AGRICULTURAL CREDIT MODELS: IDENTIFYING HIGH RISK APPLICANTS

By

DOMINIQUE ALYSSA BOUGARD

Submitted in accordance with the requirements for the degree

M.Agric

In the

J.I.F. Henning
H. Jordaan

FACULTY OF NATURAL AND AGRICULTURAL SCIENCES
DEPARTMENT OF AGRICULTURAL ECONOMICS
UNIVERSITY OF THE FREE STATE
BLOEMFONTEIN

DECLARATION

I, Dominique Alyssa Bougard, hereby declare that this dissertation, submitted for the degree of Master of Agriculture (M.Agric) in Agricultural Management at the University of the Free State, is my own, independent work and has not previously been submitted by me to any other university.

I, Dominique Alyssa Bougard, hereby declare that I am aware that the copyright is vested in the University of the Free State.

I, Dominique Alyssa Bougard, hereby declare that all the royalties with regard to intellectual property that was developed during the course of and/or in connection with the study at the University of the Free State, will accrue to the University.

______________________ ____________________

Dominique Alyssa Bougard Date

Bloemfontein

ACKNOWLEDGEMENTS

Firstly, I would like to thank my parents, Charmaine and Kenneth, for the continuous love, support, guidance and encouragement you have both given me throughout my studies. You both have been my pillars of strength through the good and difficult times of my life. I would not be where I am today without you both. My gratitude extends to my sister, Colleen, for her constant motivation and support.

I would like to extend my gratitude to my supervisor, Dr Janus Henning, for his excellent guidance and patience, valuable input at each stage of my dissertation, as well as for providing me with the opportunity to continue with my studies. In addition, I would like to thank my co-supervisor, Dr Henry Jordaan, for his support and guidance throughout this study.

Furthermore, I would like to thank the senior staff at the department of Agricultural Economics, Ms Louise Hoffman, Ms Chrizna van der Merwe and Ms Ina Combrink for all your kindness and advice, it has made missing home a lot easier. Moreover, I would like to express my utmost thanks to all my colleagues and friends for supporting me through this research. A special thanks to my best friend and roommate, Janefer Starke, for her constant motivation, positivity and friendship throughout my studies.

Lastly, this study is based on research supported in part by the National Research Foundation of South Africa, through the grant, Unique Grant No. 94132. “Any opinion, finding and conclusion or recommendation expressed in this material is that of the author(s) and the NRF does not accept any liability in this regard.”

LIST OF ACRONYMS AND ABBREVIATIONS

ATO	Asset turnover ratio
DTA	Debt to assets
Bb	Bad/bad
Bg	Bad/good
Gb	Good/bad
Gg	Good/good
LR	Logistic regression
NETFARMRATIO	Net farm income ratio
NN	Neural networks
PA	Probit analysis
PRODCOST	Production costs
ROA	Return on assets
Tb	Total predicted bad observations
TB	Total bad observations
Tg	Total predicted good observations
TG	Total actual good observation
TN	Total amount of observations
WCTGR	Working capital to gross revenue

ABSTRACT

The objective of the research was to explore the performance of various statistical credit-scoring models, in order to identify a model that will minimise the misclassification of high-risk applicants, and identify the characteristics that influence repayment ability.

The study was conducted in South Africa, with the use of a case study of a South African financial organisation serving the agricultural sector. The data gathered for this study was gathered through a formal agreement with a commercial financial organisation. Logistic regression (LR), probit analysis (PA) and neural network (NN) were used to construct the credit-scoring models that can be used to classify credit applications in the agricultural sector.

Results of the LR indicate significance at 10% of the following variables, which may have an impact on classification: medium-term loan, credit history, debt to assets (DTA), net farm ratio, diverse 2, high risk, ownership and experience. The PA results demonstrate the following variables at 10% significance: credit history, DTA, net farm ratio, diverse 2, ownership and experience. The identification of characteristics provides confirmation of characteristics that are of importance to credit research. Financial organisations can use the identification of important characteristics as a method to provide guidance to applicants who apply for loans. Doing so will ensure that the organisation will identify characteristics that ensure that the applicant is accepted by the financial organisation. Applicants for loans can ensure that they possess characteristics that correspond to important characteristics identified by the statistical model. The results from the NN are not easily interpretable; due to "black-box" qualities it was not easy to identify the variables that have an influence on the predicted outcome. The NN did, however, outperform the LR and PA in terms of classification accuracy. Neural networks achieved the highest correctly predicted overall accuracy and a lower percentage of Type II error classifications. Logistic regression and PA have overall classification percentages of 96.06% and 3.94% respectively for classifying Type II errors. The NN had an overall classification accuracy of 98.43% and Type II classification error of 1.54%. The main conclusion from this research is that the statistical methods are able to classify credit applications in the agricultural sector and have the ability to improve accuracy in correctly classifying agricultural applicants.

Further research is need to ensure that the correct variables are included in the classification. The classification results of the models are tested and monitored over a period of time to ensure that the accuracy and prediction are acceptable according to the financial organisations. Further research is needed to select the correct variables to be used when supplying credit to smallholder farmers and financial organisations can use the identified important characteristics to provide recommendations and guidance when evaluating applications for loans. Credit applicants can also use these identified important characteristics as a point of reference before applying for the loan at the financial organisation.

Keywords: Credit, Credit Evaluation, Credit Characteristics, Classification Matrix, Logistic Regression, Probit Analysis, Neural Networks

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1. BACKGROUND AND MOTIVATION

In recent years, formal financial organisations have increased total lending to the South African agricultural sector significantly (Qwabe, 2014). Lending in the agricultural sector has increased due to the demand for credit to finance farm production activities and capital expenditure. The increase in agricultural debt is caused by a strong reliance on credit to finance capital investments, such as machinery, vehicles, livestock, implements and land (DAFF, 2015). This capital is required to support farmers' operations so that they can use the available natural resources to their maximum potential. Over the past ten years, total South African agricultural debt has increased by 71%, from R36 443,8 million in 2005 to an estimated R125 712 million in 2015 (DAFF, 2015). The increase in debt has made financial organisations more aware of the need to improve credit evaluation procedures (Salame, 2011). Smallholder farmers are reliant on credit, but struggle to access finance from financial organisations (Chisasa, 2014). The lack of credit has an effect on the productivity of these smallholder farmers. In South Africa smallholder farmers struggle to access credit due to their inability to provide collateral, which is required by financial organisations (Chisasa & Makina, 2012).

The National Credit Act of 2005 defines consumer credit as a "deferral of payment of money owed to a person, or a promise to defer such a payment; or a promise to advance or pay money to or at the direction of another person." Michael, Miller & Gegenbauer (2009) state that agricultural credit is "the advance of funds to enterprises to finance inputs, production and accompanying support operations, using certain types of security that are not normally accepted by banks or investors and which are more dependent on the structure and performance of the transaction, rather than the characteristics of the borrower." As mentioned in the definition by Michael *et al.* (2009), security has a very important role in providing access to funds, as it serves as collateral in case of default. Funds are advanced to the applicant by means of a review or evaluation process during which the security provided is considered. Borrower characteristics, such as age, education, experience, management capability and reputation, are also considered (Henning & Jordaan, 2016).

Before credit can be granted a specific evaluation process must be followed to determine the creditworthiness of the farmer. This evaluation process consists of the collection, analysis and evaluation of information, such as the farmer's credit repayment history, income and overall finance, before credit can be granted (USAID, 2005). A credit-scoring points system has been designed to evaluate the credit application by adding the points gathered from the various application features to generate a total score (Abdou & Pointon, 2011). Once the credit evaluation of the applicant has been completed and he/she has been identified as an acceptable risk, the credit officer compiles an acceptable loan structure that protects the bank from the identified weaknesses and strengths of the borrower (Abdou & Pointon, 2011). If the applicant is identified as a high-risk applicant, the credit officer will reject this applicant to protect the bank from possible financial loss. When mistakes are made during the classification of applicants, costs are incurred by the financial organisations, these costs are known as misclassification errors.

Misclassification errors and increased demand for credit have encouraged financial organisations to explore alternatives for loan classification, to improve accuracy (Abdou & Pointon, 2011). Misclassification errors refer to accepting high-risk loans and rejecting low-risk loans. To reduce misclassification errors various statistical credit-scoring models have been developed. These models have the potential to reduce the inconsistency of credit decisions and improve the credit-evaluation process (Limsombunchai, Gan & Lee, 2005). For a financial organisation accepting a high-risk applicant is more costly than rejecting a low-risk applicant (Marqués, García & Sánchez, 2013). Therefore, it is important for financial organisations to minimise their risk exposure by correctly classifying high-risk loans.

Two approaches, namely the subjective and objective approach, can be used to assess the repayment ability of an applicant. The subjective approach is reliant on the knowledge and experience of the analyst who determines the applicant's repayment ability. The analyst can discriminate and incorrectly classify the applicant based on personal knowledge, instead of observing their financial ability (Limsombunchai *et al.,* 2005; Abdou & Pointon, 2011). This approach has been found to be inefficient and inconsistent (Alaraj, Abbod & Al-Hnaity, 2015), and could lead to misclassification errors. Therefore, to ensure more accurate and consistent loan classifications, the use of objective approach is advised. Credit-scoring models can reduce the need for human judgement (Marqués *et al.,* 2013), and reduce inconsistency and misclassification errors. Not only will human judgement be reduced and inconsistency be improved but the costs associated with misclassification of applicants will be reduced. The misclassification of applicants has contributed to extensive research into statistical credit models, in an attempt to identify and suggest models that reduce misclassification and, consequently, costs for financial organisations.

1.2. PROBLEM STATEMENT AND OBJECTIVES

Ample international research has focused on loan classification, and has explored the variables that influence access to credit, and principles, theories and operational frameworks for credit-evaluation techniques (Bandyopadhyay, 2007; Abdou & Pointon, 2011; Marqués *et al.,* 2013, Henning & Jordaan, 2016). These factors have been used in statistical models, such as discriminant analysis, linear regression, genetic programming, logistic regression (LR), decision tree, probit analysis (PA), expert systems, k-nearest neighbours, kernel density, support vector machine and neural networks (NN) (Abdou & Pointon, 2011) in different sectors, including the agricultural sector, to predict ability of a prospective borrower to repay a loan.

Researchers have compared and explored various methods to improve accuracy in the evaluation of credit applications. Research applied these statistical credit-scoring models to different sectors in international financial organisations in countries such as Thailand (Limsombunchai *et al.*, 2005), France (Jouault & Featherstone, 2006), Spain (Marqués *et al.*, 2013), India (Bandyopadhyay, 2007), Egypt (Abdou & Pointon, 2011), Canada (Nayak & Turvey, 1997) and the United States of America (Quaye, Haratrska & Nadolnyak, 2015). These statistical credit-scoring models have proved to be efficient and effective compared to the subjective approach. West (2000) states that even a fraction of a percent increase in credit-scoring accuracy can be regarded as a significant accomplishment. This improvement does not seem large; but compared to the number of credit applications that must be assessed, even this small improvement will have an effect on accuracy. Even though default is a rare occurrence in agricultural lending, when default does occur, the values are high and related to the performance of the farm (September, 2009).

Salame (2011) examined the performance of NN, LR and decision trees in terms of misclassification rates of credit default in agriculture. The results show that there are small differences between misclassification errors and the various models used. Limsombunchai *et al.* (2011) compared LR with NN to determine misclassification rates of credit default in agriculture. These models demonstrate successful results. Attempts to identify a statistical credit-scoring model that can predict high risk loans accurately has not received as much research attention, especially when the South African agricultural sector is considered.

In South Africa, literature was found on the development of a credit-risk model for agriculture-based structured finance lending transactions (Lubinda, 2010). Henning & Jordaan (2015, 2016) considered the factors used by financial organisations to evaluate agricultural credit applications. Henning (2016) used NN to classify agricultural loan applicants, however, he did not compare different statistical models to assess which model performed best in terms of accuracy and reducing misclassification of high-risk applicants. Few attempts have been made to identify a statistical model that can accurately classify high-risk loans, especially in agricultural credit research in South Africa. Thus, there is no scientific evidence available, specifically in South Africa, regarding the best-performing statistical model for assessing credit applicants.

The aim of the research is to explore the performance of various statistical credit-scoring models to identify a model that will minimise the misclassification of high-risk applicants, and identify the characteristics that influence repayment ability.

1.3. CHAPTER OUTLINE

The rest of this dissertation is organised in four remaining chapters. **Chapter 2** provides the relevant literature on the role of credit, credit-evaluation procedure, misclassification and variables used in statistical credit-scoring models. Included in Chapter 2 are the two credit-granting approaches, namely, the subjective and objective approaches. The chapter also provides an introduction of the various statistical credit models and identifies the statistical models that are, according to literature, the most accurate. These models identified by literature are further selected according to their advantages and disadvantages. **Chapter 3** provides an overview of the data collection, characteristics of respondents and the methodology used to generate the results. **Chapter 4** gives a presentation and discussion of the results obtained. The final chapter, **Chapter 5,** includes a summary, final conclusions made from the study and possible recommendations that can improve this study.

CHAPTER 2

LITERATURE REVIEW

2.1. INTRODUCTION

Chapter 2 provides an overview of the relevant literature on the evaluation techniques and methods that can be used to reduce misclassification errors involving high-risk agricultural loan applicants. Firstly, the role of credit will be discussed, before the credit-evaluation process is explained. The credit-evaluation process is discussed further, revealing the two approaches that can be used to evaluate applicants. This discussion also includes a comparison of various popular statistical models, including the various advantages and disadvantages of the selected statistical models. Lastly, misclassification and the characteristics used to predict repayment ability of applicants is discussed.

2.2. ROLE OF CREDIT

According to Spencer (1997) “credit implies a promise by one party to pay the other for money borrowed or goods and services received.” Access to credit is considered to be an important necessity for economic development and improving standards of living (Petrick, 2005). Individuals, families, government bodies and business firms apply for credit in order to purchase resources, pay for goods and services and meet operating expenses (Marqués *et al.,* 2013; Yakubu & Affoi, 2014). Government obtains credit to meet various kinds of capital and recurrent expenses, and business firms require credit to purchase machinery and other equipment (Yakubu & Affoi, 2014). The agricultural sector does not differ much from other sectors in terms of needs for access to credit. The agricultural sector is, however, influenced by different factors, which may influence the repayment ability of the applicants from the sector. Credit can also be used to revive economic activities that have suffered setbacks caused by natural disasters or unforeseen weather patterns (Ademu, 2006).

Financial capital is an important backbone for any business, including the agricultural sector. Agriculture is more dependent on capital (credit and equity) than any other sector of the economy due to the trend of change, from subsistence to commercial farming, and seasonal variations in farm returns (Mahmood, Khalid & Kouser, 2009). Credit capital refers to capital that is borrowed and must be repaid at a later stage, while equity capital is capital that is generated from investments by shareholders (Gitman *et al.,* 2014: 259). Equity capital is funds that consist of long-term funds provided by the firm's owners, that is, the shareholders; these funds do not need to be repaid but the owners receive a profit in the form of shares (Gitman *et al.,* 2014: 259). Credit is an important input for agricultural development, as it permits farmers to accept new investments and/or to accept new technology (Kumara, Singh & Sinha, 2010). This enables farmers to increase productivity and efficiency within agricultural businesses. The agricultural sector is heavily dependent on credit to ensure that production continues. The associated risk of this sector is very high, and the agricultural sector is regarded as having a higher degree of credit risk than other sectors in the economy (September, 2009). Various risks, such as climate change, seasonal nature of agriculture, modernised technology, excessive division of agricultural land, perishable nature of agricultural products, and fluctuation in demand and prices for agricultural products have an influence on farmers, as it will affect their ability to repay their credit (September, 2009).

2.2.1. ROLE OF CREDIT IN THE SOUTH AFRICAN AGRICULTURAL SECTOR

In South Africa credit is provided by informal organisations, formal organisations, land and development banks and by other organisations. Bank credit is known as the borrowing capacity provided to a farmer, individual or organisation in the form of a loan by the financial organisation. These organisations are important to the economy, as they make credit available to investors who have profitable ideas. Credit is important in the South African agricultural sector – this is indicated by the increase in capital assets and investments that contribute to an upward trend in farming debt. Farmers require credit to finance capital assets and investments, this increase in demand for credit has caused the total farming debt to increase, as demonstrated Figure 2.1. Access to credit is important for the agricultural industry, and Figure 2.1 demonstrates the role of credit over a ten-year period.

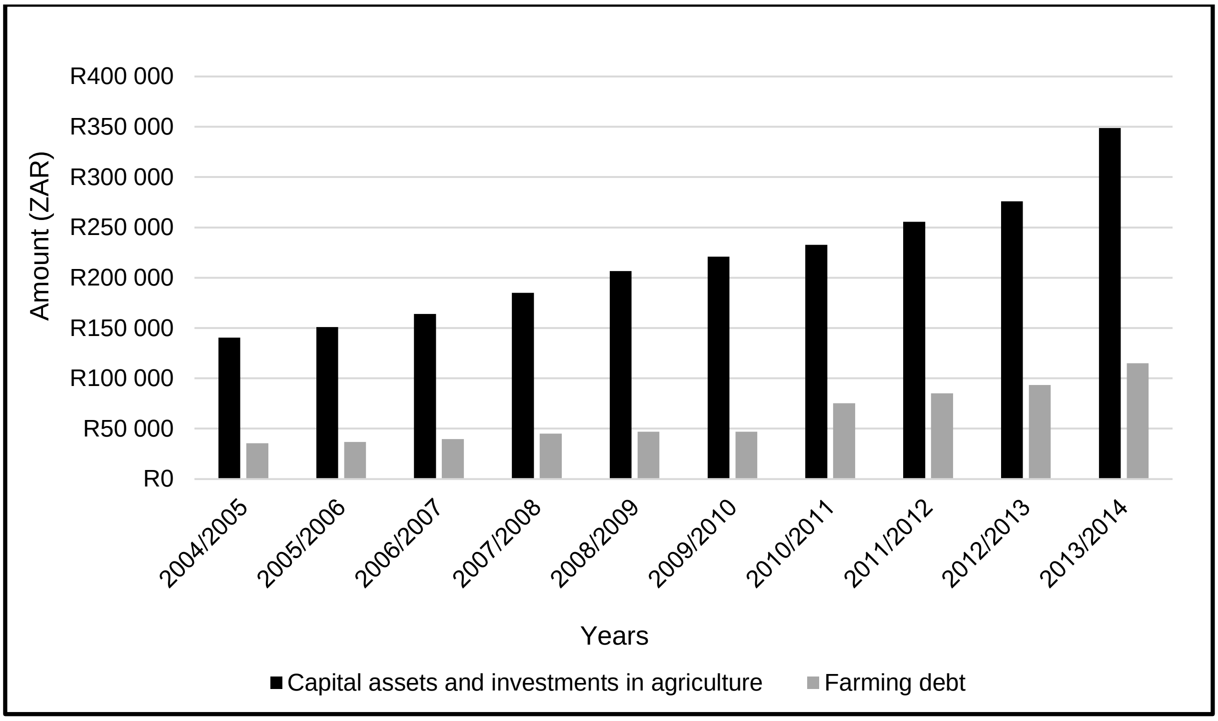

Figure 2.1: Farming debt vs capital investment June 2004 to July 2014

Source: (DAFF, 2006; 2007; 2008; 2009; 2010; 2011; 2012; 2013; 2014)

From 2005 to 2015, the nominal cost of intermediary agricultural inputs, such as fertilisers, diesel, seed and chemicals, increased by 247%, which implies an increasing rate of 30.8% per annum (ABSA, 2015). As production costs are largely financed, it is expected that the demand for working capital will continue to grow (ABSA, 2015). In the event of crop failure in 2015, a farmer would take approximately three years to repay this debt, compared to 1981, when farmers took approximately two years to repay the debt. This implies that a producer is not likely to recover after a disastrous production year without debt restriction or risk mitigation, such as production insurance or crop insurance, which is particularly important for high-risk areas (ABSA, 2015).

Capital assets and investments have increased considerably from 2004/2005 to 2013/2014; this is caused by the increase in demand for machinery, implements and vehicles (DAFF, 2006; 2007; 2008; 2009; 2010; 2011; 2012; 2013; 2014); this could have a direct influence on agricultural debt. This constant increase in capital assets and investments has resulted in the constant increase in agricultural debt illustrated in Figure 2.1.

Since 2004/2005 agricultural debt has increased year on year. Agricultural debt increases from 2009/2010 this may have been caused mainly by changes in values of the livestock industry, vehicles, fixed improvements and machinery (DAFF, 2009). The increase in accumulated debt was also exacerbated by the 2009 economic recession, which affected many farmers' debt-repayment ability. Two major events impacted the agricultural sector in 2014, namely changes in land ownership and weakening of the rand against major global currencies (ABSA, 2015). These events provided additional risk and opportunities for the agricultural trade, which affected input costs, such as fuel, fertilisers, seed and equipment, which are highly correlated with the rand/dollar exchange rate (ABSA, 2015). The fluctuating exchange rate had an influence on input production costs, which increased the production costs for crop farmers. Currently drought is ravaging several sub-Saharan African countries, which has resulted in crop damage and culling of livestock (SACAU Outlook, 2016). The drought has caused crop quality to decrease, which results in a lower price and a decrease in crop yield. The decrease in crop yield and price means input costs are higher than the output production.

When credit is made available, banks are able to provide a social service, expand capital investment and improve living standards (Adekanye, 1986). The success and failure of a financial organisation is not only related to credit risk, but also to its ability to manage reputation risk, operational risk, liquidity risk, market risk and legal risk. Therefore, financial organisations have become more aware of the need to improve credit-evaluation procedures (Salame, 2011).

2.3. CREDIT EVALUATION

Agricultural businesses are characterised by cyclical performance, seasonal production patterns, high capital intensity, annual payments of agricultural loans, leased farmland and involvement in government programmes (Katchova & Barry, 2005). Due to these characteristics, agricultural lending losses may not be frequent, but may be large, depending on the performance of the farm (September, 2009). Therefore, the main aim of credit evaluation is to increase return with the lowest risk (BiiiCPA, 2015). During the evaluation process the analyst categorises applicants into two groups, known as "good credit" and "bad credit". Leea, Chiub, Luc and Che (2002) mention that the acceptable applicants are likely to repay the financial obligation and be accepted. The "bad credit" applications are likely be rejected due to the high possibility of default. When bad credit is accepted it leads to lower bank revenue and loss in bank capital, consequently it causes an increase in bank losses, which can lead to bankruptcy or insolvency (Abdou & Pointon, 2011). As the decision-making process has an influence on the financial organisation, it is important for financial organisations to make the correct decisions when evaluating credit applications. Different approaches are available that can be used individually or complementary, to assist financial organisations and credit officers in the decision-making process.

2.3.1. CREDIT-GRANTING APPROACHES

Two credit-granting approaches can be used to determine the repayment ability of credit applicants. These approaches are known as the subjective and objective approaches.

2.3.1.1. SUBJECTIVE APPROACH

The subjective approach is performed on a judgmental basis, by the credit analyst determining the creditworthiness of the applicant based on personal knowledge and experience (Marqués *et al.,* 2013). The subjective approach suffers from inconsistent decisions and inaccuracy, which are made by different credit analysts for the same application (Marqués *et al.,* 2013). This approach also suffers from high training costs that occur when the credit analyst must undergo training before he/her can approve an applicant. Due to these shortcomings, increased demand for credit and development of computer technology, financial organisations have been encouraged to explore objective approaches and to attempt to predict the probability of default accurately (Marqués *et al.*, 2013).

2.3.1.2. OBJECTIVE APPROACH

The earliest financial tools were developed in 1950 by mail-order institutions and United States retailers for the purpose of risk evaluation (Abdou, Pointon & El-Masry, 2008). A statistical credit-scoring model is a quantitative evaluation technique used by financial organisations to evaluate the creditworthiness of applicants or firms that apply for loans (Abdou *et al*., 2008). The aim of a statistical credit-scoring model is to correctly classify credit applicants in accepted or rejected groups.

Financial organisations use statistical credit-scoring models for loan processing and pricing, credit monitoring, calculating inputs and decision-making management (Bandyopadhyay, 2007). These statistical models have been used to issue credit cards, auto loans and mortgage loans (Mester, 1997). This has stimulated remarkable growth in the consumer credit industry over the past few decades (Abdou *et al*., 2008). Without these statistical models, lenders would not have been able to improve their performance (Abdou *et al*., 2008). Statistical models provide the credit analyst with tools to help with the decision-making process (Abdou & Pointon, 2011) and have the ability to reduce human judgment, and improve consistency and accuracy (Marqués *et al.,* 2013). Statistical credit-scoring models can improve cash flow, evaluate the credit risk, support management decisions and reduce probability of default (Thomas, Edelman & Crook, 2002). Therefore, a statistical credit scoring model that has a high percentage of correctly classified applicants needs to be identified.

❖ STATISTICAL CREDIT-SCORING MODELS

There are various statistical models that are available and have been applied to credit research. Due to contradictory results there is no overall best statistical model for creating credit-scoring models (Abdou, Pointon & El-Masry, 2007). The success of the various models depends on the characteristics used, facts about the problem, the data structure and the extent to which classes can be separated by the objectives and characteristics in the classification (Hand & Henley, 1997). Logistic regression and PA have results comparable to that of sophisticated models. When building the scoring models new users must ensure that the most suitable techniques from the selection of models are available, keeping in consideration the differences between various methods (Desai, Crook & Overstreet, 1996; Hand & Henley, 1997; Ong, Haung & Tzeng, 2005), and the importance of a binary variable of “good” and “bad” (Desai *et al.,* 1996; Banasik, Crook & Thomas, 2003; Yang, Wang, Bai & Zhang, 2004).

International research has explored and applied various statistical models that can be used to improve the accuracy of evaluation of credit applicants, where the variables are selected from the applicants and not from the financial organisation. In the South African agricultural sector, less research has been performed on the accuracy of the credit evaluation in predicting high-risk loans. Abdou *et al.* (2007) state that, in a new banking environment, it would be suitable to first explore some of the traditional techniques, such as PA and LR. In credit research, the LR and PA are usually used with other statistical models for the purpose of comparing results (Abdou *et al.,* 2008). Furthermore, the PA is considered to be a successful alternative to the LR (Oriema, 2010).

In comparison, Limsombunchai *et al.* (2005) claim that NNs provide the best models for screening agricultural applications, as they have the lowest misclassification costs. Abdou *et al.* (2008) deduced that NNs have the highest average correct classification rate and the lowest estimated misclassification costs. Neural network models that are trained by a back-propagation learning algorithm outperformed those involving multiple discriminant analysis, linear discriminant analysis, decision trees, LR and k-nearest neighbours (Tsai & Wu, 2008).

Thus, PA, LR and NN have been used in credit scoring and are successful in terms of prediction accuracy. To gain an understanding of the statistical models used in the agricultural credit sector, the advantages, disadvantages and functioning of these identified models are discussed next.

PROBIT ANALYSIS

The PA can be used to determine factors that influence the probability that a farmer will default (Quaye *et al.,* 2015). This model finds the probability unit value of the binary coefficients and was specifically designed to investigate dependent variables in the regression (Abdou *et al.,* 2008). A linear combination of independent variables is transformed from a normal distribution into its cumulative probability value, which equals 0 or 1 (Abdou *et al.,* 2008). The model reduces the constraint that the effect of the independent variables is constant across different predicted values of the dependent variable. In small samples the PA has advantages over LR (Anang, Sipiläinen, Bäckman & Kola, 2015). This models assumes that only the values of 0 and 1 are observed for the dependent variable. The estimates are determined by the use of the probit function (Abdou & Pointon, 2011)

The PA has similar advantages and disadvantages as the LR, therefore, it can be used as an alternative (Abdou *et al.,* 2008). The main difference between the PA and LRs is the cumulative distribution function. The PA makes use of the standard normal distribution and the LR makes use of the logistic distribution to determine the distribution function.

LOGISTIC REGRESSION

Logistic regression is a commonly used statistical model, where the probability of the binary outcome (0 or 1) is associated with the independent variables used. The procedure estimates the coefficients of the linear equation to determine the probability of odds ratio for each independent variable. The linear combination of independent variables is coordinated by the log of the probability odds. The objective of the LR in credit scoring is to determine the restricted probability of the characteristics by means of the information provided on the credit applicant (Lee & Chen, 2005).

The LR has the capability to predict default of the applicant and identify the characteristics related to the applicant's behaviour (Li & Zhong, 2012). Logistic regression is able to remove redundant variables, identify relationships that are invisible and take into consideration the correlation between variables (Li & Zhong, 2012). It is also able to analyse variables simultaneously and individually and the user can verify the sources of error and optimise the model (Li & Zhong, 2012). The LR has dominated literature and has been used widely due to its simplicity (Limsombunchai *et al.,* 2005). The LR can be interpreted easily in terms of the odds ratios, and this is an advantage over the PA. Another advantage is that sampling of the independent variables only change the constant of the LR (Tufféry, 2011, 478).

Logistic regression, however, also has some disadvantages. The preparation of the variables takes a long time and the credit analyst must use pre-selection to determine the more important variables when there are numerous variables; independent variables must be linearly independent; and the approach is not able to handle missing values of continuous variables. This model is only able to handle missing values when the data is divided into classes and the missing data is divided into groups (Tufféry, 2011: 477-478). This model is also sensitive to extreme values of continuous variables (Tufféry, 2011: 477-478). Despite the disadvantages it has been found that the LR model is a good substitute for NN, as it is more accurate in some cases (West, 2000). Thus, this model can be used as a successful alternative for NN and has demonstrated high accuracy.

NEURAL NETWORKS

Neural networks attempts to replicate the functioning of the human brain (Abdou *et al.,* 2008). The neural network consists of many inputs, known as independent variables that are multiplied by a weight. Information is then summed up and transformed into a neuron. The result is then processed and it becomes the independent variable for another neuron (Thomas *et al.,* 2002). Techniques such as training and operation modes are used to recognise patterns and learn from its mistakes (Stergiou & Siganos, s.a.). The training operation mode is defined as the ability of the neuron to be trained to use or not to recognise the taught input pattern, so the associated outputs become the current input (Stergiou & Siganos, s.a.). This model has been designed and is ideally suited for agricultural data modelling, which is often complex and nonlinear (Sharma & Chopra, 2013).

There are several types of NN that have the ability to outperform prediction accuracy of traditional models. Two most widely used NNs are known as feed-forward and back-propagation NNs (Sharma & Chopra, 2013). In a feed-forward NN, the information moves forward in one direction while connecting pathways, from the input layer through the hidden layers to the final output layer. There are no feedback loops, as the output from each layer does not affect the same or next layer. In a back-propagation neural network there is at least one feedback loop, therefore, there are neurons with self-feedback links. This means that the output of the neuron is fed back into itself as an input (Sharma & Chopra, 2013). The back-propagation neural network is most frequently used (Thomas *et al.,* 2000).

The feed-forward and back-propagation NNs can be either a single-layer or multi-layer NN. A multiple layer NN consists of an input layer, more than one hidden layers and an output layer (Stergiou & Siganos, s.a.). Each neuron in each hidden layer has a set of weights applied to its input or independent variable. This may differ from those applied to the same independent variable entering different neurons in the hidden layers (Thomas *et al.,* 2002). The outputs from each neuron in the hidden layer have weights applied and become the inputs for the neurons in the next hidden layer. Once the output layer determines a value or total it is then compared with the average total cut-off score. The output layer provides a result that is used to predict if the credit applicant will be accepted or rejected. A back-propagation learning algorithm uses gradient descent to adjust the weights, so to minimise errors between the network output values and targeted output values (Limsombunchai *et al.,* 2005). It has been found that one hidden neuron is sufficient to provide the model with the desired accuracy (Baesen, Van Gestel, Viaene, Stepanova, Suykens & Vanthienen, 2003; found in Alaraj, Abbod & Hunaiti, 2014).

The main advantage of this NN model that it can map input patterns to the associated output patterns; it is a robust system that is fault tolerant and therefore able to handle incomplete, noisy or partial patterns (Alaraj *et al.,* 2014). This model processes information in parallel at high speed and in a distributed manner, it is able to recognise complex patterns between variables and does not require prior assumptions about the distribution variables (Eletter, Yaseen & Elrefae, 2010; Alaraj *et al.,* 2014).

The disadvantage of the NN is that it lacks explanatory capability, as it is not able to give an explanation as to why the loan was accepted or rejected (Salame, 2011). The NN's results are improper, as the values will be changed until they become proper and acceptable (Oden, Featherstone & Sanjoy, 2006). Moreover, the decision of topology is important and the problematic long training processes are criticised (Alaraj *et al.,* 2014). Neural networks are complex and often at risk of over-training (Tufféry, 2011: 499). Due to over-training it is not able to extract the subset of the most relevant variables from the set of all the potential variables. Prediction accuracy might be affected, as there is no official method to select the appropriate parameters (Alaraj *et al.,* 2014).

DISCUSSION

Models that were identified were compared according to various advantages and disadvantages, to determine the models that are best suited and adapted for the specific problem. An LR approach has been applied to agricultural financial organisations, has a high prediction accuracy rate and is a good alternative to NN, as it is more accurate, in some cases (West, 2000). The PA is able to find the ability to predict default accurately and take into consideration the correlation between variables. Probit analysis has also been selected, as it is considered as a successful alternative to the LR (Oriema, 2010). Neural networks will be used, as it is ideally suited for agricultural data modelling, which is often complex and nonlinear (Sharma & Chopra, 2013). Neural networks are able to continuously learn and recognise complex patterns. Neural networks have been used to predict the likelihood of agricultural applicants defaulting, and they demonstrate a high accuracy compared to other models. These models have also been used to reduce misclassification of agricultural applicants.

2.4. MISCLASSIFICATION

Misclassification means high-risk applicants can be unfairly selected and low-risk applicants can be unfavourably denied. These errors can be costly to a financial organisation, which is explained by the influence on the overall performance of profitability of the loan portfolio. There are two types of misclassification errors, known as Type I and Type II errors. Type I error involves denying a loan to an applicant who is able to repay the loan obligation (Nayak & Turvey, 1997). Type II errors refer to loans that are granted to applicants who have a high probability of defaulting on the loan repayment. Both these misclassification errors cause the lender to lose expected profits. Type I errors are less costly to a financial organisation than Type II errors.

If approved, high-risk applicants (Type II error) default on the specified obligations, which leads to lower bank revenue, loss in bank capital and, subsequently, increases in bank losses, which can ultimately cause bankruptcy or insolvency (Abdou & Pointon, 2011). Associated with Type II errors include loss of principal and interest on principal during the period of litigation and foreclosure (Nayak & Turvey, 1997). Various indirect costs, such as insurance coverage, legal fees, administration and property taxes, which may not be fully recoverable, also contribute to loan losses (Nayak & Turvey, 1997).

The classification matrix is a popular method for evaluating measures of misclassification, as shown by Paliwal and Kumar (2009). The classification matrix shown in Table 2.1 classifies the credit applications according to four categories: Good/good (Gg), Good/bad (Gb), Bad/good (Bg) and lastly Bad/bad (Bb). In Table 2.1, G represents the actual good observations and g is a statistical-model-predicted good outcome. B is actual bad observations and b is the statistical-model-predicted bad outcome. Gg indicates that the statistical model predicted a good outcome while the actual observations were also good. Gb is actual good observations while the statistical model predicted a bad outcome. Bg is actual bad observations, which the statistical model predicted as good and Bb is actual bad observations that were predicted as bad by the statistical model (Abdou *et al.,* 2007). TG, total actual good observations, is followed by TB, total actual bad observations. Tg refers to the total good observations that were predicted by the statistical model. The total predicted bad observations generated by the statistical model is represented as Tb, and TN is the total number of actual observations, representing the total amount of actual good observations (Abdou *et al.,* 2007). The Type I error (Gb) can be explained as rejecting a loan that must be granted and a Type II error (Bg) can be explained as granting a loan that must be rejected, as set out in Table 2.1.

Table 2.1: Classification matrix to evaluate the accuracy and misclassification of credit models

		Model testing		
		good (g)	**bad (b)**	**Total**
Actual observations	**Good (G)**	Gg	Gb	**TG**
	Bad (B)	Bg	Bb	**TB**
	Total	**Tg**	**Tb**	**TN**

Source: Abdou *et al.* (2007)

The high-risk applicants in Table 2.1 are known as Bg, where the statistical model predicted good, but, in reality, the borrower defaulted. Du Jardin (2012) mentions that the selection of variables included in credit-scoring models will have a significant effect on the accuracy of classifying applicants. The next section provides a review of different characteristics used in credit-scoring models to predict the applicant's ability to repay a loan.

2.5. CHARACTERISTICS USED IN STATISTICAL CREDIT-SCORING MODELS TO PREDICT REPAYMENT ABILITY

Research often considers factors that influence access to credit by gathering information from farmers. This information is often not obtained from commercial or agricultural banks, but rather from the client (Henning & Jordaan, 2015). Credit research indicates that various variables are included when financial organisations evaluate agricultural loan applicants in South Africa. Typically, financial organisations evaluate the applicants according to the 5 Cs, which include character of borrower (reputation), collateral, capital (leverage), capacity (volatility of earnings), and condition (macroeconomic cycle) obtained (Bandyopadhyay, 2007). The 5 Cs are widely documented to be a good indicator of the ability of a person to repay a loan. Each of the 5 Cs consists of many sub-divided components, which are used collectively to categorise a new applicant.

The credit analyst in the financial organisation analyses characteristics, such as the character of the borrower and collateral (Culp, 2013). Financial organisations evaluate the character of the borrower by analysing characteristics, such as gender, age and marital status (Marqués *et al.*, 2013), number of dependants, education level, occupation, loan duration, monthly income, loan amount, house ownership, bank accounts, monthly income, purpose of loan, and date of first business account (Steenackers & Goovaerts, 1989; Leea *et al.,* 2002; Banasik *et al.,* 2003 Chen & Huang, 2003; Hand, Sohn & Kim, 2005; Sarlija, Bensic, & Zekic-Susac, 2009; Sustersic, Mramor & Zupan, 2009; found in Abdou & Pointon, 2011). Credit analysts place significant emphasis on the borrower′s personal characteristics (i.e. integrity, production management ability and honesty) and financial information, when making decisions about the approval of credit applicants and the required level of credit (Olagunju & Ajiboye, 2010). Henning and Jordaan (2015) found that a South African financial organisation evaluates the following borrower characteristics: farmer’s age, date of first business (loyalty), farmer’s experience, education/qualification and sustainability of the enterprise.

Capital and capacity are used to evaluate the financial performance of the borrower’s enterprise according to financial ratios. Capital refers to funds that are available to operate farm businesses – this is determined by reviewing balance sheets and other financial ratios, which include cash-flow generation (Henning & Jordaan, 2015). Capacity refers to the applicant’s ability to repay the loan and to bear the financial risk of the loan. Historical projected profitability and farm cash flow is used to measure repayment capacity (Henning & Jordaan, 2015). The financial ratios include applicant’s liquidity (net working capital, quick ratio and current ratio), solvency (debt-to-equity ratio and leverage ratio), profitability (return on assets (ROA) and return on equity), repayment capacity (interest coverage, interest expense ratio and debt repayment ratio) and efficiency (capital turnover and gross ratio) (Limsombunchai *et al.*, 2005). According to Henning and Jordaan (2015), South African financial organisations assess applicants according to past and current financial information (liquidity, solvency, profitability, financial efficiency and repayment ability). Financial information, gathered through financial analysis, can be refined further to minimise multi-collinearity between the ratios (Durguner, Barry & Katchova, 2006). Therefore, Durguner *et al.* (2006) include the following ratios: working capital to gross revenue (WCTGR), net worth, ROA, asset turnover ratio (ATO), depreciation expenses ratio and operating expense ratio.

DTA and debt-to-equity ratios are all mathematically equivalent, therefore only one of the ratios need to be used (Blocker, Ibendahl & Anderson, 2010). The DTA ratio has been selected as it provides an indication if there is sufficient collateral available to cover the debt. Working capital to gross revenue, ROA and ATO were chosen to minimise multi-collinearity between the ratios (Durguner *et al.,* 2006). The net farm income ratio was used instead of net worth, as ratio measurements eliminate the economies of scale (Hoppe, 2015). Therefore, a more realistic comparison of farm performances against one another can be observed (Nieuwoudt, 2016). The cash-flow ratio is included in this research as it is considered as an important ratio by financial organisations. It demonstrates how much cash flow is required to cover production costs. The production expenses utilised during farm production are demonstrated by using the production-cost ratio. The production-cost ratio provides an indication of the amount of production cost used over the total sales and therefore this ratio was used instead of the operating-expense ratio. Other factors, such as account standings and credit record, are also evaluated by South African financial organisations and other financial organisations (Henning & Jordaan, 2015).

Collateral represents the security agreement that the serves as a final source of repayment to the lender should the borrower default on the loan agreement. Financial organisations carefully select the risk-of-return relationships of the loan request – the risk increases with larger amounts and/or higher quality collateral (Henning & Jordaan, 2015). Agricultural collateral information, such as collateral and farm ownership, are assessed by South African financial organisations (Henning & Jordaan, 2015).

Lastly, the credit analyst needs to consider agricultural conditions, which refer to the intended purpose of the loan. Factors that are considered are cyclical performance, seasonal production patterns, farm typography, commodity, geographical location, and participation in government programmes, lease of farmland, high capital intensity and annual payments of loans (Kim, 2005; Bandyopadhyay, 2007). South African financial organisations observe condition characteristics, such as type of farming enterprise, associated industry risk, loan amount and use of fund repayment terms (Henning & Jordaan, 2016).

2.6. CONCLUSION

Based on the literature reviewed, it is evident that the agricultural industry is reliant on credit, as credit enables the famer to expand a business to its maximum potential. Credit ensures that farmers can continue with production farm activities with borrowed capital (production loans) and can repay the debt after production has been completed. In terms of access to credit, the agricultural sector does not differ much from other sectors, however, this sector is influenced by different factors, which may influence the repayment ability of the applicants in the sector. Credit can also be used to revive economic activities that have suffered from setbacks caused by natural disasters or unforeseen weather patterns (Ademu, 2006).

There are two approaches, subjective and objective, that can be used to evaluate credit applications. The subjective approach is reliant on knowledge and the experience possessed by the analyst to determine repayment ability. The objective approach provides the credit analyst with a tool to help with the decision-making process (Abdou & Pointon, 2011), has the ability to reduce the role of human judgment, and improve consistency and accuracy (Marqués *et al.,* 2013). Various statistical models have been used in the objective approach for credit research, however, due to contradictory results there is no overall best model. The selection of the model depends solely on the details of the problem, data structure and characteristics used. The LR approach has been applied to agricultural financial organisations, and has dominated literature due to its simplicity. This model also has a high prediction accuracy rate and is a good alternative to NN, and in some cases, more accurate (West, 2000). The PA is a model that is considered to be a successful alternative to the LR. This model also has the ability to predict accurately and can take into consideration the correlations between variables. Neural network demonstrates high accuracy compared to other models and is ideally suited for agricultural data modelling, which is often complex and nonlinear (Sharma & Chopra, 2013). An advantage of this model is its ability to continuously learn and recognise complex patterns. This research selected the above-mentioned models to predict successful agricultural applicants and reduce misclassification costs.

Literature also shows that numerous characteristics influence the ability to repay the loan. The characteristics were selected according to the 5 Cs of credit, which are characteristics of the borrower, collateral, conditions, capital and capacity. This framework was selected to determine which characteristics are important to use when granting credit to farmers. Often, credit research considers factors that influence access to credit by gathering information from farmers instead of obtaining the information from commercial or agricultural banks (Henning & Jordaan, 2015). Numerous variables were identified to be used in statistical models for different purposes. The variables used to predict default of agricultural applicants in this research include purpose of the loan, amount, period of repayment, date of first business, credit history, collateral, financial information (WCTGR, DTA, ROA, net farm income ratio, ATO, production-cost ratio and cash-flow ratio) (Durguner *et al.,* 2006; Henning & Jordaan, 2015), farm diversification (enterprises available on the farm), industry risk association, ownership, age of applicant, years of farming experience and education. These variables were selected in accordance with variables considered to be important by South African financial organisations.

CHAPTER 3

DATA AND METHODS

3.1. INTRODUCTION

Chapter 3, firstly, provides a description of the data that was used in this research. Secondly, the characteristics of the respondents are described to indicate the distribution of the characteristics found in the data set. Lastly, the three selected methods, LR, PA and NN, are described regarding their ability to predict high-risk loan applicants in the agricultural industry.

3.2. DESCRIPTION OF DATA

This research is based on data collected by Henning (2016). Data collected for this research considers a specific financial organisation that is involved in the agricultural sector. To ensure the accuracy and relevance of the data, a formal agreement was reached with the financial organisation, which agreed to provide the researcher with credit application information from actual applicants and the classification decision made by the organisation. The agreement stipulated that all the data obtained from the organisation had to remain confidential, and that no personal information (i.e. individual names or business names) that could be used to identify the relevant clients, would be made available. A total of 127 credit applications were obtained (between July 2015 and December 2015) by the researcher with the assistance of the financial organisation. The data includes observations from several provinces in South Africa. The variables included in this research were confirmed by the representative of the organisation as important to their credit classification. The variables considered in this research are similar to the variables identified by Henning and Jordaan (2016). The information was obtained and confirmed as capturing the relevant information in the classification decision-making by individuals from the financial organisation.

3.3. CHARACTERISITICS OF RESPONDENTS

A total of 127 applicants were observed. The following section provides information on the observed applicants. These variables include the purpose for which the loan was required, amount of credit required, period of repayment of loan, business loyalty, credit history and collateral. Financial information of the farm was also considered in terms of ratio measures, such as solvency (DTA ratio), liquidity (WCTGR), profitability (ATO, ROA, net farm ratio) and efficiency (production costs, cash-flow ratio), diversification on the farm, namely, the number of enterprises on the farm, and associated industry risk as categorised by the organisation. Personal information about the farmer included ownership, age of farmer, years of farming experience and education/qualification. The dependent variable is binary, as it takes on the value 1 when an application is approved or 0 when rejected.

3.3.1. LOAN PURPOSES AND APPLICATION PERIOD

Loans in the agricultural sector are used for different purposes. In some instances, loans are used to access inputs for the production process, to finance assets, such as machinery and equipment, or to buy land (DAFF, 2015). These loans can be categorised according to the repayment period, such as, short-term production loans and overdrafts, and medium-term loans for machinery and equipment (i.e. vehicles, tractors, plough and harvesters) and breeding livestock. Long-term loans are mostly used to purchase agricultural land. To ensure that there are sufficient observations in each purpose category, the categories were identified as short, medium and long term. For discussion purposes, three categories were created to ensure that there were sufficient respondents for each category. However, continuous variables were used for statistical modelling purposes.

Loan applications for working capital, production costs and increasing overdrafts are categorised as short-term loans. The short-term loan category consists of loan applications for periods between 1 and 12 months. Medium-term loans consist mostly of loan applications that have a duration of between 12 and 120 months. These loans include loans to obtain farm machinery and vehicles, farm development, to acquire livestock, and for diversification activities. Long-term loans are required for periods longer than 120 months, and these loans are generally needed for farm and property purchases. The distribution between the categories of loans is shown in Table 3.1, which demonstrates that most of the loan applications were medium-term loans (60), followed by short-term loans (43).

Table 3.1: Distribution of loan applicants in short, medium and long-term categories

Loan purposes and application period	Total respondents n = 127
Short-term loans (0 - 12 months)	43
Medium-term loans (13 - 120 months)	60
Long-term loans (121 - 240 months)	24
Longest period	180 months
Shortest period	2 months
Average period	85 months

The longest repayment period in the data is 180 months, and the shortest period is 2 months; the average repayment period is 85 months for the 127 loan applicants. According to previous research, the longer the repayment period, the more likely the applicant is to repay the loan (Awunyo-Vitor, 2012). The possible reason for this is that the longer period relates to smaller annual or monthly payments. This has a smaller influence on current cash flow, however, it does influence the total repayment amount owed.

3.3.2. LOAN SIZE

For the purpose of this discussion the variables are categorised into smallest, largest and average loans, however, continuous variables are used for statistical modelling. Table 3.2 demonstrates that short-term loans involve the largest amounts, compared to medium-term loans, which involve the smallest amounts, specifically in the average and largest category. Short-term loans involve the highest average amount compared to average-sized medium-term amounts. This demonstrates that farmers require more finance for short-term loans (production activities) than they do for medium-term loans.

Table 3.2: Distribution of the largest, smallest and average loan sizes for short, medium and long-term categories

Loan size	Smallest	Largest	Average
Short-term	R 0.00	R 32 000 000.00	R 4 663 617.95
Medium-term	R 200 000.00	R 23 000 000.00	R 4 157 865.00
Long-term	R 2 100 000.00	R 52 000 000.00	R 12 775 000.00

These short-term loans are usually repaid at the end of the production season from the income that has been generated from the sale of the product. According to Table 3.1 more applicants apply for medium-term loans, but, on average, the size of the loan applied for is smaller than the average short-term loan, as indicated in Table 3.2. The long-term category consists mainly of farm purchases, demonstrated by the astronomical amounts. These loans are repaid over a period of 120 to 180 months. The smallest short-term loan of R0.00 refers to clients who are restructuring their finance. It was found that the larger the loan size, the lower the probability of repayment default (Awunyo-Vitor, 2012). Thus, as seen in Table 3.2, largest medium-term loan is smaller than the largest long-term loan. Short-term loans are larger than medium-term loans, as short-term loans go towards input costs, which are needed for production inputs (production loans). The cost of production input costs, such as fertiliser, diesel, seed and chemicals, has increased over time, and this may have an influence on the size of the loan (short-term loans) applied for.

3.3.3. BUSINESS LOYALTY AND CREDIT HISTORY

3.3.3.1. Business loyalty

It is important for the financial organisation to have a good relationship with the credit applicant, as this provides various advantages to the credit applicant. The client will continue to do business with the financial organisation if the client is satisfied with the manner in which the business has been conducted. For the purpose of this discussion the variables are categorised, however, statistical modelling of the variables was kept continuous. The number of years the client has been with the financial organisation demonstrates the loyalty of the client, shown in Table 3.3.

Table 3.3: Number of years the client has been with the financial organisation

Years in business with financial organisation	Number of respondents n=127
0	25
1 – 15	49
16 – 30	36
31 – 45	12
46 – 60	5
New applicants	25
Longest period	60 years
Shortest period	0 years
Average period	14 years

There were 25 applicants who were clients of other financial organisations, hence these clients do not have a reputation record with the new financial organisation (illustrated by category 0). Most of the applicants (49) have been in business or have had an account with the financial organisation for 1 to 15 years. Only 5 applicants have had accounts with financial organisations for more than 46 years; the longest period is 60 years. This indicates that most of the applicants have been doing business with the financial organisation for some time, which indicates that these applicants have a reputation record.

3.3.3.2. Credit history

Credit history is divided into two categories, namely, accepted and other, shown in Table 3.4.

Table 3.4: Description of credit history of credit applicants

Description of credit history	Number of respondents n = 127
Acceptable	115
Other	12

*Other includes not acceptable and absence of credit history

As indicated in Table 3.4, 115 respondents were considered to have acceptable credit history by the financial organisation. The other 12 respondents were considered as either undetermined or not acceptable. The clients who are classified as undetermined in the “Other” category have no records with the financial organisation. Limsombunchai *et al.* (2005) found that an applicant who has a longer relationship with a financial organisation has a higher probability of defaulting on debt repayment, and this has a significant negative influence on the repayment ability of an applicant. Previous research indicates that clients with a poor credit history are associated with loan delinquencies (Addae-Korankye, 2014). Commercial banks prefer accepting farmers who have proven track records (Abdesamed & Wahab, 2014). New clients have no credit history with the financial organisation or in some cases lack a formal credit history indication.

In the “Other” category, 12 applicants were considered as either undetermined or not acceptable. Further investigation into the 12 applicants from the “Other” category showed that 6 were new applicants while the other 6 were existing clients in the financial organisation. From the 6 new clients, 3 clients were denied credit and the other 3 were approved by the financial organisation. Out of the 6 existing clients in the financial organisation, 3 applicants were rejected and 3 were approved. This could have an influence on the outcome of this variable, meaning that if the applicants have a poor credit history and have been rejected by the financial organisation this would show a positive influence on acceptability in a financial organisation. However, if the applicants have an undetermined credit history and were accepted it could have a positive influence on the repayment ability.

3.3.4. COLLATERAL

Collateral is considered to be a security measure for the financial organisation if the applicant is unable to repay the loan. The loan appraisal process in a formal financial organisation is based on bankability of farming enterprises and their heavy reliance on “traditional collateral” (Qwabe, 2014). It has been found that collateral has a positive influence on debt repayment ability (Kohansal & Mansoori, 2009; Anigbogu, Onugu, Onyeugbo & Okoli, 2014). Commercial banks prefer farmers or business firms that offer collateral in the form of hard assets (Abdesamed & Wahab, 2014). The financial organisation can repossess the land or the machinery to collect the unpaid funds due in terms of the loan agreement. In Table 3.5 two categories are used to discuss the collateral status of the loan application. The first category is termed secure. This category indicates that the applicant has sufficient collateral for the loan requested. The second category is an indication that the clients do not have sufficient collateral available and should therefore be required to provide extra collateral or another form of security, such as enterprise diversification.

Table 3.5: Indication whether respondents' collateral is sufficient

Collateral	Number of respondents n = 127
Secured	124
Not secured	3

As shown in Table 3.5, 124 applicants have sufficient collateral available and are therefore secured. Three applicants do not have sufficient collateral available, therefore, they are considered as not secured. Not secured applicants would need to provide additional collateral or different measures to mitigate risk in their farming enterprise before receiving the loan. The more collateral the farming enterprise can provide the more security there is for the financial organisation to grant the loan. There are different measures that can be used to mitigate risk, for example, enterprise diversification and production insurance.

3.3.5. FINANCIAL PERFORMANCE INDICATORS

Loans must be repaid with finance returns from farming operations. Therefore, it is important to consider the financial performance of the business. The performance is measured by a financial analysis consisting of different ratios. Financial ratios are used to determine the financial worth or strengths and weaknesses of the applicant. The financial information is gathered from various financial statements: statement of comprehensive income, statement of financial position, as well as projected and current cash-flow statements. Financial information is of great importance, as it is used to determine if the applicant will have the capacity to repay the loan according to stipulated agreement terms and conditions. Financial ratios are used to eliminate the problems that arise when each business is evaluated according to values, and not in proportion to each other. Financial ratios are divided into various categories, such as solvency, liquidity, profitability and financial efficiency. These categories are further divided into categories such as solvency (DTA), liquidity (WCTGR), and profitability (net farm income (NETFARMRATIO), ROA, production-cost ratio (PRODCOST), financial efficiency (ATO) and cash-flow ratio (CASHFLOW)). For the purpose of this discussion the variables are categorised, however, continuous variables are used for statistical modelling. The maximum, minimum and average financial ratios are tabulated in Table 3.6.

Table 3.6: Summarised financial ratio indicators indicating the financial performance of the applicants

Financial ratios	Maximum	Minimum	Average
DTA	1.7505	0.0000	0.3535
WCTGR	6.3816	-2.3668	0.2485
NETFARMRATIO	2.2666	-0.1383	0.3352
ROA	1.9683	-0.0089	0.1330
PRODCOST	4.9749	0.0000	0.6628
ATO	1.9683	-0.0089	0.1330
CASHFLOW	2.3900	0.0000	1.1297

The DTA ratio compares the farm debt obligations owned by the farmer to the value of the farm assets (FFSC, 2011). Applicants who have a smaller DTA ratios are more likely to be accepted than those who have high DTA ratios (Quaye *et al.,* 2015). Quaye *et al.* (2015) also found that farmers who have a high net farm income ratio are less likely to be delinquent. Working capital to gross revenue is a better method to determine liquidity than merely making use of working capital, as this ratio takes into consideration the amount of livestock, differentiates between crops and considers the size of the farm (Craven, Nordquist & Klair, 2011). Therefore, the higher this ratio is, the more acceptable the applicant is to the financial organisation, as this applicant is classified in a lower risk category (Durguner *et al.,* 2006).

ROA is determined by the income earned by the business compared to the assets used in the business (Sebe-Yeboah & Mensah, 2014). The higher this ratio is, the more effectively the assets are used, and the farmer is classified as a lower risk (Durguner *et al.,* 2006). Therefore, the applicant is considered to be more acceptable to a financial organisation. The ATO is defined as a measurement of how efficiently the farm assets are used to generate revenues from the farm operation (Blocker *et al.,* 2003). An applicant is more likely to be accepted by a financial organisation if the ATO ratio is higher, as this applicant is considered to be a lower risk (Durguner *et al.,* 2006). The PRODCOST ratio refers to the amount of costs generated by production compared to the amount of revenue generated by production. This ratio has an influence on the repayment ability of the applicant – applicants with a higher production-cost ratio are more likely to be accepted by a financial organisation. Lastly, the CASHFLOW ratio is important as it measures the cash generated from the operations and compares it to total liabilities. Therefore, the lower this ratio is, the lower the financial flexibility and the higher the probability of default (Calomiris, Hubbard & Stock, 1986; Kajananthan & Velnampy, 2014).

3.3.6. ASSOCIATED INDUSTRY RISK LEVEL AND DIVERSIFICATION (NUMBER OF FARM ENTERPRISES)

3.3.6.1. Associated industry risk level categorised by financial organisation

The risk level of the industry is divided into three categories, namely, high, medium and low risk. The various industries were divided into categories according to the various risks. According to Table 3.7, 78 of the 127 applicants are considered to be medium risk by the financial organisation, making this the largest number of respondents.

Table 3.7: Associated industry risk level categorised by financial organisation

Industry projection (risk level)	Number of respondents n = 127	Percentage
High	26	20,47%
Medium	78	61,42%
Low	23	18,11%

To reduce the risk level of a farming business, farmers can have more than one enterprise on the farm, which means the cost of the business is spread over the various enterprises. In agriculture the main purpose of diversification is to reduce the risk of overall return by choosing a mixture of activities that have a low or negative correlation with net returns (Culas & Mahendrarajah, 2005). Diversification is also considered as one of the more common methods to reduce risk and uncertainty (Miller, Dobbins, Prichett, Boehlje & Ehmke, 2004). Therefore, the chance of a larger loss from a given hazard is reduced by having more than one farm enterprise (Miller *et al.,* 2004). Chirwa (1997) found that the degree of diversification is significantly related to agricultural credit repayment ability – more enterprises reduce uncertainty and reduce risk compared to a single enterprise. Therefore, financial organisations may accept an applicant with more than one enterprise as it is seen as a mechanism to reduce risk and spread the cost. In Table 3.8, Diversification 1 includes farms that have only one enterprise, Diversification 2 includes farms with more than have two enterprises and Diversification 3 is farms with three or more enterprises.

Table 3.8: Distribution of number of enterprises

Number of enterprises	Number of respondents n = 127	Percentage
Diversification 1	36	28%
Diversification 2	59	46%
Diversification 3 or more	32	25%

As shown in Table 3.8, the largest proportion (46%) of famers have two enterprises on their farms, to spread the cost risk and income. Interestingly, 28% of the applicants have only one enterprise, with only 25% of the applicants having three or more enterprises on the farms to spread the risk. This indicates that most farmers use diversification as a mechanism to reduce risk and spread cost risk.

3.3.7. CLIENTS' AGE, EXPERIENCE, EDUCATION AND OWNERSHIP

This section is divided into three sections: clients' age and experience; education; and ownership.

3.3.7.1. Age and experience

Characteristics such as age and experience are essential characteristics that the farmer has no control over. Farmers are expected to learn the necessary skills early in their farming life and become more skilful over time (Phelan, 2014). Table 3.9 demonstrates the clients' age distribution. The relation between client's age and experience is that knowledge is gained through experience. It has been mentioned that experience comes with age, which is a non-psychological factor that influences exploiting opportunity decisions (Phelan, 2014, Henning & Jordaan, 2015). The age of a famer has an influence on his or her decision-making procedure. Older farmers tend to hesitate to adopt new, innovative management skills, while younger farmers are more daring (Haden & Johnson, 1989). Age is often found to have a negative influence of the farmer's repayment ability of the loan (Nwankwo, 2004; Onyenucheya & Ukoha, 2007; Oladeebo & Oladeebo, 2008; Nwachukwu, Alamba & Oko-Isu, 2010; found in Ajah, Eyo & Ofen, 2014). However, age was also found to have a positive influence on the repayment ability of farmers (Arene, 1993). For the purpose of this discussion the characteristics are categorised as shown in Table 3.9, however, continuous variables are used for statistical modelling.

Table 3.9: Age distribution of respondents

Clients' age	Number of respondents n = 127	Percentage
21 – 30	4	3,15%
31 – 40	27	21,26%
41 – 50	35	27,56%
51 – 60	37	29,13%
61 – 70	18	14,17%
71 – 80	5	3,94%
81 – 90	1	0,79%
Maximum age	81 years	
Minimum age	28 years	
Average age	51 years	

Most of the respondents are between the ages 51 and 60 years (29.13%). The second-largest distribution of the respondents is between 41 and 50 years (27.56%). There are only four respondents who were between the ages 20 and 30. The average age of the farmers is 51 years, with the eldest individual being 81 and the youngest 28 years of age. This shows that most of the farmers in this data set are older and have gained farming knowledge through experience. Table 3.10 shows the distribution of farm experience.

Table 3.10: Distribution of respondents' experience in the industry

Experience (years in farming)	Number of respondents n = 127	Percentage
0 – 9	13	10%
10 - 19	39	31%
20 - 29	32	25%
30 - 39	27	21%
40 - 49	10	8%
50 - 59	5	4%
60 - 69	1	1%
Maximum farming experience	60 years	
Minimum farming experience	0 years	
Average farming experience	24 years	

Table 3.10 shows that most (31%) of the respondents have 10 to 19 years of experience in the farming industry. The average farmer has about 23 years of farming experience, ranging from 0 to 60 years of farming experience. Table 3.10 indicates that only 10% of the respondents have 10 years or fewer farming experience. High levels of experience in farming are demonstrated, as 90% of the respondents have more than 10 years' experience. Experience cannot be learned, adjusted or taught in a short period; this implies that younger, less experienced farmers are at a disadvantage (Henning & Jordaan, 2015). It is commonly found that farming experience may have a positive influence on the applicant's ability to be accepted by a financial organisation (Arene, 1993; Nwankwo, 2004; Olagunjiu & Adeyemo, 2007; Oladeebo & Oladeebo, 2008; Afolabi, 2010, Nwachukwu *et al.,* 2010; found in Ajah *et al.,* 2014). However, Nwankwo (2004) found that farm experience has a negative influence on the applicant's ability to be accepted by financial organisations. Based on the distribution of experience in Table 3.10, most of the applicants have more than 10 years' experience and have gathered more knowledge through experience.

3.3.7.2. Education

Farmers who are more highly educated tend to be more successful and to receive the same or better returns from farming than less educated farmers (Mishra, Hisham & Johnson, 1999). Level of education usually has a positive influence on repayment ability of applicants (Nwankwo, 2004; Olagunjiu & Adeyemo, 2007; Oladeebo & Oladeebo, 2008). Therefore, it is expected that education will also have an influence in the classification of repayment ability of the applicant farmers. Education information was collected by determining exactly what type of education each applicant possessed – this was obtained by a list of categories. The analysis indicated multi-collinearity existed between the variables, therefore the applicants were divided into two categories, that is, tertiary education or no tertiary education. Tabulated in Table 3.11 is the distribution of the respondents' education level. Table 3.11 shows that 87 of the 127 respondents received a tertiary education, while the other applicants have no tertiary education (they have matric or less). This shows that most of the farmers in this data set have a tertiary education, therefore they may be more successful or receive better returns from farming.

Table 3.11: Distribution of respondents' educational level

Educational level	Number of respondents n=127	Percentage
Tertiary	87	68.50%
No tertiary education	40	31.50%

3.3.7.3. *Ownership*

The ownership category is divided into two categories, namely, owner and manager. It is important for the financial organisation to know who owns the assets of the business, since it will determine who takes responsibility for the repayment of the loan. Access to loans and loan sizes are usually correlated with land ownership, particularly in underdeveloped, formal financial systems (Durguner *et al.,* 2006). An increase in land ownership leads to lower leverage and liquidity, a lower rate of return on assets and a greater portion of the borrower's economic rate of return occurring as unrealised capital gains on farm land (Durguner *et al.,* 2006). Petracco and Perder (2009) found that land tenure increases access to credit, because of the enhanced land security provided; land tenure is a method that farmers can use, by offering land as collateral. Therefore, it is found that ownership has a positive influence on being accepted by a financial organisation. Land tenure provides the financial organisation with an extra form of security, this indicates that the financial organisation's willingness to accept an applicant depends on the type of land ownership.

Table 3.12: Role of client in the business when applying for a loan

Ownership	Number of respondents n=127	Percentage
Owner	120	94%
Manager	7	6%

Table 3.12 shows that 94% of the applicants who had applied for loans are owners of the farming business, while 6% of the applicants are managers.

3.3.8. FINAL DECISION

The final decision refers to the dependent variable in the statistical model. For the purpose of the research the applications can be either rejected or accepted. Only 11% of the applicants were declined, while the other 89% were approved by the financial organisation; this is demonstrated in Table 3.13.

Table 3.13: Final decision in determining the repayment ability of the loan applicants

Final decision	**Total number of respondents n = 127**	**Percentage**
Approved	113	89%
Declined	14	11%

Next, the focus shifts to the description of the methods used to meet the aim of this research.

3.4. METHODS

The credit data was analysed by making use of LR, PA and NN models. These models were compared, based on their accuracy of correctly classifying agricultural credit applications, especially high-risk applications as determined by the financial organisation.

3.4.1. LOGISTIC REGRESSION

The LR model aims to find the relationships between the outcome-dependent variable (accept or reject) and a set of categorical and continuous attributes of the credit applicants. The LR equation is depicted in Equation 1.

$$\log\left[\frac{p_i}{1-p_i}\right] = \propto + \beta_1 x_i + \beta_2 x_2 + \dots + \beta_n x_n$$

Equation 1

Let p_i be the probability of the default of an agricultural borrower i, the intercept term is represented as $\propto$, and β_i represents the respective coefficient in the linear combination of independent variables x_i for $i = 1 - n$, which includes financial ratios and borrower characteristics. The dependent variable is the logarithm of the odds. The outcome of interest is determined by the logarithm of the ratio of the two probabilities, $\log\left[\frac{p_i}{1-p_i}\right]$, (Leea *et al.,* 2002). The probability of a value of one for the dichotomous outcome (approved or rejected) is calculated by a given set of independent variables, shown in Equation 2:

$$Z = \frac{1}{1 + e^{-Z}}$$ **Equation 2**

In Equation 2, Z refers to the probability that the dichotomous outcome where the agricultural applicant will be detonated: 1 if the applicant is approved or 0 if rejected, shown in Equation 3:

$$Z = \propto + \beta_1 x_i + \beta_2 x_2 + \ldots + \beta_n x_n$$ **Equation 3**

In credit scoring, the objective of an LR is to determine the conditional probability of a specific observation within a class, depending on the independent variables used for the credit applicant (Leea *et al.,* 2002). The LR was performed in STATA 11 by means of a binary LR.

3.4.2. PROBIT ANALYSIS

The probability P_i of a farmer being an acceptable risk can be expressed in terms of the cumulative distribution of a standard normal random variable (Quaye *et al.,* 2015) shown in Equation 4 and Equation 5:

$$P_i = prob[Y_i = 1\, X] = \int_{-\infty}^{\propto + x_{i\beta_1}} (2\pi)^{-1/2} \exp\left(-\frac{t^2}{2}\right) dt$$ **Equation 4**

Where:

$$P_i = \varphi(x_i \beta)$$ **Equation 5**

The Y_i represents the dependent variables, in this case to determine if the credit applicant is an accepted credit applicant (y=1) or rejected credit applicant (y=0). X represents the independent variables. The marginal effects are used to interpret the relationship between a specific variable and the probability of the outcome.

For continuous variables, the marginal effects while keeping the other variables constant is algebraically expressed in Equation 6:

$$\frac{\partial P_i}{\partial x_{ik}} = \emptyset(x_i\beta)\beta_k$$ **Equation 6**

The marginal effects for dummy variables (d) (which represent the discrete changes in the predicted probability) are demonstrated in Equation 7:

$$\Delta=\varphi(x\beta, d=1)-\varphi(x\beta, d=0)$$ **Equation 7**

The φ represents the cumulative distribution function while the $\emptyset$ represents the probability of density function. The probability of density function is defined as the probability that this random variable will take on a given value (Quaye *et al.,* 2015). This refers to the independent variables that have an influence on the function of the borrower characteristics that will have an influence on the borrowers' repayment ability. The formula can be expressed algebraically as in Equation 8:

$$\varphi^{-1}(P_{LD}) = \beta_i X_i = \beta_0 + \beta_1 B_i + \beta_2 L_i + \beta_1 Z_i + \varepsilon_i$$ **Equation 8**

The LD in the formula represents the loan delinquency, Bi contains specific borrower characteristics, Li contains loan-specific variables, Zi contains lender-specific variables, β_i represents the estimable parameters, while the ε represents the error, which is assumed to have a variance of 1 and is distributed as standard normal (Quaye *et al.,* 2015). The PA was performed in STATA 11 by means of a binary PA.

3.4.3. NEURAL NETWORKS

A multi-layer perceptron consists of an input and an output layer. The input layer consists of p number of independent variables (*x*) and the output layer consists of a single output neuron. The perception can be calculated by the function shown in Equation 9:

$$\mathrm{u_k} = w_{k0}x_0 + w_{k1}x_1 + w_{k2}x_2 + \ldots + w_{kp}x_p = \sum_{\mathrm{q=1}}^{\mathrm{p}} \mathrm{w_{kq}x_q}$$ **Equation 9**

The input layer consists of various variables, characteristics used in credit applications, $\mathrm{x_q}$ (*q* = 1,……,*p*) are known as a signal. The variables used in this research consist of farmer and loan characteristics, farm financial performance, and industry information. In the training of the

NN, weights or synaptic weights (*w*) are used, these are indicated by the subscripts (*k, p*), where *k* indicates the neuron and the specific weight, and *p* is indicated as the variable.

The weights that are assigned to the variables can be either negative (inhibitory) or positive (excitatory). A negative value decreases the value of negative $\mathrm{u}_{\mathrm{k}}^{1}$, while a positive value will increase the value (Thomas *et al.,* 2002). The value of x_0 is assigned a positive one (+1), meaning that the value of $w_{k0}x_0$, known as the bias or intercept of the specific layer, is w_{k0} and increases or decreases the $\mathrm{u}_{\mathrm{k}}^{1}$ by a constant value (Thomas *et al.,* 2002). An activation function is used to transform the output value ($\mathrm{u}_{\mathrm{k}}^{1}$). This activation function can be set by the operator according to the specific problem at hand. The output y_k of the neural is equal to the result of the neuron, as indicated in Equation 10:

$$y_k = \mathrm{F}(u_k)$$

Equation 10

Equation 11 mathematically represents a multi-layer perceptron that consists of more than one neuron, where F indicates the layer, and the subscript gives the exact number of the associated layer.

$$y_k = F_1(\sum_{\mathrm{q}=0}^{\mathrm{p}} \mathrm{w}_{\mathrm{kq}}\mathrm{x}_{\mathrm{q}})$$

Equation 11

The result (y_k) becomes an input for the second layer, which is presented below in Equation 12:

$$z_v = F_2(\sum_{\mathrm{k}=1}^{\mathrm{r}} \mathrm{k}_{\mathrm{vk}}\mathrm{y}_{\mathrm{k}} = F_2(F_1(\sum_{\mathrm{q}=0}^{\mathrm{p}} \mathrm{w}_{\mathrm{kq}}\mathrm{x}_{\mathrm{q}})))$$

Equation 12

The output of the neuron is illustrated by z_v, F_2 is the specific activation function in the output layer. K_{vk} is the weight used to connect neuron k and neuron v in the output layer, y_{k} (Thomas *et al.*, 2002). In this research an activation function that provides values between 0 and 1 are applicable. Therefore, a logistic activation function (above in Equation 13) is used when the output of the neuron needs to be mapped to the interval (0, 1) (Günther & Fritsch, 2010), as the case with classification of credit applications as either approved (1) or rejected (0).

$$F_{(u)} = \frac{1}{1 + e^{-u}}$$

Equation 13

Training in the network is performed through calculations of weight vectors, and back-propagation, which is one of the most frequently used methods (Thomas *et al.*, 2002; Mohammadi & Zangeneh, 2016). The training process starts with equal weights that are randomly selected (Günther & Fritsch, 2010), while a training pair is selected and the input variables ($\mathrm{x_q}$) are used to determine z_v. The difference between the z_v values and the known outputs (training outputs) (o_v) are calculated – this procedure is referred to as a forward pass (Thomas *et al.*, 2002). The forward pass and back propagation differ slightly. In back propagation the error is distributed back through the network in proportion to the contribution made by each weight, and adjusting the weights to reduce the portion of the error (Mohammadi & Zangeneh, 2016). This procedure continues for all the existing cases and only stops once a certain criterion is reached, normally the minimum error (Mohammadi & Zangeneh, 2016), which can be stipulated. For binary approaches, the cross entropy error function is shown below in Equation 14:

$$E = -\sum_{l=1}^{L}\sum_{h=1}^{H}(\mathrm{y}_{lh}\log(o_{lh}) + (1 - \mathrm{y}_{lh})(\log 1 - (o_{lh})) \qquad \textbf{Equation 14}$$

Cross entropy measures the difference between the determined output generated by the model and the observed output provided. *l = 1, …..,L,* indexes the observations of the input output combinations and *h* = 1, …..,*H* illustrates the output nodes (Günther & Fritsch, 2010).

The neuralnet package in R-Studio was used to train the neural network. The neuralnet package in R-Studio provides an opportunity to define the number of neurons and hidden layers (Günther & Fritsch, 2010). The number of hidden layers increases the complexity of the network, and therefore the network that provides the highest correctly classified results will determine the number of hidden layers. The software uses a function that consists of various arguments specified in the script, which includes the following (Günther & Fritsch, 2010):

Hidden layer – refers to a vector that specifies the number of hidden neurons and hidden layers. The number of layers used in the NN is determined by the network that holds the lowest misclassification error. Therefore, to determine the best-fitted network, the number of layers were varied.

Threshold – is an integer that stipulates the threshold for the derivatives of the error function. An error function is used as a criterion for the stopping of the NN, with the intention to minimise the error function. If no other number is stated, then the **default** is set to **0.01.**

Algorithm – refers to a string that contains the algorithm type that can be identified in the network. The algorithm referred to in this research is called the back-propagation network and is therefore specified as **"backprop"**.

Err.fct – the error function which is used as a stopping point for the NN can be specified between two functions. The **cross entropy ("ce"**) was used as error function, as the response of the data is binary.

Act.fct – the output values of the network were expected between 0 and 1, where 0 indicates a rejected applicant and 1 an accepted applicant. Therefore, the activation function, **"logistic"**, was chosen as acceptable.

Linear.output – as the output or determination of the applicant's repayment ability is illustrated as accepted or rejected, the output should be stated by the activation function, which maps the output between 0 and 1 and which illustrates if the applicants are accepted or rejected. The default setting is stated otherwise in the neuralnet package, and therefore the linear output was stated as **"FALSE"**.

The classification matrix is a popular method used to determine the estimated misclassification costs of the applicants. The method to calculate these results are demonstrated below.

3.4.4 CLASSIFICATION MATRIX

The misclassification errors are calculated by making use of the classification matrix in Table 3.14. The classification matrix presents the combinations of the number of actual and predicted observations in the dataset. The actual observations were gathered from the financial organisation (Good and Bad) and are read horizontally. The model testing predicted observations (good and bad) were generated by the statistical models and are read vertically. The Good (G) actual observations are applicants who have been accepted by the financial organisation. Bad (B) refers to the actual observations that have been rejected by the financial organisation. Predicted good (g) observations are applicants who are predicted as good by the statistical model. The statistically predicted bad observations (b) are observations that are predicted as bad by the statistical model. This is demonstrated in Table 3.14.

Table 3.14: Classification matrix used for classification purposes

		Model testing		
		good (g)	**bad (b)**	**Total**
Actual observations	**Good (G)**	Gg	Gb	**TG**
	Bad (B)	Bg	Bb	**TB**
	Total	**Tg**	**Tb**	**TN**

Source: Abdou *et al.* (2007)

From this matrix a number of useful rates can be calculated (Abdou, 2009):

The correctly classified good rate is calculated by Gg divided by TG shown in Equation 15. Gg shows the statistical model predicted a good outcome while the actual good observations were also good. TG refers to Total actual good observations.

Correctly classified good rate = (Gg/TG) **Equation 15**

The correctly classified bad rate is calculated by Bb divided by TB shown in Equation 16. Bb refers to B, the actual bad observations, and b is the statistical model predicted outcome as bad. TG symbolises total actual bad observations.

Correctly classified bad rate= (Bb/TB) **Equation 16**

Type I error rate, shown in Equation 17, is calculated by Gb divided by TG. Gb is the actual good observations while the statistical model predicted a bad outcome. TG as described above refers to total actual observations.

Type I error rate = (Gb/TG) **Equation 17**

Type II error rate, shown in Equation 18, is calculated by Bg divided by TB. Bg is the actual bad observations which the statistical model predicted as good. TB as described above refers to total actual bad observations.

Type II error rate=(Bg/TB) **Equation 18**

The focus of this research will be the identification of Type II errors to reduce misclassification. Further a discussion of the results generated from the NN, LR and PA will be demonstrated in Chapter 4.

CHAPTER 4

RESULTS AND DISCUSSION

4.1. INTRODUCTION

This chapter presents the results of the analyses as well as a discussion of the results. The chapter is divided into four sections. In the first three sections the results of the LR, PA and NN analyses are presented. The last section is devoted to the results of the classification comparison between the three statistical models.

4.2. LOGISTIC REGRESSION

A LR was performed in STATA 11 to generate the results. Table 4.1 illustrates the results from the LR analysis to assess the relationship between the selected characteristics of the respondents and the probability of being granted credit. It is important to note that the indication of the coefficients and significance in Table 4.1 do not indicate the variation in relative importance between the variables. The results provide an indication of the relationship between the variation observed in explanatory variables and the probability to be judged acceptable.

The McFadden R-squared and likelihood ratio statistics determine the goodness of fit for the LR. The higher the McFadden R-squared the better the fit of the model, therefore a McFadden R-squared value of 0.2 to 0.4 is considered highly satisfactory (Van der Merwe, 2011). The overall model is a good fit, as the McFadden R-squared value is 0.56. The likelihood ratio statistic is equivalent to the F test in the Linear Regression Model (Gujarati, 2003). According to the likelihood ratio statistic (-19.293945) this indicates that the overall model has a significant impact on predicting the probability of correctly classifying credit applicants.

The results shown in Table 4.1 indicate that 8 of the 22 variables are significant when considering a significance level of 10%. These variables include medium term, credit history, DTA, net farm ratio, Diversification 2 (diverse 2), high-risk, ownership and experience.

Table 4.1: Determinants in classification of credit applicants (standardised data)

Variable	Coefficient	Standard Error	Z statistic	Probability	
Intercept	***5.7800***	***3.7634***	***1.54***	***0.125***	
Characteristics of borrower					
Business loyalty	0.3471	0.7632	0.45	0.649	
Age	-1.2178	0.7605	-1.60	0.109	
Experience	2.3331	1.0207	2.29	0.022	**
Education	-0.9856	1.1879	-0.83	0.407	
Collateral					
Collateral	4.8875	4.4212	1.11	0.269	
Ownership	4.2535	2.2227	1.91	0.056	*
Capacity					
DTA	-2.1343	1.0058	-2.12	0.034	**
WCTGR	0.3875	1.1081	0.35	0.727	
ATO	2.4393	2.0431	1.19	0.232	
ROA	-0.8501	1.5570	-0.55	0.585	
NETFARMRATIO	-2.8594	1.0964	-2.61	0.009	***
PRODCOST	-0.0986	1.9553	-0.05	0.960	
CASHFLOW	0.6185	0.5221	1.18	0.236	
Condition					
High risk	4.8770	2.8052	1.74	0.082	*
Medium risk	1.7565	1.5418	1.14	0.255	
Diverse 2	-3.0771	1.7537	-1.75	0.079	*
Diverse 3	-2.9162	2.3222	-1.26	0.209	
Medium term	-5.1965	3.0180	-1.72	0.085	*
Long term	-5.1621	3.4200	-1.51	0.131	
Repayment period	-0.1597	1.0788	-0.15	0.882	
Loan amount	-0.1553	0.4025	-0.39	0.700	
Capital					
Credit history	-4.4392	1.7393	-2.55	0.011	**
Goodness of fit					
McFadden R-squared	0.5622				
Likelihood Ratio Statistic	19.2934				
Prob (Likelihood Ratio Statistic)	0.0007				

The ***, **, * indicate the significance of 1%, 5% and 10% respectively

The 5 Cs of credit are used by the credit analyst as a framework for an in-depth credit analysis, and refer to characteristics of borrower, collateral, capacity, condition and capital.

Borrower characteristics refer to personal characteristics of an individual – the borrower's risk attitude is an important element and a negative evaluation could lead to rejection of an application (Henning & Jordaan, 2015). This information defines who the borrower is and identifies the characteristics of the enterprise (Gitman *et al.,* 2014). This includes characteristics such as loyalty, age, education and experience. The significant variable among borrower characteristics is experience.

Collateral refers to the amount of assets that the business owns and which is used as security by the financial organisation when a loan is granted (Gitman *et al.,* 2014). The characteristics included in collateral are variables such as collateral and ownership (Henning & Jordaan, 2015). The study found collateral to be insignificant, although ownership was found to be significant. Collateral could be insignificant because the majority of credit applicants in the dataset have secured collateral.

Capacity refers to the applicant's ability to repay the requested credit to the financial organisation. This is determined by making use of financial ratios such as DTA, WCTGR, ATO, ROA, net farm income, production-cost and cash-flow. The capacity characteristics found to be significant include DTA and net farm ratio.

Condition is defined as characteristics that are industry-specific to economic conditions; this includes high-risk industry, medium-risk industry, diversification (diverse 2 and diverse 3), medium term, long term, loan amount and repayment period (Gitman *et al.,* 2014; Henning & Jordaan, 2015). The variables significant among condition characteristics included high risk, medium term and diverse 2.

Lastly, capital is the leverage ratio used to determine the applicant's debt relative to equity (Gitman *et al.,* 2014). Characteristics, such as credit history, are categorised as capital (Henning & Jordaan, 2015). Credit history was found to be significant in this study.

The results indicate that applicants who apply for medium-term loans are more likely to be rejected ($p<0.1$) than applicants who apply for short-term loans. Short-term loans are considered to be the base category. Long-term loans, in this case, are classified as insignificant to the financial organisation. Personal factors, such as education, experience, education and age, usually have a significant influence on the classification of applicants. Education, referred to as tertiary, is commonly found to have a positive influence on the acceptance or repayment ability of applicants (Nwankwo, 2004; Olagunjiu & Adeyemo, 2007; Oladeebo & Oladeebo, 2008; Quaye *et al.,* 2015). However, in this case, education is classified as insignificant at 10%. Other personal factors, such as ownership and experience, have an influence on classification in this specific case. Experience ($p<0.05$) is positively related to being classified as approved by the financial organisation. The positive influence of farming experience generally meets the a priori expectations of previous studies (Arene, 1993; Nwankwo, 2004; Afolabi, 2010). Applicants are more likely to be approved by the financial organisation when they have more experience. Interestingly, neither age nor education was found to be significantly related to repayment ability.

Ownership ($p<0.05$) is positively related to being classified as approved by the financial organisation. A positive influence by ownership meets the expectation of previous research, where an increase in land tenure increases the financial organisation's willingness to grant credit to the applicant (Durguner *et al.,* 2006; Petracco & Perder, 2009)

The DTA ratio compares a farm's debt obligations (owed by the farmer) to the value of the farm assets (FFSC, 2011). It is important for the applicant to have more assets than debt in the farm business. The higher the DTA ratio is, the larger the risk exposure will be in the farming business (FFSC, 2011). Should the applicant have a higher DTA ratio, the application is more likely to be rejected by the financial organisation ($p<0.05$). Applicants who have higher-valued assets are more likely to be granted credit, as higher-valued assets is a measurement of security for the financial organisation. Net farm income ratio has a negative influence ($p<0.001$) on the likelihood of being classified as approved. Should the net farm income decrease in the farm enterprise the likelihood of the application being rejected may increase (Quaye *et al.,* 2015). Therefore, an applicant with a higher net farm income is more likely to be accepted by a financial organisation.

Diversification refers to the number of enterprises that are operated on the farm. The number of farm enterprises was classified into three categories, diverse 1, 2 and 3 or more. Diverse 1 was used as the base variable; this variable was compared to diverse 2 and diverse 3 or more. The results indicate that applicants with two enterprises on their farms are more likely to be rejected ($p<0.1$) than applicants who have a single enterprise (diverse 1). This result is inconsistent with other findings. It is commonly found that having more than one enterprise in a farm business reduces the chance of loss from a given hazard (Miller *et al.,* 2004). Culas and Mahendrarajah (2005) state that diversification's main aim is to reduce the risk of overall return by selecting a mixture of activities that have net returns with a low or negative correlation. The results may be an indication that the particular bank prefers clients who specialise in a single enterprise. However, Chirwa (1997) found that the degree of diversification has a significant influence on agricultural credit repayment ability, because more farm enterprises reduce risk and increase security. In this specific case, diverse 3 is insignificant and does not have an influence on being approved as an acceptable applicant.

An applicant who is interested in investing in a high-risk industry is more likely to be accepted by the financial organisation ($p<0.1$) than an applicant who is interested in investing in a low-risk industry, as the applicants may have smaller DTA ratios, or have secured credit history. The majority of the applicants used in this case study could have a secured credit history, which could have an influence on the outcome.

The relationship between the applicant and the financial organisation may also have an influence on the classification of the applicant. This is indicated by the negative influence of credit history of an applicant ($p<0.1$). Therefore, applicants with a poor credit history are more likely to be rejected by the financial organisation. Previous research indicates that clients with poor credit history are associated with loan delinquencies (Addae-Korankye, 2014). Applicants with a poor credit history are more likely to be rejected by financial organisations to prevent loan delinquencies.

4.3. PROBIT ANALYSIS

A PA was used as the second specification to assess agricultural credit applicants. Table 4.2 illustrates the results of the PA's classification of the credit applications. As for the PA, it is important to note that the sizes of the coefficients and significance in Table 4.2 do not show the difference in relative importance of the variables. The results provide an indication of the relationship between observed explanatory variables, and not of the probability to be judged acceptable.

The McFadden R-squared and likelihood ratio statistics determine the goodness of fit for the PA. A McFadden R-squared value of 0.2 to 0.4 is considered to be highly satisfactory, however, the higher the value the better the model fit (Van der Merwe, 2011). The PA had a McFadden R-squared value of 0.55, which shows that the overall model is a good fit. The likelihood ratio statistic is equivalent to the F test in the linear regression model (Gujarati, 2003). According to the likelihood ratio statistic (-19.8901), this indicates that the overall model has a significant impact on predicting the probability of correctly classifying credit applicants.

Table 4.2 indicates that 6 of the 22 variables are significant at 10%. These variables are credit history, DTA, net farm ratio, diverse 2, ownership and experience.

Table 4.2: Determinants in classification of credit applicants by means of a PA (standardised data)

Variable	Coefficient	Standard Error	Z statistic	Probability	
Intercept	***2.8654***	***1.9895***	***1.44***	***0.150***	
Characteristics of borrower					
Business loyalty	0.1964	0.4069	0.48	0.629	
Age	-0.5842	0.3958	-1.48	0.140	
Experience	1.0735	0.4819	2.23	0.026	**
Education	-0.4188	0.6242	-0.67	0.502	
Collateral					
Collateral	2.2572	2.1633	1.04	0.297	
Ownership	2.0289	1.0979	1.85	0.065	*
Capacity					
DTA	-0.9382	0.4670	-2.01	0.045	**
WCTGR	0.3306	0.6324	0.52	0.601	
ATO	1.1138	0.7678	1.45	0.147	
ROA	-0.4055	0.7396	-0.55	0.583	
NETFARMRATIO	-1.4200	0.5305	-2.68	0.007	***
PRODCOST	-0.0301	1.0190	-0.03	0.976	
CASHFLOW	0.2544	0.2899	0.88	0.380	
Condition					
High risk	2.2847	1.5296	1.49	0.135	
Medium risk	0.7980	0.8288	0.96	0.336	
Diverse 2	-1.6578	0.9237	-1.79	0.073	*
Diverse 3	-1.3895	1.1575	-1.20	0.230	
Medium term	-2.3058	1.4844	-1.55	0.120	
Long term	-2.2952	1.7475	-1.31	0.189	
Loan amount	-0.0746	0.2236	-0.33	0.739	
Repayment period	-0.2255	0.5399	-0.42	0.676	
Capital					
Credit history	-2.1748	0.8706	-2.50	0.012	**
Goodness of fit					
McFadden R-squared	0.5487				
Likelihood ratio statistic	-19.8901				
Prob (Likelihood ratio statistic)	0.0010				

The ***, **, * indicate the significance of 1%, 5% and 10% respectively

The characteristics used in this statistical model are divided according to the five categories known as the 5 Cs of credit, in this case, characteristics of the borrower, such as loyalty, age, experience and education. Collateral characteristics include collateral and ownership. Condition characteristics include high risk, medium risk, medium term, long term, diversification (diverse 2 and diverse 3), repayment period and loan amount. Capacity characteristics include DTA, WCTGR, ATO, net farm ratio, production costs and cash flow. Lastly, capital includes the characteristic, credit history. Of these 22 variables only 6 variables indicated significance at 10%, these are ownership, experience, diverse 2, DTA, net farm ratio and credit history.

Personal qualities, such as experience ($p<0.01$), have a positive relationship with the likelihood of an application being approved by the financial organisation. Results suggest that loan applicants with more years of experience are more likely to be granted loans. Usually, farming experience has a positive influence on applicants' applications being approved by a financial organisation (Arene, 1993; Ezeh, 2003; Nwankwo, 2004; Afolabi, 2010). Ownership ($p<0.1$) has a positive relationship with the likelihood of being approved by the financial organisation. The positive influence of ownership on likelihood of having a loan application approved meets the expectation of previous research, which found that an increase in land tenure increases the financial organisation's willingness to grant credit to the applicant (Durguner *et al.,* 2006; Petracco & Perder, 2009).

The number of enterprises associated with the farm was found to have an influence on whether the applicant is classified as approved. Diverse 1 was used as a base category and was compared to diverse 2 and 3 or more. The results indicate that applicants who have two enterprises on the farm are more likely to be rejected ($p<0.1$), compared to diverse 1. Again, this is inconsistent with other findings, because having more than one enterprise in a farm business is expected to reduce the chance of a loss from a given hazard (Miller *et al.,* 2004). Chirwa (1997) found that the degree of diversification has a significant influence on agricultural credit repayment ability. Similar to the LR, diverse 3 was found to be insignificant.

The financial organisation is interested in knowing whether the applicant has sufficient or high-value assets compared to debt when the applicant applies for credit. The DTA ratio compares the farm's debt obligations to the value of the farm assets (FFSC, 2011). It is important for the applicant to have a lower DTA ratio in a farm business. The higher the DTA ratio is, the larger the risk exposure will be in the farming business (FFSC, 2011). Therefore, if the applicant has a high DTA ratio, it is more likely that the application will be rejected by the financial organisation ($p<0.05$).

Net farm income ratio has a negative relationship ($p<0.001$) with the likelihood of being classified as approved. This finding is similar to the findings of Quaye *et al.* (2015), who found that, as net farm income of the farming business decreases the likelihood of the applicant being rejected may increase. Other research found that net farm income has an influence on the creditworthiness of the applicant (Onyenucheneya & Ukoha, 2007). Results from this research show that applicants who have a lower return on net farm income are less likely to be accepted ($p<0.001$). Therefore, it is important for credit applicants to have a higher return on net farm income when they apply for loans from financial organisations.

Financial organisations determine the credit history of the applicants from previous credit history records. The relationship between the financial organisation and the applicant also have an influence on the classification of the applicant. The results indicate that a poor credit history has a negative influence on an applicant when applying for credit ($p<0.05$). Therefore, applicants who have a poor credit history are more likely to be rejected by a financial organisation. Previous research indicates that clients with a poor credit history are associated with loan delinquencies (Addae-Korankye, 2014).

4.4. NEURAL NETWORKS

A back-propagation NN was used to train and determine the classification accuracy of the agricultural credit classification model in R-Studio. Basic information used to generate the result in the trained NN reached a minimum error of 3.45. A total of 44 553 steps was needed for the absolute partial derivatives of the error function to reach a minimum threshold of less than 0.01. Figure 4.1 shows the plot of the trained NN, which demonstrates the single hidden layer that consists of two hidden neurons, synaptic weights and intercept values.

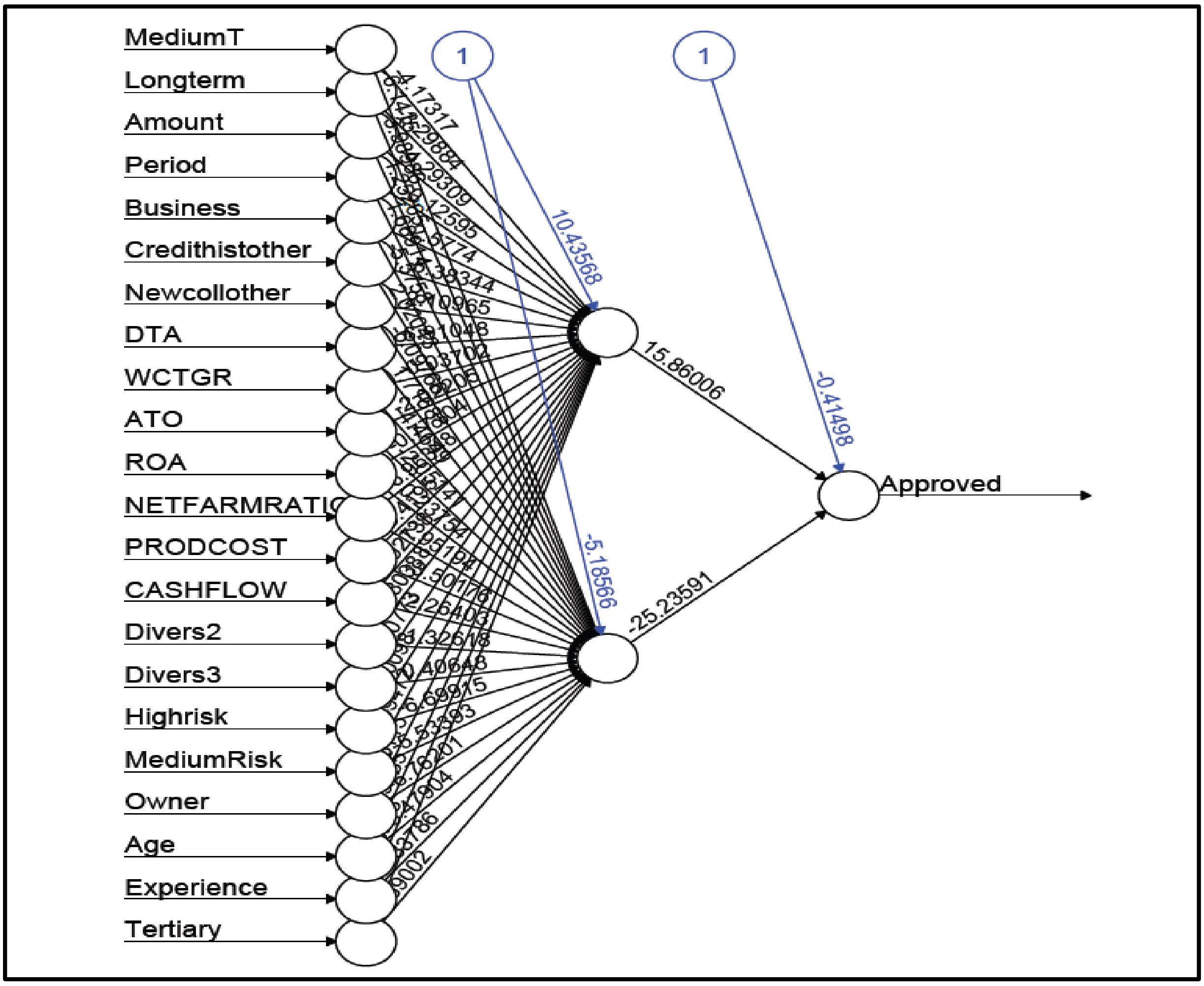

Figure 4.1: Plot of trained neural network including trained synaptic weights and basic information about training process

The synaptic weights found between each input and hidden layer in Figure 4.1 are unclear and are shown in Table 4.3. Weights were used to minimise the error function by either increasing or decreasing their values, the product of the weights and the associated inputs. Outputs from the sum product of the weights and inputs were then transformed by the logistic activation function into output values of 0 (rejected) and 1 (accepted) (Thomas *et al.,* 2000). The weights indicate the effect and direction of the inputs in the NN, specifically in the two neurons. Variable weights, in respect of each neuron that was generated in the NN, are tabulated in Table 4.3.

Table 4.3: Weights generated from neural network

	Neuron 1	Neuron 2
Medium term	10.436	-5.186
Long term	-4.173	6.141
Amount	-3.299	3.99
Period	-4.293	1.253
Business (Loyalty)	-0.126	1.689
Credit history	-2.577	-5.376
Collateral	-5.383	2.02
DTA	2.11	-6.098
WCTGR	6.91	1.717
ATO	-0.037	-4.465
ROA	-2.022	-2.051
NETFARMRATIO	-0.771	0.238
PRODCOST	-6.963	2.952
CASHFLOW	-5.419	-3.502
Divers2	2.027	2.264
Divers3	-5.804	1.326
High risk	1.008	0.406
Medium risk	1.201	-6.699
Owner	-1.495	-5.534
Age	6.181	-6.762
Experience	-3.494	-3.479
Tertiary	5.092	1.838

Outputs generated from each neuron is not transferred to the next neuron, but the output in the hidden layer is transformed into a value with the aid of a logistic activation function. The value is determined by the weight coefficients, inputs and intercept of each neuron, the output (Neuron 1) from the hidden layer is then transformed again by the activation function and transferred to Neuron 2. If this was to be the final outcome, the output would be the result or generated outcome. Intercepts, coefficient weights and inputs have an influence on the output generated, which is influenced by the activation function. In Table 4.3, the positive weight values are transformed into values closer to 0 and negative weight values closer to 1. This transformation is activated by a logistic activation function. The basic information in the neural network, such as the hidden layers and intercept, is demonstrated in Table 4.4.

Table 4.4: Basic information in neural network

Intercept to Approved	-0.415
Hiddenlayer1 to Approved	15.86
Hiddenlayer2 to Approved	-25.236

A disadvantage of NN is that the results are like black boxes, which are difficult to interpret and fail to provide reasons for the results attained (Salame, 2011). The results of the NN cannot be interpreted in terms of the likely influence of the different variables on the application being classified in the approved category. For this reason, the influence of the variables are not discussed, the following section discusses the classification ability of the NN. The inability to interpret the results of the NN is a huge disadvantage of the method, as it can be expected that the financial organisation provides reasons for reaching the specific classification decision.

The trained NN classified the applications into two categories: either approved (1) or rejected (0). The actual observed classification and model-testing classifications of the NN are discussed, after which the classification accuracy of the NN, PA and LR are compared.

4.5. MISCLASSIFICATION COMPARISION

The accuracy by which credit providers classify their credit applications is very important. Organisations not only allocate time and resources to the process, but there are also costs involved when applications are classified incorrectly. Table 4.5 demonstrates a classification table with the actual observation and model-testing observed classifications of applicants who were justifiably approved, rejected and misclassified. The information provided for the misclassified applicants can be determined in terms of Type I and Type II errors. A cut-off value of 0.5 was used to determine the classification group of the applicants for all three statistical models. The cut-off value of 0.5 is considered to be the standard or predefined cut-off for LR, PA and NN (Limsombunchai *et al.,* 2005; Abdou *et al.,* 2007). The results are presented according to the four categories: Good/good (Gg), Good/bad (Gb), Bad/good (Bg) and lastly Bad/bad (Bb).

Table 4.5: Logistic regression classification table for agricultural credit applications

		Model testing		
		good (g)	**bad (b)**	**Total**
Actual observations	**Good (G)**	113	0	**113**
	Bad (B)	5	9	**14**
	Total	**118**	**9**	**127**

Table 4.5 shows that 113 of 127 applicants were classified correctly, 9 applicants were rejected correctly and 0 applicants were rejected who should have been approved. Five of the applicants were misclassified as good but they were actually rejected by the financial organisation. These 5 applicants have a common denominator: all 5 misclassified and rejected applicants have a vulnerable WCTGR of less than 10%, while the ideal rating is more than 30% (CFFM, 2014). Four of the 5 applicants have a vulnerable ATO rating of less than 30%, when the ideal rating is more than 45% (CFFM, 2014). Therefore, it can be established that, due to these two factors, the model regarded these applicants as high risk. The LR procedure classified with an overall accuracy of 96.06% and a total misclassification rate of 3.94%. When only the Type II errors are considered, the LR procedure has a 3.94% misclassification rate. Literature states that higher costs are associated with Type II errors. Ensuring that applications are classified correctly can make a significant contribution to credit providers. The results in Table 4.6 are compared to the results achieved by the PA procedure in Table 4.6, to determine which method provides the highest accuracy in classifying Type II errors.

Table 4.6: Probit analysis classification table for agricultural credit applications

		Model testing		
		good (g)	**bad (b)**	**Total**
Actual observations	**Good (G)**	113	0	**113**
	Bad (B)	5	9	**14**
	Total	**118**	**9**	**127**

There was no difference in classification scores for the LR and PA. The PA achieved a total misclassification rate of 3.94%, while the model achieved an overall accuracy rate of 96.06%. As demonstrated in Table 4.6, 113 applicants were classified correctly, 9 applicants were rejected correctly. The model has a 0% misclassification rate for Type I errors. No applicants were rejected who should have been approved by the financial organisation in the classification. When only the Type II errors are considered, the LR procedure has a 3.94% misclassification rate. Five of the applicants were misclassified as good when they were actually rejected by the financial organisation. The 5 applicants who were misclassified in the PA were the same applicants misclassified in the LR. All 5 the applicants have a vulnerable WCTGR of less than 10%, 4 of the 5 applicants have a vulnerable ATO rating of less than 30%. A higher ATO rating suggests that assets are used more effectively to generate revenues (FFSC, 2011). Therefore, it can be established that these two factors caused the model to refer to these applicants as posing high risk – these two variables are considered important for classifying applicants. These results are further compared to the results of the NN procedure to determine the highest accuracy in classifying Type II errors (high risk loans in the rejected category) (see Table 4.7).

Table 4.7: Neural network classification table for agricultural credit applications

		Model testing		
		good (g)	**bad (b)**	**Total**
Actual observations	**Good (G)**	111	0	**111**
	Bad (B)	2	14	**16**
	Total	**113**	**14**	**127**

The NN has a total misclassification rate of 1.54%. The results represent an improvement compared to the misclassification of the regression results. As shown in Table 4.7, 111 of 127 applicants were correctly classified in the Good/good category, 14 applicants were correctly rejected (Bad/bad). The model has a 0% misclassification rate for Type I errors, this refers to applicants who were rejected but who should have been approved according to the financial organisation's classification (Good/bad). Lastly and importantly, only 2 applicants were misclassified (1.54%) as being approved, but were actually rejected (Bad/good). These two applicants who were misclassified were not the same applicants rejected by the LR and PA. The model has identified two applicants as being high risk, because these 2 applicants have a vulnerable WCTGR ratio of less than 10% and a vulnerable ATO rating of less than 30% – the one applicant's is 6% and the other applicant's 17%. It may be due to these two variables that the model classified these applicants as high risk. The NN therefore has an overall accuracy rate of 98.43%.

4.6. CONCLUSION

Classification results from the NN cannot be interpreted in terms of variation in the variables and the influence thereof on the predicted outcome. The reasons for a specific decision are thus not clear, and cannot be explained to the applicant. Therefore, the "black boxes" can have an influence on the adaptation of the method for credit classification. The advantages of the LR and PA is that the variation in variables can be explained and the significance of the various variables can be explained. The variables of the LR and PA were investigated further to determine the significance of the various variables. The results of the LR indicated that 8 of the 22 variables, namely, medium term, credit history, DTA, net farm ratio, diverse 2, high risk, ownership and experience, are significant at 10%. The PA's results indicate that 6 of the 22 variables, namely, credit history, DTA, net farm ratio, diverse 2, ownership and experience, are significant at 10%. The two variables that were not significant in the PA were medium term and high risk. The selection and identification of variables are important, as they can be used as conformation for credit research. Financial organisations can use these characteristics to guide or recommend applicants when they apply for loans. Therefore, financial organisations can use these variables as recommendations to determine which variables are of importance to the financial organisation.

The NN received an overall accuracy of 98.43% compared to 96.06% for the LR and PA, which means the NN outperformed the LR and PA. The NN also achieved the highest accuracy rate when considering the high-risk applicants (Type II). This model achieved a misclassification rate of 1.54% compared to the LR and PA, which achieved a misclassification rate of 3.94%. In terms of the main objective of the research, NN does prove to have the best classification accuracy for the specific case. Neural networks were able to correctly predict the lowest number of Type II errors, which, according to literature, has the highest costs associated with it. These results were compared to that of literature that produced the same outcome, namely, that NN outperforms the LR and PA (e.g. Blanco, Mejias, Lara and Rayo, 2013; Mohammadi and Zangeneh, 2016).

CHAPTER 5

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1. SUMMARY AND CONCLUSIONS

Financial organisations have increased lending to the South African agricultural sector to finance production activities, capital expenditures and capital investments, such as machinery, vehicles, livestock, implements and land (DAFF, 2015). Both commercial and smallholder farmers require credit to support their operations, so that available natural resources can be used to its maximum potential. Smallholder farmers are reliant on credit but often struggle to access finance from financial organisations due to their inability to provide collateral (Chisasa & Makina, 2012; Chisasa, 2014). The increase in the demand for credit has made financial organisations more aware of the need to explore alternative evaluation procedures to reduce debt (Salame, 2011). A specific evaluation process must be followed to determine the creditworthiness of a farmer before credit can be granted to applicants. This process consists of the collection, analysis and evaluation of information, such as farmers' credit repayment history, income and overall finance, before credit can be granted (USAID, 2005). During the evaluation process, mistakes are made in the classification of applicants. The costs incurred by the financial organisation are known as misclassification errors, which are caused by accepting high-risk loans and/or rejecting low-risk loans.

The repayment ability of an applicant can be evaluated by making use of either a subjective or an objective approach. The subjective approach relies on human judgement. The objective approach, on the other hand, has the ability to reduce human judgement, inconsistency, and misclassification errors, and provides a tool for the credit analyst to accept or reject the applicant. The challenge is to select a model that will incorporate all the necessary information used to evaluate an agricultural applicant, and one that will suit the agricultural industry. The agricultural industry is unique, as various environmental factors (e.g. natural disasters and unforeseen weather patterns) that affect the farmers' loan repayment ability should be taken into consideration.

Researchers have compared and explored various methods to improve accuracy of the procedure for evaluation of credit applicants. These statistical models have been applied to international financial organisations in various locations and have proven to be more efficient and effective than the subjective approach. Even a fraction of a percent increase in credit scoring can be regarded as a significant improvement, give number of credit applicants assessed annually (West, 2000). In South Africa literature has explored factors used by a financial organisation to evaluate agricultural applicants (Henning & Jordaan, 2015, 2016). The accuracy of the various statistical methods was not compared (Henning, 2016) and it was recommended that a statistical model be found that has the ability to predict high-risk applicants correctly. The objective of this study was to explore the performance of different statistical models to identify a model that can minimise the misclassification of high-risk applicants, and identify the characteristics that influence repayment ability.

With the assistance of the financial organisation a total of 127 credit applications were obtained from various provinces in South Africa. The credit data used for the study was collected with a formal agreement from a specific financial organisation that is involved in the South African agricultural sector. This formal agreement ensured that all the data obtained from the financial organisation remains confidential. Various variables were provided by the financial organisation, which includes purpose of the loan, amount, period of repayment, business loyalty, credit history, collateral, financial information, farm diversification (number of enterprises on farm), associated industry risk, ownership of business, age of applicant, years of farming experience and education. The final outcome was considered as the dependent variable, which is either approved or rejected. LR, PA and NN were chosen to classify high-risk applicants, as these procedures demonstrated the highest degree of accuracy and are mostly used to classify agricultural applicants in credit research. STATA 11 was used to generate the results for the LR and PA, while R-Studio was used to generate the results for the NN.

The results indicate that NN achieved an overall accuracy of 98.43%, while LR and PA achieved an overall accuracy rate of 96.06%. Further investigation into the accuracy for classification of Type II errors between these three methods shows that NN had a misclassification rate of 1.54%, which is lower than the 3.94% found for the LR and PA. The outcome of the comparison shows that NN outperformed LR and PA, which confirms previous research (e.g. Blanco *et al.,* 2013; Mohammadi & Zangeneh, 2016). Results show that, for the LR and PA, variation in variables and the significance thereof can be explained. However, the variation and significance of variables in a NN cannot be interpreted. The specific decision for granting an application is unclear and the decision to reject the application cannot be explained.

In the field of credit the 5 Cs are well known as representing capital, capacity, collateral, borrower characteristics and condition; these were used as a framework by the analyst to assess the credit applicant. The variables of the LR and PA were investigated further to determine which characteristics are significant for predicting the probabilities of the applicant being correctly accepted. The variables that were significant at 10% for both the LR and PA are medium term, credit history, DTA, net farm ratio, diverse 2, high risk, ownership and experience. The difference between the LR and PA is that the LR has two extra variables that were significant, namely, medium term and high risk.

The LR and PA had similar results, however, there was a slight difference between the significance in variables. Research found that clients with poor credit history are associated with loan delinquencies (Addae-Korankye, 2014). The results generated from this case indicate that applicants with poor credit history are more likely to be rejected by a financial organisation. The main purpose of diversification in an agricultural enterprise is to reduce the risk of overall return. The results achieved by the two models are contradictory to the norm, as an applicant with more than one enterprise (diverse 2) is more likely to be rejected by the financial organisation than an enterprise with one enterprise (diverse 1). Most applicants who have more than two enterprises are more likely to be accepted by a financial organisation than an applicant with one enterprise. The results show that the loan of an applicant with a larger DTA ratio is more likely to be rejected by the financial organisation. Quaye *et al.* (2015) found that, should net farm income decrease in the farm enterprise, the likelihood of the application being rejected will increase; this is similar to the results found in this study. The results generated indicate that applicants who have more years of experience are more likely to be accepted by a financial organisation, once again confirming previous research (Arene, 1993; Nwankwo, 2004; Afolabi, 2010). Another personal quality, ownership, had a positive influence on the likelihood of being accepted by a financial organisation. Education and age were found to be insignificant, which is unusual, as other research found at least one of these variables to be significant.

The results generated by the LR, PA and NN show that the NN is more accurate than the LR and PA. Therefore, NN should be used by financial organisations to predict Type II errors, however, due to “black-boxes”, the variables that are important cannot be interpreted. The PA and LR have a lower overall accuracy than the NN, but can be used to indicate which characteristics are of importance for repayment ability. Financial organisations can use the important characteristics that can be identified to provide recommendations and guidance to evaluate applications for loans. Credit applicants can also use these identified important characteristics as a point of reference before applying for the loan at the financial organisation. These characteristics can be used as a form of reference for further credit research. Thus, it is important for a financial organisation to select a statistical model that meets what they would like to achieve best.

5.2. RECOMMENDATIONS

The following conclusions are drawn from the study:

- Implications of the research include that the proposed method can assist in reducing the waiting time needed for decision-making, which is one of the advantages of the objective approaches.
- Statistical models demonstrate accuracy and should be used as a tool during the decision-making procedure for accepting or rejecting agricultural loan applications.
- Statistical procedures ensure that decisions are consistent with past and present classification, which is an important element to ensure that there is no discrimination of any kind when credit is granted.
- It is important to ensure that the correct variables are selected when determining the repayment ability of the agricultural applicant. The characteristics that were found to be important for both LR and PA are the "medium term", "credit history", "DTA", "net farm ratio", "diverse 2", "high risk', "ownership" and "experience." The difference between the LR and PA is that the LR has two additional variables: "medium term" and "high risk". When the correct variables are selected financial organisations can use the important characteristics that are identified to provide recommendations and guidance to evaluate applications for loans. The credit applicants can also use these identified important characteristics as a point of reference before applying for the loan at the financial organisation.

From the study the following recommendations can be made for further research:

- The results generated by the model should be confirmed by further testing and application, by tracking and monitoring the repayments of the applicants over time. The testing of the model should include testing classification results against the actual performance of the applicants over time; this would determine if the model is actually classifying the applicants correctly. A factor that makes this difficult is that only Type II errors can be tracked, because if an applicant is rejected there is no method that can determine whether the classification was correct. Type II errors can be tracked, because the approved applicant is a client of the organisation and is identifiable.

– Previous research identified that interest rates have an influence on the repayment ability of agricultural loan applicants. Therefore, further research should be done into methods that interest rates can be incorporated into statistical credit models to determine which applications are accepted or rejected by financial organisations. It is also important to investigate the effects of interest rates on the financial output of the statistical models. The statistical model should also take into consideration fluctuations in interest rates, which can be incorporated into a statistical model.

– Unforeseen weather conditions, such as drought and natural disasters, have an influence on applicants' ability to repay loans. Therefore, it is important to further investigate the use of weather derivatives as a financial instrument that can be used by financial organisations or individuals as part of a risk-management strategy. It is important to investigate the use of this instrument further and to incorporate it into the statistical model to reduce the associated risk.

– Previous research has found that smallholder farmers are reliant on credit, but often struggle to access credit from financial organisations. Various statistical models have been identified and tested for commercial farming, therefore, the identified model can be applied to smallholder farmers. However, there are various variables, such as collateral, which some smallholder farmers cannot provide, therefore, further investigation should be performed to select the correct variables to be used when supplying credit to smallholder farmers.

– Further research is needed into statistical models that can be applied to non-profit organisations, such as government organisations. These statistical models can be compared to determine which models suit an organisation best, to ensure that the correct applicants are granted loans or subsidies. This will ensure that the capital invested in smallholder farming or a potential commercial farmer is the correct investment, especially by government organisations.

REFERENCES

Abdesamed, K.H. & Wahab, K.A. 2014. Financing of small and medium enterprises (SME): Determinants of bank loan application. *African Journal of Business Management* 8(17): 717-727.

Abdou, H.A. 2009. Genetic programming for credit scoring: The case of Egyptian public sector banks. *Expert Systems with Applications* 36(2009): 11402-11402.

Abdou, H.A. & Pointon, J. 2011. Credit scoring, statistical techniques and evaluation criteria: A review of the literature. *Intelligent Systems in Accounting, Finance and Management* 18(2-3): 59-88.

Abdou, H. A., Pointon, J. & El-Masry, A. 2007. On the applicability of credit scoring models in Egyptian banks. *Banks and Banking Systems* 2(1): 4-20*.*

Abdou, H.A., Pointon, J. & El-Masry, A. 2008. Neural nets versus conventional techniques in credit scoring in Egyptian banking. *Expert Systems with Applications* 35(2008): 1275-1292.

ABSA. 2015. *ABSA Outlook 2015*. ABSA. [Online]. Retrieved from: http://www.agrisa.co.za/pdf/Absa_Eng.pdf [26 October 2016].

Addae-Korankye, A. 2014. Causes and controls of loan default/delinquency in microfinance institutions in Ghana. *American International Journal of Contemporary Research* 4(12): 37-45.

Adekanye, F. 1986. *Elements of banking in Nigeria.* Lagos: F and A Publishers.

Ademu, W.A. 2006. *The informal sector and employment generation in Nigeri*a. Proceedings of the 2006 Annual Conference of the Nigeria Economic Society, Calabar. 22-24 August 2006.

Afolabi, J.A. 2010. Analysis of loan repayment among small scale farmers in Oyo State. *Journal of Social Sciences* 22(2): 115-119.

Ajah, E.A., Eyo, E.O. & Ofem, U.I. 2014. Analysis of creditworthiness and loan repayment among bank of agricultural loan beneficiaries (poultry farmers) in Cross River State, Nigeria. *International Journal of Livestock Production* 5(9): 155-164.

Alaraj, M., Abbod, M. & Hunaiti Z. 2014. *Evaluating consumer loans using neural networks ensembles*. In Proceedings of the 2014 International Conference on Machine Learning, Electrical and Mechanical Engineering. United Arab Emirates, Dubai. 8-9 January 2014.

Alaraj, M., Abbod, M. & Al-Hnaity, B. 2015. *Evaluation of consumer credit in Jordanian banks: A credit scoring approach*. 17th UKSIM-AMSS International Conference on Modelling and Simulation, Cambridge, United Kingdom. 25-27 March 2015.

Anang, T.A., Sipiläinen, T., Bäckman, S. & Kola, J. 2015. Factors influencing smallholder farmers' access to agricultural microcredit northern Ghana. *African Journal of Agricultural Research* 10(24): 2460-2469.

Anigbogu, T.U., Onugu, C.U., Onyeugbo, B.N. & Okoli, M.I. 2014. Determinants of loan repayment among co-operative farmers in ANKA North LGA of Anambra State, Nigeria. *European Scientific Journal* 10(22): 168-190.

Arene, G.J. 1993. An analysis of loan repayment potentials of smallholder soyabean group farmers in Nigeria. *Zeitschrift für ausländische Landwirtschaft,* 32(2): 160-69.

Awunyo-Vitor, D. 2012. Determinants of loan repayment default among farmers in Ghana. *Journal of Development and Agricultural Economics* 4(13): 339-345.

Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J. & Vanthienen, J. 2003. Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the Operational Research Society* 54: 627-635.

Banasik, J., Crook, J. & Thomas, L. 2003. Sample selection bias in credit scoring models. *Journal of the Operational Research Society* 54(8): 822-832.

Bandyopadhyay, A. 2007. *Credit risk models for managing bank's agricultural loan portfolio.* Munich Personal RePEc Archive Paper No. 5358 [Online]. Retrieved from: http://mpra.ub.uni-muenchen.de/5358/ [27 June 2015].

BiiiCPA 2015. *The Basel iii Accord.* Basel Committee on Banking Supervision. [Online]. Retrieved from: http://www.basel-iii-accord.com/ [18 October 2015].

Blanco, A., Mejias, R., Lara, J. & Rayo, S. 2013. Credit scoring models for microfinance industry using neural networks: Evidence from Peru. Expert System Application 40(1): 356-364.

Blocker, A., Ibendahl, G. & Anderson, J. 2010. *Interpreting farm financial ratios.* Mississippi State University. [Online]. Retrieved from: http://www.cottoninc.com/fiber/AgriculturalDisciplines/AgriculturalEconomics/Farm-Finance/Interpreting-Farm-Financial-Rates/Interpreting_Farm_Financial_Ratios.pdf [22 September 2016].

Calomiris, C.W., Hubbard, R.G. & Stock, J.H. 1986. The farm debt crisis and public policy. *Brookings Papers on Economic Activity* 2. [Online]. Retrieved from: http://scholar.harvard.edu/files/stock/files/the_farm_debt_crisis_and_public_policy.pdf [30 October 2016].

CFFM. 2014. *Farm finance scorecard*. Center for Farm Financial Management. [Online]. Retrieved from: http://www.cffm.umn.edu/Publications/pubs/FarmMgtTopics/FarmFinanceScorecard.pdf [10 October 2016].

Chen, M-C & Huang, S-H. 2003. Credit scoring and rejected instances reassigning through evolutionary computation techniques. *Expert Systems with Applications* 24(4): 433-441.

Chirwa, E.A. 1997. An econometric analysis of the determinants of agricultural credit payment in Malawi. *African Review of Money Finance and Banking* 1(2): 107-122.

Chisasa, J. 2014. The finance-growth nexus in South Africa's agricultural sector: A structural equation modelling approach. *Banks and Bank Systems* 9(4): 38-47.

Chisasa, J. & Makina. D. 2012. Trends in credit to smallholder farmers in South Africa. *International Business and Economic Research Journal* 11(7): 771-784.

Craven, R., Nordquist, D. & Klair, K. 2011. *The benefits of financial benchmarking to farmers in the United States.* Congress Proceedings – Non peer reviewed papers and posters, 268-275. Christchurch, New Zealand.

Culas, R. & Mahendrarajah, M. 2005. *Causes of diversification in agriculture over time: Evidence from Norwegian farming sector*. In Proceedings: The future of Rural Europe in Global AgriFood Systems. 11th Congress of the European Association of Agricultural Economists, Copenhagen, Denmark. 24–27 August 2005.

Culp, T. 2013. *Agricultural credit analysis.* Senior Project [Online]. Retrieved from: http://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=1119&context=agbsp [14 October 2015].

DAFF (Department of Agriculture, Forestry & Fisheries). 2006. Trends in the agricultural sector 2015. Pretoria: Department of Agriculture, Forestry & Fisheries.

DAFF (Department of Agriculture, Forestry & Fisheries). 2007. Trends in the agricultural sector 2015. Pretoria: Department of Agriculture, Forestry & Fisheries.

DAFF (Department of Agriculture, Forestry & Fisheries). 2008. Trends in the agricultural sector 2015. Pretoria: Department of Agriculture, Forestry & Fisheries.

DAFF (Department of Agriculture, Forestry & Fisheries). 2009. Trends in the agricultural sector 2015. Pretoria: Department of Agriculture, Forestry & Fisheries.

DAFF (Department of Agriculture, Forestry & Fisheries). 2010. Trends in the agricultural sector 2015. Pretoria: Department of Agriculture, Forestry & Fisheries.

DAFF (Department of Agriculture, Forestry & Fisheries). 2011. Trends in the agricultural sector 2015. Pretoria: Department of Agriculture, Forestry & Fisheries.

DAFF (Department of Agriculture, Forestry & Fisheries). 2012. Trends in the agricultural sector 2015. Pretoria: Department of Agriculture, Forestry & Fisheries.

DAFF (Department of Agriculture, Forestry & Fisheries). 2013. Trends in the Agricultural Sector 2015. Pretoria: Department of Agriculture, Forestry & Fisheries.

DAFF (Department of Agriculture, Forestry & Fisheries). 2014. Trends in the agricultural sector 2015. Pretoria: Department of Agriculture, Forestry & Fisheries.

DAFF (Department of Agriculture, Forestry & Fisheries). 2015. Trends in the agricultural sector 2015. Pretoria: Department of Agriculture, Forestry & Fisheries.

Desai, V.S., Crook, J.N. & Overstreet, G.A. 1996. A comparison of neural networks and linear scoring models in the credit union environment. *European Journal of Operational Research* 95(1): 24-37.

Du Jardin, P. 2012. The influence of variable selection methods on the accuracy of bankruptcy prediction models. *Bankers, Markets & Investors* 116(January): 20–39. [Online]. Retrieved from: https://mpra.ub.uni-muenchen.de/44383/1/MPRA_paper_44383.pdf [20 October 2016].

Durguner, S., Barry, PJ. & Katchova, A.L. 2006. *Credit scoring models: A comparison between crop and livestock farms*. Paper presented at American Agricultural Economics Association Meeting, Long Beach, California. 23–26 July 2006.

Eletter, S.F., Yaseen, S.G. & Elrefae, G.A. 2014. Neuro-based artificial intelligence model for loan decisions. *American Journal of Economics and Business Administration* 2(1): 27-34.

Ezeh, C.I. 2003. Credit worthiness and determinants of loan repayment of smallholder farmers in Abia State, Nigeria. *Journal Trap. Agricultural Research* 5: 10-13.

FFSC. 2011. *Financial guidelines for agricultural producers*. USA: Farm Financial Standards Council.

Gitman, L. J., Smith, M.B., Hall, J., Mikina, D., Malan, M., Marx, J., Mestry, R., Ngwenya, S. & Strydom, B. 2014. *Global and southern African perspectives: Principles of managerial finance* (2nd Edition). Pearson Publishers.

Gujarati, D.N. 2003. *Basic econometrics*. (4th Edition). New York, NY: The McGraw-Hill.

Günther, F. & Fritsch, S. 2010. Neuralnet: Training of neural networks. *The R Journal* 2(1): 30-38.

Haden, K.L. & Johnson, L.A. 1989. Factors which contribute to the financial performance of selected Tennessee dairies. *Southern Journal of Agricultural Economics* July: 105-112.

Hand, D.J. & Henley, W. E. Statistical classification methods in consumer credit scoring: A review. *Journal of Royal Statistical Society* 160(3): 523-541.

Hand, D.J., Sohn, S.Y. & Kim, Y. 2005. Optimal bipartite scorecards. *Expert Systems with Applications* 29(3): 684-690.

Henning, J.I.F. 2016. *Credit scoring model: Incorporating entrepreneurial characteristics*. PhD Thesis. Bloemfontein, Republic of South Africa: University of the Free State.

Henning, J.I.F. & Jordaan, H. 2015. Investigating factors considered in agricultural credit applications, what are currently considered by commercial banks? In H. Watson, M. Lipari, S. Gendron, M-C. Bouchard, S. Couture, N. Nadeau. *Proceedings I, Healthy Agriculture for a Healthy World.* 20th International Farm Management Congress, Quebec City, Canada. 12–17 July 2015.

Henning, J.I.F. & Jordaan, H 2016. Determinants of financial sustainability for farm credit applications – a Delphi study. *Sustainability* 8(1): 77. Retrieved from: http://www.mdpi.com/2071-1050/8/1/77 [23 October 2016].

Hoppe, R. 2015. Profit margin increases with farm size. *Amber Waves*, 25.

Jouault, A. & Featherstone, A.M. 2006. *Determining the probability of default of agricultural loans in a French bank.* Presentation at the American Agricultural Economics Association Annual Meeting, Long Beach, California. 23-26 July 2006.

Kajananthan, R. & Velnampy, T. 2014. Liquidity, solvency and profitability analysis using cash flow ratios and traditional ratios: the telecommunication sector in Sri Lanka. *Research Journal of Finance and Accounting* 5(23):164-170.

Katchova, A.L. & Barry, P.J. 2005. Credit risk models and agricultural lending. *American Journal of Agricultural Economics* 87(1): 194-205.

Kim, J. 2005. *A credit risk model for agricultural loan portfolios under the new Basel Capital Accord*. PhD Dissertation. USA: University of Texas A&M.

Kohansal, M.R. & Mansoori, H. 2009. *Factors affecting on loan repayment performance of farmers in Kharasan-Razavi province of Iran*. A paper presented at a conference on International Research on Food Security. Natural Resource Management and Rural Development, University of Hamburg. 6-8 October 2009.

Kumara, A., Singh, K.M. & Sinha, S. 2010. Institutional credit to agriculture sector in India: Status, performance and determinants. *Agricultural Economics Research Review* 23(2): 253-264.

Lee, T. & Chen, I. 2005. A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines*. Expert Systems with it Applications* 29(1): 734-752.

Leea, T-S., Chiub, C-C., Luc, C-J. & Che, I-F. 2002. Credit scoring using the hybrid neural discriminant technique. *Expert System with Applications* 23(2002): 245-254.

Limsombunchai, V., Gan, C. & Lee, M. 2005. An analysis of credit scoring for agricultural loans in Thailand. *American Journal of Applied Sciences* 2(8):1198-1205.

Li, X-L. & Zhong, Y. 2012. An overview of personal credit scoring: Techniques of future work. *International Journal of Intelligence Science* 2: 181-189.

Lubinda, M. 2010. *The evaluation of credit risk in structured finance lending transactions in agriculture.* Master's Dissertation. Bloemfontein, Republic of South Africa: University of the Free State.

Mahmood, A.N., Khalid, M. & Kouser, S. 2009. The role of agricultural credit in the growth of livestock sector: A case study of Faisalabad. *Pakistan Veterinary Journal* 29(2): 81-84.

Marqués, A.I., García, V. & Sánchez, J.S. 2013. A literature review on application of evolutionary computing to credit scoring. *Journal of the Operational Research Society* 64(2013): 1384-1399.

Mester, L.J. 1997. What is the point of credit scoring? *Federal Reserve Bank of Philadelphia Business Review* September/October: 3-16.

Michael, W., Miller, C. & Gegenbauer, I. 2009. The use of structured finance instruments in agriculture in Eastern Europe and central Asia. Agricultural Management, Marketing and Finance Services (AGSF) Rural Infrastructure and Agro-Industries Division. Rome, Italy: Food and Agriculture Organisation of the United Nations.

Miller, A., Dobbins, C., Prichett, J., Boehlje, M. & Ehmke, C. 2004. Risk management for farmers*. Staff Paper* 4-11. Department of Agricultural Economics, Purdue University [Online]. Retrieved from: http://ageconsearch.umn.edu/bitstream/28640/1/sp040011.pdf [9 August 2016].

Mishra, A.K., Hisham, S.E-O. & Johnson, J.D. 1999. Factors contributing to earnings success of cash grain farms. *Journal of Agricultural and Applied Economics* 31(3): 623-637.

Mohammadi, N. & Zangeneh, M. 2016. Customer credit risk assessment using artificial neural networks. *International Journal of Information Technology and Computer Science (IJITCS)*, 8(3): 58-66. DOI: 10.5815/ijitcs.2016.03.07

Nayak, G.N. & Turvey, C.G. 1997. Credit risk assessment and opportunity costs of loan misclassification. *Canadian Journal of Agricultural Economics* 45(1997): 285-299.

Nieuwoudt, S. 2016. *Entrepreneur characteristics and financial performance.* Master's Dissertation. Bloemfontein, Republic of South Africa: University of the Free State.

Nwachukwu, I.N., Alamba, C.S. & Oko-Isu, A. 2010. Determinants of institutional credit repayment performance among farmers in Afikpo North LGA of Ebonyi State, Nigeria. *AAB BIOFLUX* 2(3): 211-226.

Nwankwo, U.M. 2004. *Impact of community banks on women farmers' poverty level in rural Abia State Nigeria*. Master's Dissertation. Abia State, Nigeria: University of Uturu.

Oden, O.O., Featherstone, A.M. & Sanjoy, D. 2006. *Predicting credit default in an agricultural bank: Methods and issues*. Paper presented at the Southern Agricultural Economics Association Annual Meetings, Florida. 5-8 February 2006.

Ong, C.S., Huang, J.J. & Tzeng, G.H. 2005. Building credit scoring models using genetic programming. *Expert Systems with Applications* 29(1): 41-47.

Oladeebo, J.O. & Oladeebo, O.E. 2008. Determinants of loan repayment among smallholder farmers in Ogbomoso Agricultural Zone of Oyo State, Nigeria. *Journal of Social Sciences* 17(1): 59-62.

Olagunjiu, F.I. & Adeyemo, R. 2007. Determinants of repayment decision among small holder farmers in southwestern Nigeria. *Pakistan Journal of Social Sciences* 4(5): 677-686.

Olagunju, F.I. & Ajiboye, A. 2010. Agricultural lending decision: A Tobit regression analysis. *African Journal of Food, Agriculture, Nutrition and Development* 10(5): 2515-2541.

Onyenucheya, F. & Ukoha, O.O. 2007. Loan repayment and credit worthiness of farmers under the Nigerian Agricultural Cooperative Rural Development Bank NACRDB. *Agricultural Journal* 2(2): 265-270.

Oriema, O.J. 2010. *Behavioural credit scoring model for credit cardholders*. Master's Dissertation. Nairobi, Kenya: University of Nairobi.

Paliwal, M. & Kumar, U.A. 2009. Neural networks and statistical techniques: A review of applications. *Expert Systems with Applications* 36(1): 2-17.

Petracco, C.K. & Perder, J. 2009. *Evaluating the impact of land tenure and titling on access to credit in Uganda.* International Food Policy Research Institute (IFPRI) Discussion Paper [Online]. Retrieved from http://www.ifpri.org/publication/evaluating-impact-land-tenure-and-titling-access-credit-uganda [26 October 2016].

Petrick, M. 2005. Empirical measurement of credit rationing in agriculture: A methodological survey. *Journal of Agricultural Economics* 33(2005): 191-203.

Phelan, C. 2014. *Understanding the farmer: An analysis of the entrepreneurial competencies required for diversification to farm tourism*. PhD Thesis. Lancashire, England: University of Central Lancashire.

Quaye, F., Haratrska, V. & Nadolnyak, D. 2015. Farmer credit delinquency in southern US: Factors and behaviour prediction. Southern Agricultural Economics Association. 2015 Annual Meeting, Atlanta Georgia. January 31-3 February 2015.

Qwabe, N.P. 2014. *Lending to small-scale farmers in South Africa: A case for best practices in formal institutions.* Master's Dissertation. Pretoria, Republic of South Africa: University of Pretoria.

SACAU Outlook. 2016. SACAU to take up southern African drought. *Southern African Confederation of Agricultural Unions*, 22 January 2016. [Online]. Retrieved from http://www.sacau.org/blog/2016/01/22/sacau-to-take-up-southern-african-drought/ [30 November 2016].

Salame, E. 2011. *Applying data mining techniques to evaluate applications for agricultural loans*. PhD Thesis. Nebraska, USA: University of Nebraska.

Sarlija, N., Bensic, M. & Zekic-Susac, M. 2009. Comparison procedure of predicting the time to default in behavioural scoring. *Expert Systems with Applications* 36(5): 8778-8788.

Sebe-Yeboah, G. & Mensah, C. 2014. A critical analysis of financial performance of agricultural development bank (ADB, Ghana). *European Journal of Accounting Auditing and Financial Research* 2: 1-23.

September, M.T. 2010. *Credit risk management: loans to high risk agricultural clients in central South Africa.* (Master's Dissertation). Bloemfontein, Republic South Africa: University of the Free State.

Sharma, A. & Chopra. 2013. Artificial neural networks: Applications in management. *Journal of Business and Management* 12(5): 33-40.

Spencer, H.M. 1997. *Contemporary macroeconomics*. (3rd Edition). New York: Worth Publishers.

Steenackers, A. & Goovaerts, M.J. 1989. A credit scoring model for personal loans. *Insurance: Mathematics and Economics* 8(8): 31-34.

Stergiou, C. & Siganos, D. s.a. *Neural networks*. [Online]. Retrieved from: http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html#Historical_background [18 October 2015].

Sustersic, M., Mramor, D. & Zupan, J. 2009. Consumer credit scoring models with limited data. *Expert System with Applications* 36(3): 4736-4744.

Thomas, L.C., Edelman, D.B. & Crook, L.N. 2002. *Credit scoring and its applications*. Philadelphia: Society for Industrial and Applied Mathematics.

Tsai, C-F. & Wu, J-W. 2008. Using neural networks enables for bankruptcy prediction and credit scoring. *Expert Systems with Applications* 34(2008): 2639-2649.

Tufféry, S. 2011. *Data mining and statistics for decision making.* (10th Edition). United Kingdom: Wiley Publishers.

USAID. 2005. *The credit process.* USAID-Funded Economic Governance II Project [Online]. Retrieved from: http://pdf.usaid.gov/pdf_docs/Pnadq084.pdf [29 March 2016].

Van der Merwe, E. 2011. *Economic literacy as a factor affecting allocative efficiency.* (Master's Dissertation). Bloemfontein, Republic South Africa: University of the Free State.

West, D. 2000. Neural network credit scoring models. *Computers & Operations Research* 27(2000): 1331-1152.

Yakubu, Z. & Affoi, A.Y. 2014. An analysis of commercial banks' credit on economic growth in Nigeria. *Current Research Journal of Economic Theory* 6(2): 11-15.

Yang, Z., Wang, Y., Bai, Y. & Zhang. X. 2004. Measuring scorecard performance. *Computational Science* 3039: 900-906.