A Model for Predicting Customer Value Towards Upselling in Medium and Large Scale Businesses (MLBs)

Adeyemo, O.A. & Akinselure, A.
1,2Department of Mathematical and Physical Sciences
Afe Babalola University (ABUAD)
Ado-Ekiti, Ekiti State, Nigeria
E-mail: befittingcrown@gmail.com

ABSTRACT

Medium and Large Scale Businesses (MLBs) in Nigeria continue to deploy expensive mass market strategies which has not yielded much result. There is a need to determine the strategy for predicting the customer value that best identify the right customer for the right marketing campaign. This study proposed a model to predict customer value towards upselling in MLBs. The proposed model involves benchmarking the Multiple Linear Regression Model against the Multilayer Perceptron Neural Network. Furthermore, both models are tested on a transactional dataset from a Microfinance bank using the determined predictor values such as customer patronage duration and amount of monthly repayment among others. Consequently, the results from both models show that the Savings balance predictor variable as the most important towards determining customer value for upselling in MLBs. Most businesses can conveniently determine their customer value towards upselling using the proposed model tested in this study to increase their customers’ lifetime value and overall business revenue.

Keywords: Customer Value, Upselling, Predictor, Medium and Large Scale Businesses, Market Strategy, MLPNN.

1. INTRODUCTION

The important parts of Customer Relationship Management (CRM) activities are understanding customers’ profitability and retaining profitable customers and in order to cultivate the full profit potentials of customers, many businesses already try to measure and use customer value in their management activities (Rosset et al., 2002; Hawkes, 2000). Upselling is a technique in CRM and marketing where a seller persuades the customer to purchase more expensive items, upgrades or include additional items all in an attempt to make a more profitable sale, thereby marketing a profitable service or product (Sharma and Sharma, 2015). In recent times, the focus on products and services has shifted largely to a focus on customers (Vargo and Lusch, 2006).
Additionally, this focus was confirmed by parallel evolutions in the understanding of the inter-relationships between customer satisfaction, trust and commitment (Garbarino and Johnson, 1999). In the evaluation of customer profitability, marketers are often reminded of the 80/20 rule which states that 80% of the profits are produced by top 20% of profitable customers and 80% of the costs are produced by top 20% of unprofitable customers (Duboff, 1992).

The Medium and Large Scale Businesses MLB plays a very important role in economic growth and are considered as the backbone of industrial development in any nation especially in developing countries. The advances in Information Technology allow for the following; customer data, recent customer actions, and business objectives that is used to create models for optimizing customer interactions. This would have many benefits, the first being improving the Return on Investment (ROI) of the campaign by increasing the conversion rate, and finally reducing campaigns costs and optimizing the effort (Nair and Mitchell, 2011).

In the last two decades, many studies have been carried out on the efforts of Medium and Large businesses in boosting economic development (Muritala et al., 2012, Siaka-Momoh, 2003 and Doguwa et al., 2011). A major part of the business problem is that business executives face profitability challenges based on execution of the appropriate marketing strategies. However, it is necessary for a business to first determine that the customer relationship has the potential to be profitable before investing the resources required to develop the relationship. The result of this leads to determination of the value of a customer to a business which will foster profitable customers to optimize marketing efforts (Hwang, 2016). Hence, obtaining information on the value of a customer can help businesses to predict their response towards an upsell offer.

This study developed a model for predicting customer value towards upselling in Medium and Large businesses, tested the model on a transactional dataset and identified the actual predictor variable for predicting customer value from the transactional dataset. A microfinance bank is used as a case study. It takes into consideration the available attributes of the transactional data obtained and does not contain demographics factors such as income, age and education. The model was tested on a dataset of active 500 clients currently at Lapo Microfinance bank, Bashiri branch in Ado-ekiti, Ekiti state.

2. LITERATURE REVIEW

A customer is the recipient of a good, service, product or an idea obtained from a seller, vendor, or supplier through a financial transaction, exchange for money or some other valuable consideration (Reizenstein, 2004; Kendall, 2007). In sales, commerce and economics, a customer is also known as a client, buyer, or purchaser. A customer may or may not also be a consumer, a customer purchases goods; a consumer uses them (Blythe, 2008; Kansal and Rao, 2006). An ultimate customer may be a consumer as well, but just as equally may have purchased items for someone else to consume. An intermediate customer is not a consumer at all (Frain, 1999; Reizenstein, 2004). The situation is complicated in that ultimate customers of so-called industrial goods and services (who are entities such as government bodies, manufacturers, and educational and medical institutions) either themselves use up the goods and services that they buy, or incorporate them into other finished products, and so are technically consumers, too. However, they are rarely called that, but are rather called industrial customers or business-to-business customers (Frain, 1999). Similarly, customers who buy services rather than goods are rarely called consumers (Reizenstein, 2004).
Customer Value can be defined as the sum of the revenues gained from a company’s customers over the lifetime of transaction after the deduction of the total cost of attracting, selling, and servicing customers, taking into account of the time value of money (Hwang et al., 2004). To cultivate the full profit potentials of customers, many companies already try to measure and use customer value in their management activities (Gloy et al., 1997; Rosset et al., 2002; Verhoef, and Donkers, 2001). Value in marketing can be defined by both qualitative and quantitative measures. On the qualitative side, value is the perceived gain composed of individual’s emotional, mental and physical condition plus various social, economic, cultural and environmental factors. On the quantitative side, value is the actual gain measured in terms of financial numbers, percentages, and money (Turel et al., 2007).

Upselling is a marketing technique aimed at persuading customers to purchase a more expensive, upgraded or premium version of the chosen item or other add-ons for the purpose of making a larger sale. Businesses teach their employees to upsell products and services and offer incentives and bonuses to the most successful personnel. While Upselling can be viewed by some as just a way to make more cash, sales and marketing professionals should be constantly thinking about how best to meet customers’ needs (Swayne and Dodds, 2011). Cross-selling involves offering the customer other products or services in a different product category. Upselling is similar to the concept of cross-selling; the difference is that upselling involves an upgrade within the same product category. Upselling and Cross-selling are important strategies for increasing revenues among current customers (Blattberg et al., 2010).

A Medium-Scale Business is a business comprising of 50–199 employees while a Large-Scale Business is a business comprising of over 200 employees (Doguwa et al., 2011). Microfinance is the provision of financial services to the poor who are traditionally not served by the conventional banks (Acha, 2012, Jaffari et al., 2011, CBN, 2005 and Conroy, 2003). Otero and Rhyne (1994) viewed microfinance as a system that involves the large scale provision of small loans and deposit services to greater number of active poor, there is need by the members of microfinance institution to allow for increase in capital to cover a wide needs of the low income earners.

2.1 Multiple Linear Regression Model
Regression is the relation between selected values of $x$ and observed values of $y$ from which the most probable value of $y$ can be predicted for any value of $x$. It is the estimation of a real value function based on finite noisy data. Linear Regression was historically the earliest predictive method and is based on the relationship between input variables and the output variable (Mishra et al., 2010).

2.2 Multilayer Perceptron Neural Network
A perceptron is a single neuron model that was a precursor to larger neural networks. It is a field that investigates how simple models of biological brains can be used to solve difficult computational tasks like the predictive modelling tasks we see in machine learning. The goal is not to create realistic models of the brain, but instead to develop robust algorithms and data structures that we can use to model difficult problems.
3. THEORETICAL MODELS USED FOR PREDICTING CUSTOMER VALUE

Customer Value has been studied under the name of LTV (Life Time Value), CLV (Customer Lifetime Value), CE (Customer Equity) and Customer Profitability (Hwang, 2016). The differences between the definitions are relatively small. Zeithaml (1988) investigated the consumer’s perception of value through a literature review and an exploratory, qualitative study using the Means-end Chains model. Verhoef and Donkers, (2001) developed a model for customer potential value from an insurance company’s dataset by comparing different statistical methods: Linear regression analysis, univariate probit, and multivariate probit.

Verstraeten et al., (2005) examined the extent to which the use of an alternative reward system would enable companies to improve the way loyal customers are currently rewarded for the benefits they deliver by using a Variable Selection Technique integrated with a Linear Regression Model.

Kim et al., (2005) showed three perspectives on Customer Value (current value, potential value, and customer loyalty. Current value provided financial viewpoint and potential value indicates cross-selling opportunity by analyzing the characteristics of segmented customer groups, they developed refined strategies for each segment. Bauer and Hammerschmidt, (2005) used the shareholder value (SHV) approach with a Reference Value model to develop an integrated, marketing-based method for corporate valuation.

Buckinx et al., (2006) showed that it is possible to predict a customer’s loyalty to a reasonable degree from the internal transactional database using a Multilinear Regression model, provided that an elaborate variable-selection procedure is performed. Lu and Han (2011) segmented telecom customers based on customer value by decision tree model and they proposed a novel customer segmentation method based on customer lifecycle.

Emtiyaz and Keyvanpour (2011) used a semi-supervised leaning approach working on a bank dataset and an insurance dataset by training with a Multilayer Perceptron Neural Network using labeled data, classify unlabeled data using labeled data, assign label to the unlabeled data with most confidence scores, model evaluation, prediction. Hwang (2016) focused on the implementation of an individualized Customer Lifetime Value model. Customer relationship with a firm is embodied by a state of the Markov chain. The suggested CLV model provided not only the lifetime value of a customer but the fractionized lifetime value of a customer according to the state.
4. METHODOLOGY

This work compared two techniques that show interesting differences and similarities. The models were used based on their ability to deliver adequate predictive performance on a validation set and also to provide an insight into the most important variables in the model as shown in figure 1.

(1) Data Preprocessing
The data for this study is a transactional data of customers running loans at a microfinance bank. The data obtained in a .xls spreadsheet format is loaded into the SPSS Data Editor which is the explorer menu to create an .sav file. SAV is a file extension used for the saved data of SPSS (Statistical Package for the Social Sciences). SAV files contain binary data which can only be used on the platform that created the file.

(2) Descriptive Analysis
Descriptive analysis is used to describe the basic features of the data in the study. They provide simple summaries about the sample and the measures. Frequency tables are used to describe the basic features of this data.

(3) Multiple Linear Regression Model
As a predictive analysis, the Multiple Linear Regression tools is used to explain the relationship between one continuous dependent variable and two or more independent variables. In this study, to predict Customer Value towards upselling, the current loan amount is selected as the dependent variable. The other variables of the dataset will be used as the independent variables.
**Multilayer Perceptron Neural Network**

The Multilayer Perceptron (MLP) procedure produces a predictive model for one or more dependent (target) variables based on the values of the predictor variables. A neural network works by taking the values of predictor or input fields and feeding them into the algorithm as an input layer. The input layer is used to create a hidden layer containing unseen nodes or units where each node is some function of the input fields. Some networks can create more than one hidden layer. The output layer contains the responses or predictions.

5. **IMPLEMENTATION**

**Descriptive Analysis**

Descriptive analysis is performed to provide a summary of the attributes contained in the whole dataset.

**Table 1: Frequency of Loan terms**

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Long Term</strong></td>
<td>136</td>
</tr>
<tr>
<td><strong>Short Term</strong></td>
<td>364</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>500</td>
</tr>
</tbody>
</table>

The table above shows that 27% of the loans obtained by customers in the data are Long term loans while 73% are Short Term loans. Table 2 shows the customer patronage duration.

**Table 2: Customer Patronage duration**

<table>
<thead>
<tr>
<th>Year</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 year</td>
<td>72</td>
<td>14.4</td>
</tr>
<tr>
<td>2 years</td>
<td>48</td>
<td>9.6</td>
</tr>
<tr>
<td>3 years</td>
<td>36</td>
<td>7.2</td>
</tr>
<tr>
<td>4 years</td>
<td>30</td>
<td>6.0</td>
</tr>
<tr>
<td>5 years</td>
<td>42</td>
<td>8.4</td>
</tr>
<tr>
<td>6 years</td>
<td>49</td>
<td>9.8</td>
</tr>
<tr>
<td>7 years</td>
<td>22</td>
<td>4.4</td>
</tr>
<tr>
<td>8 years</td>
<td>20</td>
<td>4.0</td>
</tr>
<tr>
<td>9 years</td>
<td>20</td>
<td>4.0</td>
</tr>
<tr>
<td>10 years</td>
<td>161</td>
<td>32.2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>500</td>
<td>100.0</td>
</tr>
</tbody>
</table>

5.1 **Multiple Linear Regression (MLR) Model**

**Table 3: Model Summary Table**

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.440</td>
<td>0.494</td>
<td>0.484</td>
<td>2.02549</td>
</tr>
</tbody>
</table>

Table 3 is the model summary table. R shows the strength of the relationship between the outcome variable (DV) and the predictor variables (IVs). In this case, R = 0.440, which is a fairly strong relationship. This suggests the model is a relatively good predictor of the outcome. R square value indicates the proportion of
variation in the outcome variable (amount of loan disbursed) that can be explained by the model (i.e. the predictor variables). R square = 0.194 is multiplied by 100 to give a proportion, which implies that 19.4% of the variance in the data can be explained by the predictor variables.

Table 4: ANOVA Table

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Regression</td>
<td>486.218</td>
<td>6</td>
<td>81.036</td>
<td>19.752</td>
<td>.000</td>
</tr>
<tr>
<td>Residual</td>
<td>2022.582</td>
<td>493</td>
<td>4.103</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2508.800</td>
<td>499</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. Dependent Variable: Current Loan Amount
b. Predictors: (Constant), No. instances of Credit Problem, No. of years as a customer, Term, Months since last delinquent, Savings Balance, Monthly Repayment

Table 4 shows whether the model is a significant predictor of the outcome variable (DV). The results indicated that the model was a significant predictor of amount of loan disbursed (DV), $F(6, 493) = 19.752, p = .000$ where $F$ (Regression df, Residual df) = F-Ratio, $p = Sig$

This validates the MLR model and implies that customer’s value towards upselling can be predicted using from customers’ transactional database.

5.2 Multilayer Perceptron Neural Network (MLPNN)
Using the MLPNN model as shown in figure 2, 70% and 30% partition was specified as the relative numbers for training and tests respectively. Automatic architecture selection was used as the default activation functions for the hidden and output layers of the network. Batch training was selected as it directly minimizes the total error of the outputs.

Table 5: Independence Variable Importance Table

<table>
<thead>
<tr>
<th></th>
<th>Importance</th>
<th>Normalized Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term</td>
<td>.171</td>
<td>78.6%</td>
</tr>
<tr>
<td>Savings Balance</td>
<td>.218</td>
<td>100.0%</td>
</tr>
<tr>
<td>No. of years as a customer</td>
<td>.148</td>
<td>67.9%</td>
</tr>
<tr>
<td>Monthly Repayment</td>
<td>.211</td>
<td>96.7%</td>
</tr>
<tr>
<td>Months since last delinquent</td>
<td>.135</td>
<td>62.2%</td>
</tr>
<tr>
<td>No. instances of Credit Problem</td>
<td>.117</td>
<td>53.7%</td>
</tr>
</tbody>
</table>

Table 5 above shows the importance of the predictor variables on customer value towards upselling. It could be deduced from the table that a customer’s savings balance is of the highest importance with 100% importance to determine a customer’s value towards Upselling. The next factor that affects customer value is the monthly repayment at 96.7% importance.
6. CONCLUSION AND AREAS OF FURTHER RESEARCH

Many businesses try to measure and use customer value in their management activities to determine customers’ profitability and identify profitable customers for the purpose of cultivating the full profit potentials of customers. One of such ways is by using the upselling marketing technique. MLBs have been faced with the challenge of executing the appropriate marketing strategies to drive sales and ensure business profitability. Consequently, it is important for MLBs to first determine that their customers have the potential to be profitable before investing in any marketing strategy including upselling marketing efforts.
This study proposed a conceptual model to predict customer value towards Upselling. The process involved testing the MLR and MLPNN model on a transactional dataset from a MFB to gain insight into the most important variables in the model.

The results from the ANOVA table of the MLR model validates the model that it is a significant predictor of the amount of loan disbursed (DV). This validates the MLR model and implies that customer’s value towards upselling can be predicted from customers’ transactional database. The MLR model also allows for the calculation of each customer’s value towards upselling by using their exact values in the following equation. Customer Value = \( \beta_0 + \beta_1 \text{(Term)} + \beta_2 \text{(Savings Balance)} + \beta_3 \text{(No. of years as customer)} + \beta_4 \text{(Monthly Repayment)} + \beta_5 \text{(Months since Last Delinquent)} + \beta_6 \text{(No. of instances of credit problems)} \).

The result from the MLPNN Model was in agreement with that of the MLR Model as the savings balance predictor variable emerged as the most important variable for determining customer value and hence the actual predictor for customer value. The developed model can also be applied to determine the right group of people for a political campaign.

REFERENCES

20. Lu S., Han S. (2016) *Classifying the segmentation of Telecom Customer Value Using Decision Tree Model* (Department of Management Science, Xiamen University, 361005, Xiamen, P.R.China)