Deriving Business Failure through the use of Predictive Modelling and Analytical Techniques

Philip D. Pretorius and Nomvula L. Motale
School of Mathematical and Statistical Sciences,
Faculty of Natural and Agricultural Sciences,
North West University
Vanderbijlpark, South Africa
Philip.Pretorius@nwu.ac.za

Abstract

The South African business banking sector has seen a high number of new to the bank, business banking customers closing their accounts within three years of having their account and operating their business, which then leads to them ending their relationship with the bank, either due to lack of funding, business insight or support from various stakeholders. Through the use of Predictive Modelling and Analytical Techniques, variables significant to business failure can be identified and a customer retention strategy can be created, to help business banks to successfully retain business banking customers and provide them with knowledge to help grow and groom their business. A logistic regression model is used with the objective to explain the relationship between the dependent binary variable and the independent predictive variables. For the interpretation of the logistic regression, the Gini coefficient as well as the information value is looked at. The logistic model is evaluated using the confusion matrix as well as the receiver operating characteristic chart. Then conclusions are drawn from this study.

Keywords
Business Bank, Business Failure, Customer Retention Strategy, Analytical Tools

1. Introduction

In South Africa the success or failure of a business contributes to the stability and growth of the economy. For a South African business bank whenever a business banking customer's business fails, the loss of time, cost and effort in managing that customer is absorbed by the bank. Given the current increase in the number of business closures and business failures. This then affects the country's GDP target and the National Development Plan, which is to develop entrepreneurs and to grow the economy. Business failure increases the unemployment rate of the country. Employees will be retrenched if the business cannot be sustained.

According to Castano *et al.* (2017: 60), factors such as the business age, business size, the industry the business is in, business financial standing and the business risk could lead to business failure whilst the reasons thereof differ according to the researchers focus in the study. Williams (2014) suggests that the causes of business failure are associated with management failure, failed marketing strategies, failure in customer retention, failure to manage finance as well as systems and structural failure. Businesses are said to usually fail due to internal factors and not external factors as a result of lack of good management decisions according to organisation ecology scholars who study business failure.

Business size is said to be a contributing variable in terms of business failure, based on business turnover and the number of employees a business has (Bloodgood *et al.,* 1996; Williams 2011, 2014).

Business age is seen as a predictive variable for business failure as it shows the experience and maturity of the business (Satendra *et al.,* 2012, Williams, 2014). Older businesses are said to have a better prospect of continued existence than smaller or younger businesses as they have more experience in their respective industries (Pretorius, 2009).

Industry plays a role in business failure, as the industry in which a business operates in plays a role in its ability to succeed (Campbell et al., 2012:90; Avi-Yonah *et al.*, 2017:7). Access to resources is limited to specific sectors, the same way performance differs per sector (Avi-Yonah *et al.*, 2017:12). The level of competition and influence of other factors in a sector determines whether a business will succeed in the sector or exit the sector. Businesses with greater resources will be more likely to survive and support smaller firms in the same industry, which will lead to other smaller firms surviving compared to sectors where there is a lack of support in terms of resources. However, there are cases where bigger businesses with more resources do not want to support small businesses to eliminate the competition in the same industry.

Financial resource is an important variable that is easily observed especially when it is time to release financial reports. A business' financial standing plays a role in its ability to get credit, manage its funds, business credit score and turnover (Anani, 2010). When a business owner does not invest a large capital amount, it may indicate that the owner might want to take time to learn about operating the business instead of expanding the business immediately. Therefore, one can assume that investing less capital into a business could lead to business closure (Williams, 2014; Avi-Yonah *et al.*, 2017:14).

Some of the variables above will be used in the study together with other financial variables that are relevant to the business bank, as the aim is to build a propensity model to predict business failure using financial data. A propensity model is a statistical model that is used to help predict the behaviour of a customer.

2. Literature Review

Customer longevity equals profitability: Studies have shown that the more customers a business can retain the more profitable a business can become. Only with a thorough understanding of customer retention factors can customer attrition decrease (Anvari & Amiem, 2010:17-18). Based on previous studies, there seems to be a correlation among profit and market share where they associated a greater return on investment to market share. Banks with greater revenue can build more branches across the country, which will lead to smaller banks actually closing down their business (Abir & Chokri, 2010:17). According to Bhattacharjee *et. al.* (2009) in South Africa about eight out of ten new businesses fail within three years of operating due to micro- and macro-economic factors that affect businesses. Poor economic conditions and the business industry in which the business is conducted are common causes of business bankruptcies. Businesses that run the risk of bankruptcy may find an exit route, which will allow them to be acquired such that their assets may be redeployed by forming friendly merges that are not affected by distrain.

The environment in which a business is conducted makes the business vulnerable to digital change, economical shift and regulatory changes to name a few, this puts pressure on business management's strategy, which, if not managed, could lead to business failure. The level of competition and influence of other factors in a sector determines whether a business will succeed in the sector or exit the sector. Businesses with greater resources will be more likely to survive and support smaller firms in the same industry, which will lead to other smaller firms surviving compared to sectors where there is a lack of support in terms of resources (Dias & Teixeira, 2014). According to Franco *et al.* (2009) and Castano *et al.* (2017), small businesses are said to be more flexible in terms of making decisions and making business contingency plans. This allows them to be able to call emergency meetings and make decisions quicker for the benefit of the business as there are fewer members involved in the decision making of the business. For example, most small businesses are mainly managed by family members, an individual or a small group of owners, which makes it easier for small business owners to come together to make a decision regarding their business. Large businesses find difficulty in terms of making decisions regarding a business' future prospects. Older businesses are said to have a better prospect of continued existence than smaller or younger businesses as they have more experience in their respective industries (Pretorius, 2009).

In the increasingly competitive business banking sector, banks are opposing high debt levels, competing with other banks and offering duplicate products to their business banking customers. South Africa has more than five banks that offer business banking services to customers who either operate or manage a registered business and have a business

bank account (BusinessTech, 2017). Banks play a critical role in economic growth through investment lending, offering loans, accepting deposits and ensuring that they adhere to the policies and procedures of the South African Reserve Bank (SARB), whose main function is to manage South African money and its banking system (Nhundu, 2016).

Jain *et al.* (2017) define customer retention as a measure that companies, businesses and organisations can take to reduce the number of customers that attrite by retaining as many customers as possible, whether new or existing, to ensure that they do not go to their competitors. However, most efforts are put on retaining existing customers as it is easier to adopt new retention strategies. It is usually more expensive to acquire new to bank customers, as more money is spent on acquiring new customers than retaining existing customers (De Meyer *et al.*, 2010:27-42). Retaining existing customers is more important and cost effective than acquiring new customers, as more costs are incurred at the beginning of the bank and customer relationship. Most studies have only focused on retaining customers and a business customer's happiness, however not looking at them simultaneously by associating them to one another in the form of a customer retention model. When a customer retention strategy is not well maintained, no matter how long a customer banks with a specific bank, the customer can still get out of the relationship with the bank at any time, regardless of how hard bank management and the employees work.

Through the use of data analytics and statistical methodologies, variables significant to business failure can be identified and a customer retention strategy can be created. This is done to help business banks to successfully retain business banking customers and provide them with knowledge to help grow and groom their business.

3. Logistic regression model

This study used a quantitative approach that involves a statistical and/or numerical approach into the research design, it's methodology is consistent with a researcher and are also independent of the researcher (Clark *et. al*., 2003). As a result, the data are used to measure the probability of the objective event taking place and the research gains through collecting and analyzing data.

A logistic regression is a regression analysis method where the outcome is measured with a binary variable that can only be zero or one. The objective of a logistic regression model is to explain the relationship between the dependent binary variable and the independent predictive variables. This is done by creating coefficients of a formula to predict a logit transformation to get the probability of the presence of characteristic of interest such that the prepared model becomes:

$$Logit\ (p) = b_o + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7 + \ldots + b_iX_i$$

where:

p is the probability of presence of the characteristic of interest.
X_1 is variable 1
X_i is variable i

The logit transformation is defined as the log odds, where

$$Logit(p) = \ln\,(odds) = \ln\,(\frac{p}{1-p})$$

and

$$odds = \frac{p}{1-p} = \frac{probability\ of\ presence\ of\ characteristic}{probability\ of\ absence\ of\ characteristic}$$

By taking the exponential of both sides of the regression equation as given above, the equation can be rewritten as:

$$odds = \frac{p}{1-p} = e^{b_o} \times e^{b_1X_1} \times e^{b_2X_2} \times e^{b_3X_3} \times \ldots \times e^{b_iX_i}$$

It is clear that when a variable X_i increases by one unit, with all other factors remaining unchanged, then the odds will increase by a factor e^{b_i}

$$e^{b_i(1+X_i)} \div e^{b_i X_i} = e^{b_i + b_i X_i - b_i X_i} = e^{b_i}.$$

This factor e^{b_i} is the odds ratio (O.R.) for the independent variable X_i and it gives the relative amount by which the odds of the outcome increase (O.R. greater than one) or decrease (O.R. less than one) when the value of the independent variable is increased by one unit. The prepared model will represent predictive variables of business failure.

4. Analytical Tools for Predictive Modelling

4.1 Microsoft Visual Basic Applications

Visual Basic Applications (VBA) is a computer programming language that allows the creation of user-defined functions and the automation of specific processes as well as calculations based on the user's needs. VBA controls Microsoft Office products that are VBA compatible such as Microsoft Word, Microsoft Access and Microsoft Outlook, which come standard with VBA. In order for a VBA programme to work effectively, a user must understand the macro that they are creating and its functionality to help achieve a specific goal. (Anon, 2017)

4.2 SAS

SAS is more favourable to use as it can analyse big data, read almost any data source and perform extractions while joining Microsoft Excel files to SAS data sets through an import statement. SAS gives the user the ability to create and join datasets with millions of rows, whereas Microsoft Excel is limited to just over one million rows and is not designed to do extreme analysis on big data. Through SAS Enterprise Miner and SAS Enterprise guide, a user has the ability to save all temporary datasets in a work library that stores all your datasets until the programme is closed (Shankar, 2012). SAS enterprise guide is a tool developed by the SAS Institute to perform multivariate analysis, business intelligence, predictive analytics and data management. It is a tool that can be to access data in any format including Microsoft excel and SAS data tables, manage data and manipulate existing data to get the data into the format that the user wants and data analysis using statistical techniques such as correlations, descriptive statistics of the data. SAS can be used to do the following:

- Access data in any format including Microsoft excel and SAS data tables.
- Manage data and manipulate existing data to get the data into the format that the user wants.
- Data analysis using statistical techniques such as correlations, descriptive statistics of the data.
- Allows the creation of reports in an understandable format that can be saved in different formats such as PDF (SAS Institute Inc., 2017).

5. Business and Statistical Results

For the interpretation of the logistic regression, the Gini coefficient as well as the information value will be looked at.

The information value is used to measure the predictive power of the predictor and is used to measure the predictive power without any regard to an ordering of a predictor. An information value can be computed for any predictor given that $\boldsymbol{g_k}$ or $\boldsymbol{b_k}$ is not zero.
As a formula, information value is given by:

$$IV = \sum_{k=1}^{L} (g_k - b_k) * \log(g_k / b_k)$$

Where $L \geq 2$ and where g_k and $b_k > 0$ for all $k = 1, \dots, L$

The Gini coefficient is a statistic used to measure the inequality of a distribution whereby the numerator is seen as the area below the Lorenz curve and the 45-degree line. The denominator is the area below the 45-degree line. The Gini coefficient ranges from zero to 100 percent; whereby, the closer the Gini coefficient is 100 percent better, this shows the predictability of each variable. In figure 1 the Gini coefficient and Lorenz curve is shown:

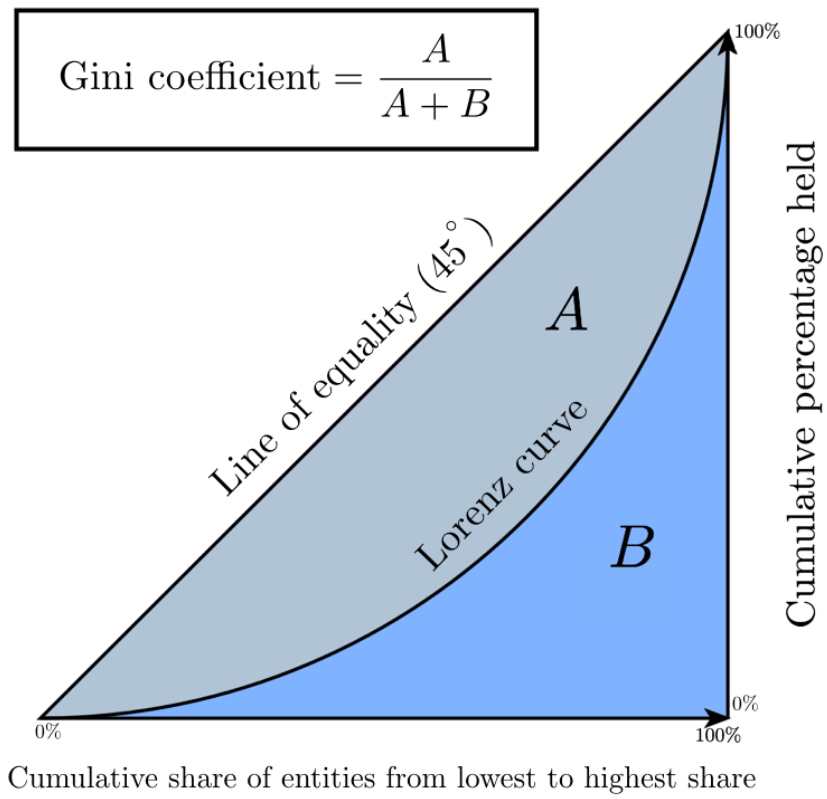

Figure 1. Information value interpretation

In Figure 2 the interpretation of the information value for a binary logistic regression model is shown. Guidelines are taken from Siddiqi (2006:81) such that:

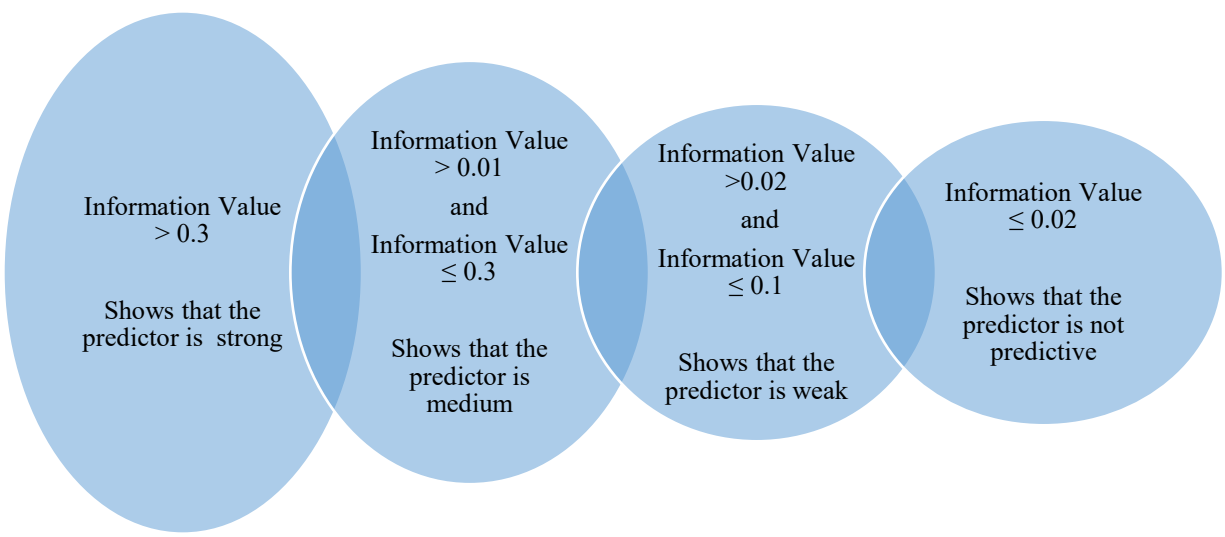

Figure 2. Information value interpretation

Table 1 shows a list of variables that were found to be predictive based on the logistic regression model. The strength of each variable is interpreted using the value of the information value, with the result that five variables were found to be strong predictors, followed by three medium predictors and two weak predictors.

Table 1: Logistic regression predictive variables

Variable	**Gini coefficient**	**Information value**	**Statistics interpretation**
Closing balance amount	41.686	1.279	Strong predictor
Last month's account risk category	25.612	1.237	Strong predictor
Profile risk category	25.577	1.232	Strong predictor
Yearly credit turnover	30.837	0.479	Strong predictor
Account age	31.418	0.349	Strong predictor
Customer type	25.35	0.228	Medium predictor
Number of sub product	20.589	0.179	Medium predictor
Number of cheque accounts	15.957	0.112	Medium predictor
Last deposited amount	10.494	0.042	Weak predictor
Customer total number of products	7.599	0.027	Weak predictor

The following variables are predictive based on the Gini coefficient, information value (IV) and WOE that measures how much of the evidence supports or undermines the hypothesis.

5.1 Closing balance amount

The closing balance amount is a variable that refers to the closing balance amount the customer had in their business banking cheque account prior to closing it.

Table 2: Interactive grouping on closing balance amount variable

Variable	**Group**	**WOE**	**Event count**	**Non-event count**	**Group rate**
Closing balance amount < -R58.17	1	-2.5013	5588	2585	0.684
Closing balance >= -R58.17	2	0.56736	6718	66857	0.091
Original Gini is 41.686 and the Information value (IV) is 1.279					

Table 2 shows that customers who had a negative closing balance amount less than R58.17 have a higher probability of closing their account due to business failure; with a group event rate of 0.684 and a WOE of -2.5013. The high negative of value of WOE correlates to the high risk of business failure whereas high positive values refer to low risk of business failure.

5.2 Last month's account risk category

Last month's account risk category variable refers to the customer's risk category code on their business cheque account at the end of every month.

Table 3: Interactive grouping on last month's account risk category variable

Variable	**Group**	**WOE**	**Event count**	**Non-event count**	**Group rate**
0,1,missing	1	0.29669	9113	69186	0.116
2	2	-3.06178	496	131	0.791
3,4	3	-4.80199	2697	125	0.956
Original Gini is 25.612 and the Information value (IV) is 1.237					

In Table 3 customers who have a risk category of 3 and 4 have a group event rate of 0.956 and a WOE of -4.80199, which shows a likelihood of their businesses failing. Risk category 3 and 4 represent business customers who are at a risk of insolvency and fraud, ideal customers to model with based on their customer behaviour are customers with risk category code of 1 and 0 as they resemble normal customer behaviour, normal profile and have not been detected or monitored for fraud and insolvency.

5.3 Profile risk category

Profile risk category variable refers to the risk category code of the customer based on the customer's overall profile behaviour based on all the products a customer holds within the bank.

Table 4: Interactive grouping on account risk category variable

Variable	**Group**	**WOE**	**Event count**	**Non-event count**	**Group rate**
0,1,missing	1	0.29623	9116	69177	0.116
2	2	-2.99847	494	139	0.78
3,4	3	-4.79365	2696	126	0.955
Original Gini is 25.577 and the Information value (IV) is 1.237					

In Table 4 customers who with a risk category of 3 and 4 have a group event rate of 0.955 and a WOE of -4.79365, which resembles a high probability of business failure. This shows there is a correlation between 'last month's account risk category code' and 'profile risk category' when comparing the WOE, Gini coefficient and IV.

5.4 Yearly credit turnover

The yearly credit turnover variable refers to the turnover a business makes in one year.

Table 5: Interactive grouping on yearly credit turnover variable

Variable	Group	WOE	Event count	Non-event count	Group rate
Yearly credit turnover <178.77	1	0.45308	6621	58776	0.101
Yearly credit turnover >=178.77	2	-1.10117	5685	10666	0.479
Original Gini is 30.837 and the Information value (IV) is 0.479					

Table 5 shows that businesses with a yearly credit turnover less than R178.77 yearly have a high event rate of 0.348 with an event count of 5685. The calculated WOE is -1.101; the smaller your turnover the higher the probability of you closing your account.

5.5 Number of products a customer has

The number of products variable refers to the number of products a business banking customer has within the bank. The vertical sales index (VSI) that is used within the bank tracks the number of products a customer holds within the bank.

Table 6: Interactive grouping on customer total number of products variable

Variable	Group	WOE	Event count	Non-event count	Group rate
Customer total number of products <2	1	-0.34975	8585	34147	0.201
Customer total number of products >2	2	0.51934	3721	35295	0.179
Original Gini is 20.589 and the Information value (IV) is 0.179					

Table 6 shows us that those customers who have less than two products within the bank have a high probability of leaving the bank, looking at their WOE of -0.34975. The more products a customer has, the less likely it is for them to close their cheque accounts and leave the bank. However, if a customer has less product holdings the more likely it is for them to end their banking relationship.

6. Model Validation

The confusion matrix is described in Table 7 for the binary case.

Table 7: Confusion matrix

True positive: customers **correctly** predicted for business failure.	**False positive**: customers **incorrectly** predicted as other.
False negative: customers **incorrectly** predicted for business failure.	**True negative**: customers **correctly** predicted as other.

The confusion matrix in Table 8 below describes the performance of the logistic regression model and confirms its predicting power:

Table 8: Confusion matrix for the fitted logistic regression model

		PREDICTED TARGET	
ACTUAL TARGET		**1**	**0**
	1	6775	10341
	0	2768	100056

Sensitivity is the percentage of customers who are correctly identified to have closed their accounts due to business failure. Specificity is the percentage of customers who are correctly identified to have not closed their accounts due to business failure. Accuracy is a measure of how often the logistic regression model is correct. The figures are calculated as follows.

$$\textbf{\textit{Sensitivity}} = \frac{True\ Positive}{True\ Positive + False\ Negative} = 0.709944$$

$$\textbf{\textit{Specificity}} = \frac{False\ Positive}{False\ Positive + True\ Negative} = 0.906329$$

$$\textbf{\textit{Accuracy = (Correctly Predicted)/Total}} = 0.890704$$

$$\textbf{\textit{Correctly Predicted}} = True\ Positive + True\ Negative = 106831$$

$$\textbf{\textit{Total}} = True\ Positive + True\ Negative + False\ Positive + False\ Negative = 119940$$

The data partitioning process is used to assess the quality of the model from the data source. Data partitioning consists of a data set that is used for training where its primary function is the first step of model fitting, whereby the rest of the observations are kept aside for empirical validations.

Table 9 gives the misclassification rate of the logistic regression model is 0.10380=10.38 percent, which is a false positive rate that shows the proportion of misclassified observations. The logistic regression has a prediction power of 89 percent, which is far greater than that of a linear regression model.

Table 9: Confusion matrix for the fitted logistic regression model

Model	Node	Model description	Train - misclassification rate	Train - squared error	Valid - misclassification rate	Valid - squared error
Y	Reg3	Logistic regression	0.10380	0.084064	0.10519	0.082795

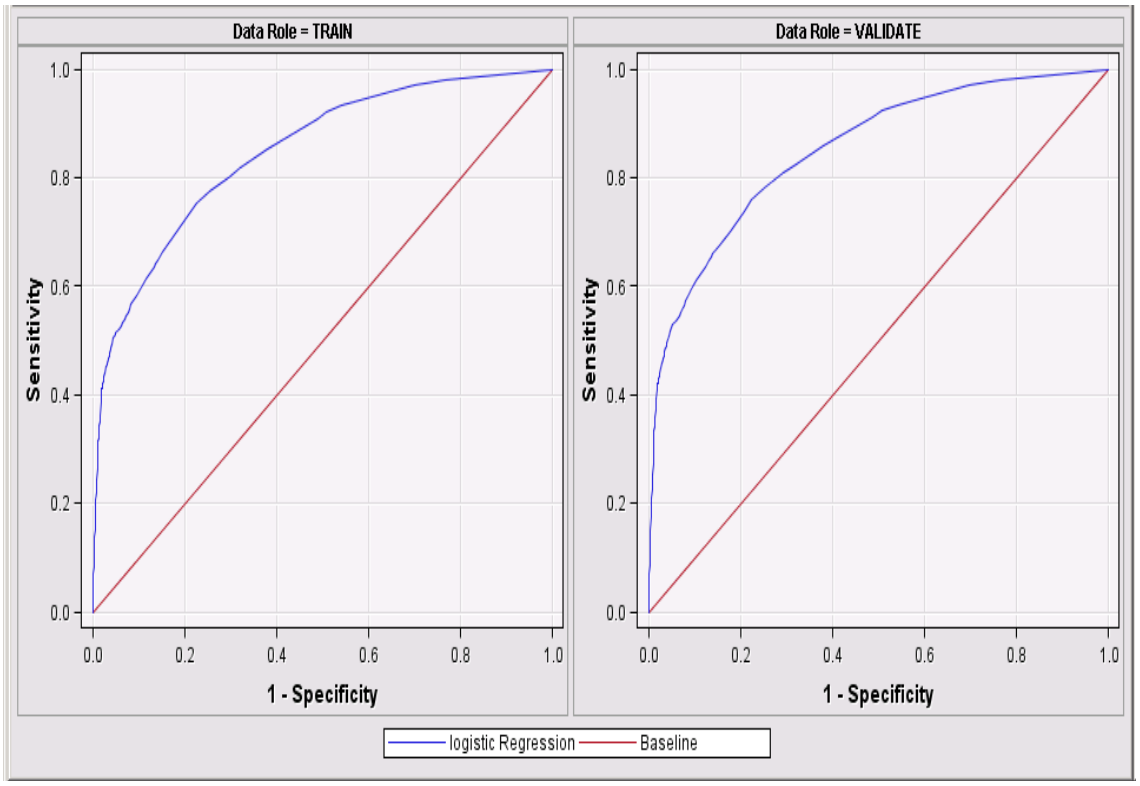

Figure 3: Logistic regression Receiver Operating Characteristic (ROC) chart

In Figure 3 is a plot of the Receiver Operating Characteristic (ROC) curve for Logistic Regression, indicating significant lift above the Baseline, illustrating that the Logistic Regression model is more sensitive than the Baseline (Random selection). The ROC curve for the validation set is similar to the training set, indicating a good model for prediction.

By determining customers who are more likely to close their business cheque accounts due to business failure, a cumulative lift chart is used to explain the effectiveness of the model. The axis of the cumulative graph reflects different outcomes whereby the vertical axis shows the lift of the chart and the horizontal axis shows the percentile. A cumulative lift curve shows the advantage of using a predictive model to determine, which customers to communicate with regarding business failure as a customer retention strategy to remedy the situation.

7. Conclusion

The conclusions are as follows:

Analytical tools and statistical methodologies can be used to predict and detect business failure across the business banking sector. The SAS tool is used to perform data analysis, data management, data manipulation and importing of data, to name a few. This helped to process thousands of rows in a short period of time thus being very effective in providing results. The SAS tool helps banks and businesses to increase their business functions.

Business banking financial significant variables were identified, to predict if a business customer's business is failing. The following variables were found to be significant strong predictors that can be used to predict and identify business failure; closing balance amount, last month's risk category, profile risk category, yearly credit turnover, account age, customer type and product holdings. The more products a customer has, the less chances of customer attrition, therefore, the more products a customer has, the less likely it is for them to end their relationship with the bank. If a customer has a long relationship with the bank the less likely it is for a customer to attrite.

Using the scoring code from the logistic regression model, probabilities of business failure can be obtained by running the scoring code on a current business banking base. The primary objective which is to test if modelling can be used to predict business failure in the South African business bank, has been achieved.

8. References

Abir, M. & Chokri, M. 2010. Dynamic of financial innovation and performance of banking firms: context of an emerging banking industry. *International Research Journal of Finance & Economics*, 51:17.

Anon. 2017. What is VBA. https://www.emagenit.com/VBA%20Folder/what_is_vba_visual_basic_for_applications.htm / Date of access: 8 Nov. 2017.

Anani, A.O. 2010. Attracting and retaining customers in South Africa's banking sector. Port Elizabeth: NMMU. (Mini-dissertation – MBA).

Anvari, R. & Amien, S.M. 2010. The customer relationship management strategies: personal needs assessment of training and customer. *European Journal of Social Sciences*, 14(1):17-33.

Avi-Yonah, R.S., Batchelder, L.L., Fleming,J.C, Gamage, D., Glogower, A.D., Hemel, D.J., Kamin, D., Kane, M., Kysar, M.R., Miller, D.S., Shanske, D., Shaviro, D. & Viswanathan, M. 2017. The games they will play: tax games, roadblocks, and glitches under the house and senate tax bills. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3084187 Date of access: 15 Dec. 2017.

Bhattacharjee, A., Higson, C., Holly, S. & Kattuman, P. 2009. Macroeconomic instability and business exit: determinants of failures and acquisitions of uk firms. *Economica,* 76(301), 108-131. doi:10.1111/j.1468-0335.2007.00662.x.

Bloodgood, J.M., Sapienza, H.J. & Almeida, J.G. 1996. The internationalization of new high-potential U.S. ventures: antecedents and outcomes. *Entrepreneurship: Theory and Practice,* 20(4):61-76.

BusinessTech. 2017. The battle of how South African banks compare. https://businesstech.co.za/news/banking/182873/battle-of-the-banks-how-sas-big-five-banks-compare/ Date of access: 3 Feb. 2018.

Campbell, N.D., Heriot, K.C., Jauregui, A. & Mitchell, D.T. 2012. Which state policies lead to us firm exits? analysis with the economic freedom index. *Journal of Small Business Management, 50*(1):87-104.

Castano, F.J. & Tascon, M.T. 2017. Selection of variables in small business failure analysis. *Journal of Quantitative Methods for Economics and Business,* 24:54-88.

Clark, V.L., Creswell, J.W., Gutman, M.L. & Hanson W.E. 2003. Advance mix methods research designs. https://books.google.co.za/books?hl=en&lr=&id=F8BFOM8DCKoC&oi=fnd&pg=PA209&dq=article+by+creswell+2003&ots=gVdTwFvyMj&sig=lLmvSWVVt4tako2OZqwT2M-8yDA#v=onepage&q=article%20by%20creswell%202003&f=false Date of access: 25 Apr. 2017

De Meyer, C. & Mostert, P. 2010. Building customer relationships as retention strategy in the South African domestic passenger airline industry. *Acta Commercii,* 10 (1): 27 – 42.

Dias, A.R. & Teixeira, A. 2014. The anatomy of business failure. a qualitative account of its implications for future business success. *University of Porto-FEP Working Papers, 550/2014*, 1-25.

Franco, M., & Haase, H. 2009. Failure factors in small and medium-sized enterprises: qualitative study from an attributional perspective. *International Entrepreneurship and Management Journal,* 6(4):503-521. Doi: 10.1007/s11365-009-0124-5

Jain, R. & Syed, T. 2017. A study on the determinants of customer retention in banking sector. *International Education and Research Journal*, 3(6). http://ierj.in/journal/index.php/ierj/article/view/1216 .

Nhundu, N. 2016. An overview of the prominent barriers to entry in South Africa's retail banking industry. http://www.competition.org.za/review/2016/2/19/an-overview-of-the-prominent-barriers-to-entry-in-south-africas-retail-banking-industry Date of access: 10 Mar, 2017.

Pretorius, M. 2009. Defining business decline, failure and turnaround: a content analysis. *The Southern African Journal of Entrepreneurship and Small Business Management,* 2(1):1. doi:10.4102/sajesbm.v2i1.15

SAS Institute Inc. 2017. SAS®9.3 Language reference: concepts. Cary, NC: SAS Institute Inc.

Satendra, T. & Singh, A. P. 2012. Brand image, customer satisfaction and loyalty intention: a study in the context of cosmetic product among the people of central India. *International Journal of Multidisciplinary Management Studies*, 2(5): 37–50.

Shankar, C. 2012. Do i have to choose between excel and SAS? https://blogs.sas.com/content/sastraining/2012/02/27/do-i-have-to-choose-between-excel-and-sas/ / Date of access: 9 Nov. 2017

Williams, D.A. 2011. Impact of firm size and age on the export behaviour of small, locally-owned firms: fresh insights. *Journal of International Entrepreneurship,* 9(2):152–174.

Williams, D.A. 2014. Resources and business failure in SMEs: does size matter? *Journal of Business and Management*, 20(2):89.

Biographies

Philip D. Pretorius is Professor in the School for Mathematical and Statistical Sciences, with a link to the School for Computer Science and Information Systems, in the Faculty of Natural and Agricultural Sciences, at the North West University in South Africa. He has published journal and conference papers. His research interests include quality, statistics and decision support systems.

Movula L. Motale completed her masters study in the banking environment, and work in the banking environment.