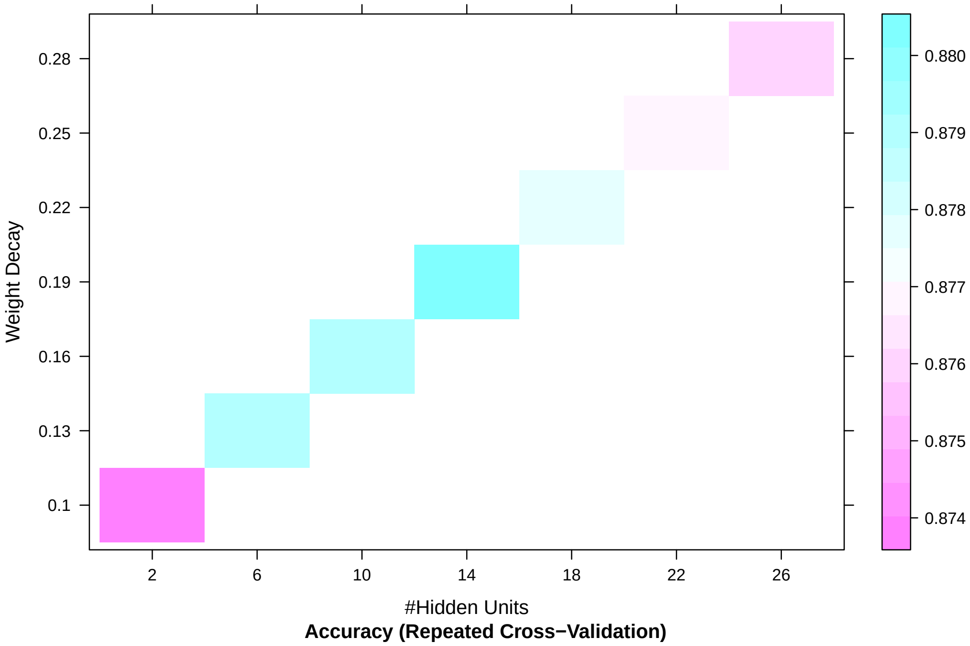

Figure 4.3. Best model structure by Accuracy

4.2.2 Network Interpretation Diagram

A network interpretation diagram for the final model show the nature of the various connections among the network nodes. Each of the of the variables in the input layer connects to all the computational nodes in the hidden layer, contributing to the final sum of inputs to the node, z, either positively or negatively as shown in Figure 4.4 below.

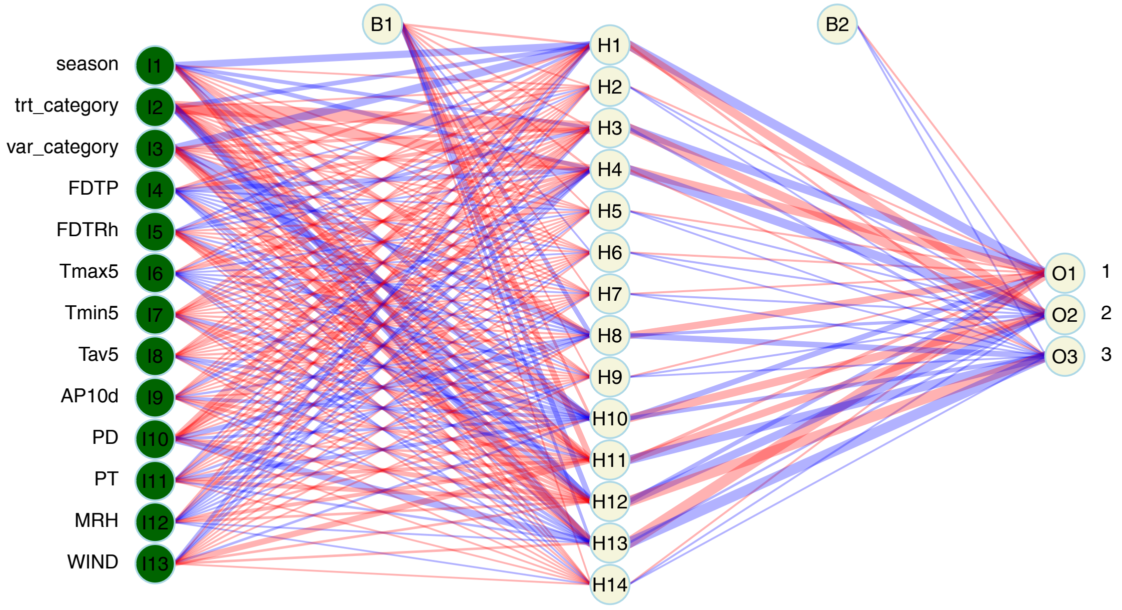

Figure 4.4. Neural network interpretation diagram for the final model with 14 nodes in the hidden layer. Positive connection weights in red and negative weights in blue.

4.2.3 Model Performance

Comparing the predicted labels with the actual data labels, the confusion matrix shows most of the classifications in the main diagonal. Out of 166 labels in the test data, 148 are correctly classified leading to 0.892 classification accuracy with a 95% confidence interval range between 0.834 and 0.935 and only 10.8% misclassification rate (Table 4.4). A significant accuracy level was obtained with a $p - value_{[Acc>NIR]}$ of 1.39×10^{-9}. Cohen's Kappa value of show that the model accuracy was 75% greater than random expected chance accuracy.

Table 4.4. Confusion Matrix

	Actual		
Predicted	Low	Medium	High
Low	34	5	0
Medium	7	109	5
High	0	1	5

Table 4.5. Statistics by Class

	Low	Medium	High
Precision	0.87	0.90	0.83
Recall	0.83	0.95	0.50
F1 Measure	0.85	0.92	0.62

Table 4.5 show how the model performed in correctly predicting the classes. The values for precision, recall and F1 metrics were generally high.

Table 4.6. Distribution of the response variable

	Disease	Risk	
Levels	Low	Medium	High
Training	30%	62%	8%
Testing	25%	69%	6%
Predicted	23%	73%	4%

Although recall value for the third class was low compared to the others, the F1 Measure shows that optimal levels of precision and recall were achieved in the model. The plot below shows a precision-recall curve with an area under curve of 0.93 which means that the model has a good class separation performance.

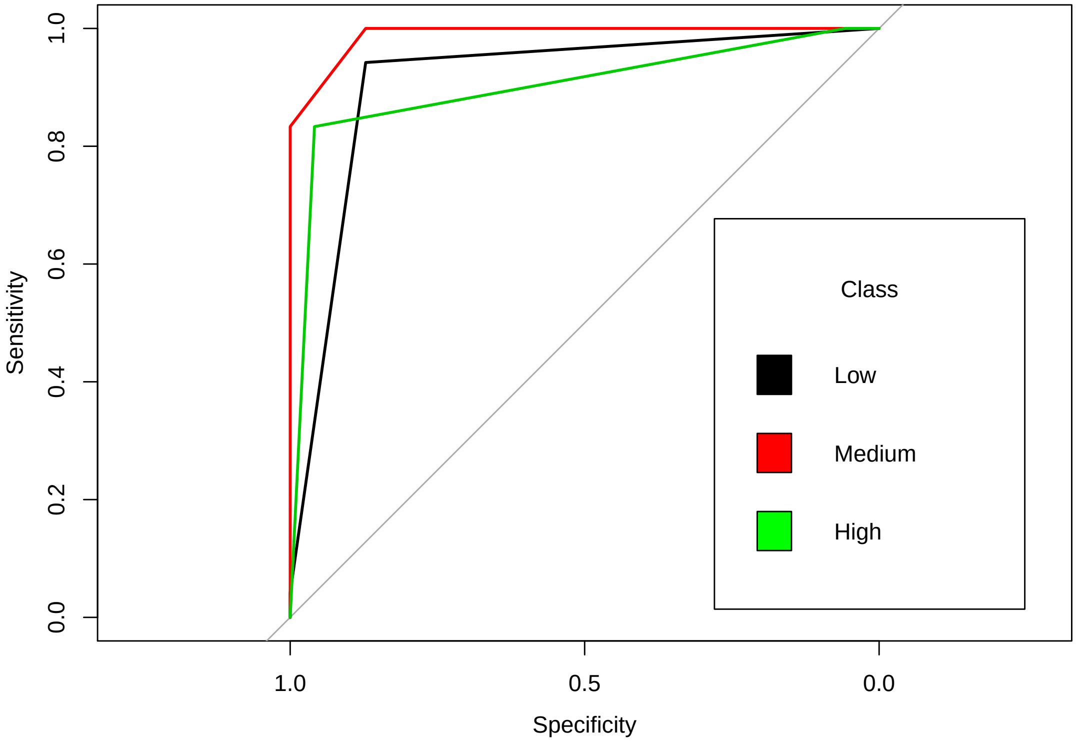

Figure 4.5. Area under curve for the three classes. Multi-class area under the curve = 0.93

4.2.4 Variable importance and feature selection

In order to select the most significant contributors to the model's performance, the scaled variable importance result show that the type of late blight management treatment used has the most significant effect on the model output followed by the number of days with precipitation of 0.25mm or higher. The number of days with favorable temperature and relative humidity is also a significant model feature

Table 4.7. Top five model features ranked by scaled average variable importance across the classes

	Feature	% Importance
1	Treatment category	100.00
2	PD(days)	61.85
3	FDTRh(days)	47.34
4	Variety category	41.95
5	Season	33.62

5 Discussion

5.1 Data preparation and overview

Most of the selected sites fall under areas with annual precipitation above 1000mm and relative humidity of above 60% as shown in Figure 1.1, conditions that favor the disease development. The regions however experience different weather patterns that have an impact on the magnitude of late blight experienced.

5.2 Model building and selection

The model selection followed a seven series of iterations with varied connectivity parameters. Each of the model was subjected to a repeated cross validation process to ensure that the final model fitted to the training set as well as generalizing to new dataset with a significant accuracy. 10 - fold cross validation was performed while training model and the best model selected with 88% accuracy on the validation set and a Kappa agreement of 76%. Cross validation introduces a form of rotational estimation while building the model to asses the performance of the model outside the sample before applying it to the test set. It ensure that a stable model is selected.

The presence of multiple computational nodes in the hidden layer, each producing a δ gradient estimate might cause the model to get stuck in one of the many local minima. To counter this, stochastic gradient descent was employed whereby weights updates were performed after running a sample of the dataset through the model as opposed to running the entire dataset before updating the weights. This ensure that the model does not get stuck in local minima and achieves the best accuracy possible as the global minimum.

The best tune for final model chosen had a [13-14-3]- structure with one hidden layer and a weight decay of 0.19 (Figure 4.4). Overall, the model was able to predict new cases with up to 89% accuracy on the test dataset and 75% agreement between the predictions and the data labels.

5.3 Model evaluation

The results show that the model fit performed significantly both in-sample and outside sample by achieving a cross validation accuracy of 0.88 and a significant out-of-sample accuracy of 0.89. This model accuracy is larger than the majority class thus the model performs better than a prediction done by always choosing the majority class. This is an important measure for the selected model because the data used to train and test the model has a high level of class imbalance as in Table 4.6, showing the 'Medium' class as the majority class and 'High' class as the minority class. A model that would always predict the majority class would have achieved a 0.69 accuracy, which is outperformed by the final model selected. A kappa statistic validates the accuracy achieved by the model against a random accuracy and shows a 75% agreement between predicted results and the actual observations.

Class metrics in Table 4.5 show that the model performed well in predicting individual classes. The probability of the model to correctly identify the classes (precision) was high for all the three classes. The recall value was small for the "High" class compared to the other classes. This could be due to the high level of class imbalance in the data as seen in Table 4.6 with the third class having the least number of occurrences overall.

There is evidence, however, that all the classes have been correctly identified. This is supported by the precision recall curve (PR-C) in Figure 4.5 above which shows an 93.3% area under curve . In a PR-C curve, 100% area shows a perfect test while a 50% area shows a no skill model.

5.4 Feature selection

Each of the input variables contributes a certain strength of signal, either positive or negative, towards achieving the outcome. As shown in Figure 4.4, the intensity of the color shows how strong an input contributes towards the final result. A variable importance rank in Table 4.7 shows which variables contribute most across the predicted classes. The model shows that the treatment used, the number of precipitation days, the number of temperature - relative Humidity disease favorable days and the genotype as among the top most important contributors to the model outcome.

Conclusion

Artificial neural network model architecture was successfully implemented to predict late blight disease pressure using data on weather conditions and data on crop management practices. Performing a 10-fold cross validation produced a stable model that fits the training data well and generalized well on new datasets. The model would play a significant role in forecasting the expected magnitude of late blight incidence to be experienced in a particular region, given that the pathogen is present. It also highlights the weather variables that require close monitoring as these have a significant influence on the disease development cycle. The model can be coupled with existing late blight models to provide a strong decision support tools for potato variety choices and the type of late blight control measures that can be employed in various agroecological zones to minimize crop damage and excessive use of fungicides. To use this model, however, a few validation trials would be required in order to calibrate the model before performing predictions.

6 Appendix 1

6.1 R code used for computation

6.1.1 Data partition

```
agro_weather <- readRDS("agro_weather.rds")
#processing data for use in training and testing model
data <- agro_weather %>%
  mutate_if(is.factor, funs(match(., unique(.)))) %>%
  dplyr::select(-audpc, -raudpc, -yield_tha, -year, -site) %>%
  na.omit() %>%
  mutate(lb_risk = factor(lb_risk)) %>%
  dplyr::select(lb_risk, everything())
# create data partition
set.seed(123)
ind <- sample(2, nrow(data), replace = T, prob = c(.8, .2))
trainingDF <- data[ind == 1, ]
testDF <- data[ind == 2, ]
```

6.1.2 Model training

```
set.seed(1234)
train_params <- trainControl(method = 'repeatedcv',number = 10,repeats = 3)

future::plan('multiprocess')
model_nnet_class <- caret::train(
  trainingDF[-1],trainingDF$lb_risk,
  method = 'nnet',
  trControl = train_params,
  maxit = 500,
  tuneGrid = data.frame(
    #iterate through different model structures to
    #select the structure with the highest accuracy
    size = seq(2,28,by=4),
    decay = seq(0.1,0.3, by= 0.03)
  ),
  #Normalize the data before training the model to improve numerical accuracy
  preProcess = c('scale','center'),
  na.action = na.omit,             # Omit any cases with missing values
  trace= FALSE,skip = TRUE
)

#Testing the model performance on new data
pred_nnet <- predict(model_nnet_class, testDF)
postResample(pred_nnet, ordered(testDF$lb_risk)) %>% round(2)
```

6.1.3 Calculating model classwise metrics

```
cm_nnet <- confusionMatrix(pred_nnet, testDF$lb_risk)
cmatrix <- cm_nnet$table
n <- sum(cmatrix)
nc <- nrow(cmatrix)
dg <- diag(cmatrix)
rowsum <- apply(cmatrix, 1, sum)
colsum <- apply(cmatrix, 2, sum)
x <- rowsum / n
y <- colsum / n
acc <- sum(dg) / n
expAcc = sum(x*y)
kap <- (acc-expAcc)/(1-expAcc)
```

Classwise performance scores

```
prec <- dg / colsum
rec <- dg / rowsum
f1 <- 2 * prec * rec / (prec + rec)
x <- round(data.frame("Precision" = prec,
                      "Recall" = rec,
                      "F1 Measure" = f1),
           2)
```

Bibliography

1. Alves, Daniel Pedrosa et al. (2017). “Artificial Neural Network for Prediction of the Area under the Disease Progress Curve of Tomato Late Blight”. In: *Scientia Agricola* 74.1, pp. 51–59. ISSN: 0103-9016. DOI: `10.1590/1678-992x-2015-0309`.
2. Beaumont, A. (1947). “The Dependence on the Weather of the Dates of Outbreak of Potato Blight Epidemics”. In: *Transactions of the British Mycological Society* 31.1-2, pp. 45–53.
3. Cook, Harold T. (1949). *Forecasting Lete Blight Epiphytotics of Potatoes and Tomatoes.*
4. Crane-Droesch, Andrew (2018). “Machine Learning Methods for Crop Yield Prediction and Climate Change Impact Assessment in Agriculture”. In: *Environmental Research Letters* 13.11, p. 114003. ISSN: 1748-9326. DOI: `10.1088/1748-9326/aae159`.
5. Gu, Y.H. et al. (2016). “BLITE-SVR: New Forecasting Model for Late Blight on Potato Using Support-Vector Regression”. In: *Computers and Electronics in Agriculture* 130, pp. 169–176. ISSN: 01681699. DOI: `10.1016/j.compag.2016.10.005`.
6. Hansen, Jens Grønbech (1995). “NEGFRY - A System for Scheduling Chemical Control of Late Blight in Potatoes”. In: *Proceedings Phytophthora 150 Sesquicentennial Scientific Conference*, pp. 201–208.
7. Henderson, Donna, Christopher J. Williams, and Jeffrey S. Miller (2007). “Forecasting Late Blight in Potato Crops of Southern Idaho Using Logistic Regression Analysis”. In: *Plant Disease* 91.8, pp. 951–956. ISSN: 0191-2917. DOI: `10.1094/PDIS-91-8-0951`.
8. Hijmans, R. J., G. A. Forbes, and T. S. Walker (2000). “Estimating the Global Severity of Potato Late Blight with GIS-Linked Disease Forecast Models”. In: *Plant Pathology* 49.6, pp. 697–705. ISSN: 0032-0862, 1365-3059. DOI: `10.1046/j.1365-3059.2000.00511.x`.
9. Kaguongo, Wachira et al. (2010). *Seed Potato Subsector Master Plan for Kenya (2009-2014).*
10. Kamuyu, Loise Muthoni (2017). “Health Status Of Potato Seed And Host Resistance Against Late Blight Disease Under Greenhouse And Field Conditions In Kenya”. University of Nairobi. 116 pp.

11. Krause, R. A. (1976). “Blitecast; a Computerized Forecast of Potato Late Blight. Implementation and Evaluation”. In: *Modeling for Pest Management: Concepts, Techniques, and Applications; US USSR Symposium.*

12. Kromann, Peter et al. (2012). “Use of Phosphonate to Manage Foliar Potato Late Blight in Developing Countries”. In: *Plant Disease* 96.7, pp. 1008–1015. ISSN: 0191-2917. DOI: 10.1094/PDIS-12-11-1029-RE.

13. Maina, Christine Njeri (2016). “Vision-Based Model for Maize Leaf Disease Identification:” Strathmore University.

14. Mariita, Micah, Johnson Nyangeri, and Jacqueline Makatiani (2016). “Assessing the Incidences of Late Blight Disease on Irish Potato Varieties in Kisii County, Kenya”. In: *Annual Research & Review in Biology* 9.6, pp. 1–8. ISSN: 2347565X. DOI: 10.9734/ARRB/2016/23617.

15. Muchiri, F. N. et al. (2009). “Efficacy of Fungicide Mixtures for the Management of Phytophthora Infestans (US-1) on Potato”. In: *Phytoprotection* 90.1, pp. 19–29. ISSN: 0031-9511, 1710-1603. DOI: https://doi.org/10.7202/038983ar.

16. Muthoni, Jane, D.O. Nyamongo Nyamongo, and M. Mbiyu (2017). “Climatic Change, Its Likely Impact on Potato (*Solanum Tuberosum* L.) Production in Kenya and Plausible Coping Measures”. In: *International Journal of Horticulture.* ISSN: 1927-5803. DOI: 10.5376/ijh.2017.07.0014.

17. Nyankanga, R. O. et al. (2004). “Farmers’ Cultural Practices and Management of Potato Late Blight in Kenya Highlands: Implications for Development of Integrated Disease Management”. In: *International Journal of Pest Management* 50.2, pp. 135–144. ISSN: 0967-0874. DOI: 10.1080/09670870410001691812.

18. Rizzo, Donna M., Susanne Conklin, and David E. Dougherty (2003). “Using Artificial Neural Networks to Predict Local Disease Risk Indicators with Multi-Scale Weather, Land and Crop Data”. In: *World Water &amp; Environmental Resources Congress 2003.* World Water and Environmental Resources Congress 2003. Philadelphia, Pennsylvania, United States: American Society of Civil Engineers, pp. 1–10. ISBN: 978-0-7844-0685-4. DOI: 10.1061/40685(2003)230.

19. Sannakki, S. et al. (2013). “A Neural Network Approach for Disease Forecasting in Grapes Using Weather Parameters”. In: *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT).* 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT). Tiruchengode: IEEE, pp. 1–5. ISBN: 978-1-4799-3926-8 978-1-4799-3925-1. DOI: 10.1109/ICCCNT.2013.6726613.

20. Scott, A. J. and M. Knott (1974). "A Cluster Analysis Method for Grouping Means in the Analysis of Variance". In: *Biometrics* 30.3, pp. 507–512. ISSN: 0006-341X. DOI: 10.2307/2529204.

21. Shaner, Gregory (1977). "The Effect of Nitrogen Fertilization on the Expression of Slow-Mildewing Resistance in Knox Wheat". In: *Phytopathology* 77.8, p. 1051. ISSN: 0031949X. DOI: 10.1094/Phyto-67-1051.

22. Sharma, Priyanka, B.K. Singh, and R.P. Singh (2018). "Prediction of Potato Late Blight Disease Based Upon Weather Parameters Using Artificial Neural Network Approach". In: *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*. 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT). Bangalore: IEEE, pp. 1–13. ISBN: 978-1-5386-4430-0. DOI: 10.1109/ICCCNT.2018.8494024.

23. Shastry, K. Aditya, H.A. Sanjay, and Abhijeeth Deshmukh (2016). "A Parameter Based Customized Artificial Neural Network Model for Crop Yield Prediction". In: *Journal of Artificial Intelligence* 9.1, pp. 23–32. ISSN: 19945450. DOI: 10.3923/jai.2016.23.32.

24. Taylor, M.C et al. (2003). "Relative Performance of Five Forecasting Schemes for Potato Late Blight (Phytophthora Infestans) I. Accuracy of Infection Warnings and Reduction of Unnecessary, Theoretical, Fungicide Applications". In: *Crop Protection* 22.2, pp. 275–283. ISSN: 02612194. DOI: 10.1016/S0261-2194(02)00148-5.

25. Toroitich, Patrick Kiplimo (2017). "A Model for Early Detection of Potato Late Blight Disease: A Case Study in Nakuru County". Strathmore University.

26. Vianna, Gizelle K, Gabriel V Cunha, and Gustavo S Oliveira (2017). "A Neural Network Classifier for Estimation of the Degree of Infestation by Late Blight on Tomato Leaves". In: 11.1, p. 7.

27. Wallin, Jack R. (1953). "The Production And Survival Of Sporangia Of Phytophthora-Infestans On Tomato And Potato Plants In The Field". In: *Phytopathology* 43.9, pp. 505–508.

28. Yang, Xin and Tingwei Guo (2017). "Machine Learning in Plant Disease Research". In: *European Journal of BioMedical Research* 3.1, p. 6. ISSN: 2428-5544. DOI: 10.18088/ejbmr.3.1.2017.pp6-9.