I.J. Information Technology and Computer Science, **2012, 1, 9-18**
Published Online February 2012 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijitcs.2012.01.02

A Data Mining-Based Response Model for Target Selection in Direct Marketing

Eniafe Festus Ayetiran
Department of Computer & Information Systems, Achievers University, Owo, Nigeria
Email: eniafef@yahoo.com

Adesesan Barnabas Adeyemo
University of Ibadan, Ibadan, Nigeria
Email: sesan_adeyemo@yahoo.com

Abstract— **Identifying customers who are more likely to respond to new product offers is an important issue in direct marketing. In direct marketing, data mining has been used extensively to identify potential customers for a new product (target selection). Using historical purchase data, a predictive response model with data mining techniques was developed to predict a probability that a customer in Ebedi Microfinance bank will respond to a promotion or an offer. To achieve this purpose, a predictive response model using customers' historical purchase data was built with data mining techniques. The data were stored in a data warehouse to serve as management decision support system. The response model was built from customers' historic purchases and demographic dataset.**

Bayesian algorithm precisely Naïve Bayes algorithm was employed in constructing the classifier system. Both filter and wrapper feature selection techniques were employed in determining inputs to the model.

The results obtained shows that Ebedi Microfinance bank can plan effective marketing of their products and services by obtaining a guiding report on the status of their customers which will go a long way in assisting management in saving significant amount of money that could have been spent on wasteful promotional campaigns.

Index Terms— **Data warehouse, Data Mining, Direct Marketing, Target Selection, Naïve Bayes**

1. Introduction

Large amounts of data are nowadays available to companies about their customers. This data can be used to establish and maintain direct relationship with the customers in order to target them individually for specific products and services offer from the company. Large databases of customers and market data are maintained for this purpose. The customers to be targeted are selected from the database given different types of information such as demographic information and information on the customers' personal characteristics like profession, age and sex and purchase history. Usually, the selected customers are contacted directly through one or some of the following means: personal contact, mail, e-mail, telephone and short message service (sms) to promote the new product or service. This type of marketing is called direct marketing. Among other companies, a growing number of banks and insurance companies are adopting direct marketing as their main strategy for interacting with their customers. Partly due to the growing interest in direct marketing, it has become an important application for data mining. In direct marketing, data mining has been used extensively to identify potential customers for new products. Using historical purchase data, a predictive response model was developed using a data mining technique to predict a probability that a customer is going to respond to a promotion or an offer.

In general, banks and financial services companies in Nigeria use mass marketing as their strategy for promoting a new service or product to their customers. In this strategy, a single communication message is broadcast to all customers through media such as print, radio or television etc. In this approach, companies do not establish a direct relationship with their customers for new product offers. This kind of sales promotion leads to a high waste as only a small proportion of customers respond to these offers. This is largely due to the fact that differences among customers are not put into consideration. As a result, in today's world where products and services are overwhelming and there is a highly competitive market, mass marketing has become less effective.

Consequent upon the ineffectiveness of mass marketing strategy, banks, financial services companies and other companies are shifting away from this strategy and are now targeting subsets of their customer base for specific product and service offers (Direct marketing)[19][21].In direct marketing, companies and organisations try to establish and maintain direct relationship with their customers in order to target them individually for specific product or service offers. Direct marketing is done by sending product offers and information directly to customers through personal contact, sending e-mails or SMS to customers, making phone calls or by addressing customers through post. Nowadays, this type of marketing is being used by growing number of companies as their main strategy for interacting with their customers. But direct marketers in a wide range of industries from banking, manufacturing among others are faced with the challenge of continually rising printing and postage costs, high telephone, internet and transport costs together with high investment cost on new product and services with decreasing response rates from customers, hence the need to target the likely respondents to their products.

Identifying customers who are likely to respond to new offers will be a difficult task through manual perusal of a large customer database. Hence, data mining technique was employed for pattern recognition in customers' dataset in order to make accurate customers' response prediction.

1.1 Data Mining

Data Mining or Knowledge Discovery in Databases can be defined as an activity that extracts some new nontrivial information contained in large databases. The goal is to discover hidden patterns, unexpected trends or other subtle relationships in the data using a combination of techniques from machine learning, statistics and database technologies. This new discipline today finds application in a wide and diverse range of business, scientific and engineering scenarios.

The overall knowledge discovery process was outlined by [22] as an interactive and iterative process involving more or less the following steps: understanding the application domain, selecting the data, data cleaning and preprocessing, data integration, data reduction and transformation, selecting data mining algorithms, data mining, interpretation of the results and using the discovered knowledge. According to [13], data mining tasks can be generally classified into two categories: descriptive and predictive. The former characterizes the general properties of the data in the database. The latter performs inference on the current data in order to make predictions.

Actually, the major reason why data mining has attracted a great deal of attention in the information industry and in the society as a whole in recent years is due to the wide availability of huge amounts of data and the imminent need for turning such data into useful information and knowledge [13]. The data mining process is sometimes referred to as knowledge discovery or KDD (knowledge discovery in databases). The term "KDD" (Knowledge Discovery in Databases) refers to the overall process of discovering useful knowledge from data. There is a difference in understanding the terms "knowledge discovery" and "data mining" between people from different areas contributing to this new field.

Knowledge discovery in databases is the process of identifying valid, novel, potentially useful, and ultimately understandable patterns/models in data. Data mining is a step in the knowledge discovery process consisting of particular data mining algorithms that under some acceptable computational efficiency limitations, finds patterns or models in data [14].

1.2 Bayesian Belief Networks and Naive Bayes Classifier

Classification is the task of identifying the class labels for instances based on a set of features (attributes). Learning accurate classifiers from pre-classified data is a very active research topic in machine learning and data mining. In the past two decades, many algorithms have been developed for learning which include decision-trees and neural-network classifiers. While Bayesian networks (BNs) are powerful tools for knowledge representation and inference under conditions of uncertainty, they were not considered as classifiers until the discovery that Naïve-Bayes, a very simple kind of Bayesian Networks that assumes the attributes are independent, given the class node, are surprisingly effective in classification task [6].

A Bayesian network, belief network or directed acyclic graphical model is a probabilistic graphical model that represents a set of random variables and their conditional dependences via a directed acyclic graph (DAG). For example, a Bayesian network could represent the probabilistic relationships between diseases and symptoms. Given symptoms, the network can be used to compute the probabilities of the presence of various diseases. Formally, Bayesian networks are directed acyclic graphs whose nodes represent random variables in the Bayesian sense: they may be observable quantities, latent variables, unknown parameters or hypotheses. Edges represent conditional dependencies. Nodes which are not connected represent variables which are conditionally independent of each other. Each node is associated with a probability function that takes as input a particular set of values for the node's parent variables and gives the probability of the variable represented by the node. For example, if the parents are m Boolean variables, then the probability function could be represented by a table of 2^m entries, one entry for each of the 2^m possible combinations of its parents being true or false. Efficient algorithms exist that perform inference and learning in Bayesian networks. Bayesian networks that model sequences of variables (*e.g.* speech signals or protein sequences) are called dynamic Bayesian networks. Generalizations of Bayesian networks that can represent and solve decision problems under uncertainty are called influence diagrams.

The Naive (or simple) Bayes classifer finds its origins in the Bayesian theory of probability. Thomas Bayes (1702-1761) created the Bayes ratio formula, the first definition of conditional probability in the 18th century, but most publications on Bayesian networks and classifiers start to appear between 1980 and 1990 as applications for machine learning (Pearl, 1988). The Bayes Naive classifier selects the most likely classification V_{nb} given the attribute values x_1, x_2,........ x_n. This results in:

$$\boldsymbol{Vnb = argmaxvj \epsilon VP(Vj)} \prod \boldsymbol{P(xi \mid Vj)} \qquad \textbf{\textit{(1.1)}}$$

We generally estimate P ($a_i|v_j$) using m-estimates:

$$\boldsymbol{P(xi|vj)} = \frac{\boldsymbol{nc + mp}}{\boldsymbol{n + m}} \qquad (\mathbf{1.2})$$

Where:
n = the number of training examples for which $v = v_j$
n_c = number of examples for which $v = v_j$ and $x = x_i$
p = a priori estimate for P (x_i|vj)
m = the equivalent sample size

A Naïve-Bayes is a simple structure that has the class node as the parent node of all other nodes (Fig 1). No other connections are allowed in a Naïve-Bayes structure. Naïve-Bayes has been used as an effective classifier for many years. Unlike many other classifiers, it is easy to construct, as the structure is given *a priori* (and hence no *structure* learning procedure is required). Naïve-Bayes assumes that all the features are independent of each other. Although this independence assumption is obviously problematic, Naïve-Bayes has surprisingly outperformed many sophisticated classifiers over a large number of datasets, especially where the features are not strongly correlated. In recent years, a lot of efforts have focussed on improving Naïve-Bayesian classifiers, following two general approaches: selecting feature subset and relaxing independence assumptions.

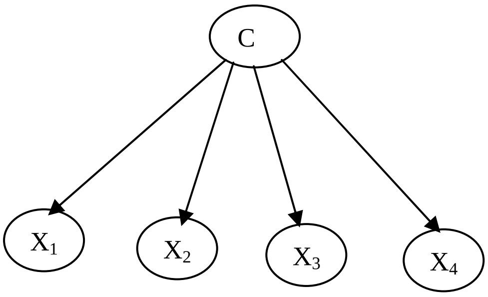

Figure 1: Simple Naive Bayes Classifier

1.2.1 Advantages of the Naive Bayes classifier

The main advantage of Bayesian classifiers is that they are probabilistic models, robust to noise found in real data. The Naive Bayes classifier presupposes independence of the attributes used in classification. However, it was tested on several artificial and real datasets, showing good performances even when strong attribute dependences are present. In addition, the Naive Bayes classifier can outperform other powerful classifiers when the sample size is small.

2. Materials And Methods

The purchase behaviour variables used in the model development are as follow.

Recency: This is the number of months since the last purchase and first purchase. It is typically the most powerful of the three characteristics for predicting response to a subsequent offer. This seems quite logical. It says that if you have recently purchased something from a company, you are more likely to make another purchase than someone who did not recently make a purchase.

Frequency: This is the number of purchases. It can be the total of purchases within a specific time frame or include all purchases. This characteristic is second to recency in predictive power for response. Again, it is quite intuitive as to why it relates to future purchases.

Monetary value: This is the total naira amount. Similar to frequency, it can be within a specific time frame or include all purchases. Of the three, this characteristic is the least powerful when it comes to predicting response. But when used in combination, it can add another dimension of understanding.

Demographic information includes customers' personal characteristics and information such as age, sex, address, profession etc

2.1 Data Warehouse Development

The data used for this research was gathered from customers' records at Ebedi Microfinance Limited. Management staff of the bank were asked questions concerning their products, services and operations which enhanced the understanding of the data gathered. Oracle database version 10.2.0.3 was used in building the data warehouse.The following procedures were adopted in the design of the data warehouse:

2.1.1Warehouse Builder Design Centre Preparation

In this phase, the project is specified. Connection to target data object was established by specifying the locations. In this case, a single data source was used and therefore a single oracle module was used [1][10[12][16]. The logical design of the warehouse was carried out to support two physical environments: testing and production environments.

2.1.2 Importing the Source Metadata

The metadata was imported from a single location. Customers' purchases and demographic data were imported from a text file.

2.1.3 Profiling Data and Ensuring Data Quality

Quality data is crucial to decision-making and planning. The aim of building a data warehouse is to have an integrated, single source of data that can be used to make business decisions. Since the data is usually sourced from a number of disparate systems, it is important to ensure that the data is standardised and cleansed before loading into the data warehouse. Data profiling is the process of uncovering data anomalies, inconsistencies and redundancies by analysing the content, structure, and relationships within the data. The analysis and data discovery techniques form the basis for data monitoring.

2.1.4 Designing the Target Schema

Relational target schema was designed and used as the target schema. A relational target schema is one that contains relational data objects such as tables, views, materialized views, and sequences. All the warehouse data are stored in these objects. Relational objects which include tables, views, materialized views, and sequences were created. Additional structures pertaining to relational objects such as constraints, indexes and partitions were also created. The objects were configured by setting the physical properties. Finally the codes that will create the data objects were generated.

2.1.5 Designing the ETL Logic

ETL (Extraction, Transformation and Loading) logic deals with designing mappings that define the flow of data from the source to target objects. After mapping design, a process flow is defined. A process flow allows activities to be linked together and describes constraints between the activities.

2.1.6 Deploying the Design and Executing the Data Integration Solution

Deployment is the process of copying the relevant metadata and generated code in the design centre to the target schema. This step is necessary because it enables the target schema to execute ETL logic such as mappings. Execution is the process of executing the ETL logic defined in the deployed objects.

2.1.7 Monitoring and Reporting on the Data Warehouse

It is essential to ensure the quality of data entering your data warehouse over time. Data auditors enable the monitoring of the quality of incoming data by validating incoming data against a set of data rules and determining if the data conforms to the business rules defined in the data warehouse.

2.2 Data Mining and Response Model Development

The dataset derived from the data warehouse for the mining activity had 14 attributes including purchases history data and customers' demographic information. The attributes, their data types and description are presented in Table 1.

Table 1: Attributes of Customers Purchases History and Demographic dataset.

Attribute	Data Type	Description
ACC_NO	Numerical	A unique number identifying each customer
SURNAME	Categorical	A customer's family name
FIRST_NAME	Categorical	A customer's given name
CONTACT_MODE	Categorical	Media through which a customer is contacted
PROFESSION	Categorical	Customer's occupation
SEX	Categorical	Customer's gender
DATEOFBIR	Categorical	Customer's date of birth
MARISTAT	Categorical	Customer's marital information
ADDRESS	Categorical	Customer's contact address
ACCOUNT TYPE	Categorical	Account type operated by a customer
OPEN DATE	Categorical	Date on which an account was opened
AMOUNT(N)	Numerical	Opening amount for an account
DATLASTTR	Categorical	Date on last transaction was carried out
TNT	Numerical	Total number of transactions by a customer

Methods for analysing and modelling data can be divided into two groups: supervised learning and unsupervised learning. Supervised learning requires input data that has both predictor (independent) variables and a target (dependent) variable whose value is to be estimated. Through various means, the process learns how to model (predict) the value of the target variable based on predictor variables. Decision trees, neural networks, support vector machines are examples of supervised learning. Supervised learning is best suited for analysis dealing with the prediction of some variable.

Unsupervised learning on the other hand, instead of identifying a target (dependent) variable treats all of the variables equally. In this case, the goal is not to predict the value of a variable but rather to look for patterns, groupings or other ways to characterise the data which may lead to understanding of the way data interrelates. Cluster analysis, correlation and statistical measures are examples of unsupervised learning.

The purpose of predictive mining is to find useful patterns in the data in order to make nontrivial predictions on new data. Two major categories of predictive mining techniques are those which express the mined results as a black box whose innards are effectively incomprehensible to non-experts and those which represent the mined results as a transparent box whose construction reveals the structure of the pattern.

Neural networks are major techniques in the former category. The latter includes methods for constructing decision trees, classification rules, bayesian network, association rules, clusters and instance-based learning [5][7][9].

In this work, the Naive Bayesian algorithm (a supervised learning algorithm) was used to model the customers' purchases dataset for target selection whose construction are easily comprehensible to experts and non-experts alike. Figure 2 presents a pictorial representation of the classifier system phases. Sixty percent (60%) of the dataset were used to build the response model and the model was applied to the remaining forty percent (40%) of the dataset for testing.

2.2.1 Feature Selection

Feature selection is a critical step in response modelling. No matter how powerful a model is, presence of irrelevant input variables lead to poor accuracy. Because customer related dataset usually contains hundreds of features or variables, many of which are irrelevant and heavily correlated with others, without feature selection, they tend to deteriorate performance of the model, as well as increase the model training time. Feature subset selection can be formulated as an optimization problem which involves searching the space of possible features to identify a subset that is optimum or near optimal with respect to performance measures such as accuracy. Feature selection can then either be performed as a preprocessing step, independent of the induction algorithm or explicitly make use of it. The former approach is termed filter, the latter wrapper. Filter methods operate independently of the target learning algorithm. Undesirable inputs are filtered out of the data before induction commences. Wrapper methods make use of the actual target learning algorithm to evaluate the usefulness of inputs. Typically, the input evaluation heuristic that is used is based upon inspection of the trained parameters and/or comparison of predictive performance under different input subset configurations. Input selection is then often performed in a sequential fashion. The backward selection scheme starts from a full input set and step-wisely prunes input variables that are undesirable. The forward selection scheme starts from the empty input set and step-wisely adds input variables that are desirable. Feature wrappers often achieve better results than filters due to the fact that they are tuned to the specific interaction between an induction algorithm and its training data, but it also tends to be more computationally expensive than the filter model.

When the number of features becomes very large, the filter model is usually chosen due to its computational efficiency. In this study, both filter and wrapper methods have been employed. Attributes such as surname, first name and mode of contact have been removed from the input as part of pre-processing due to their irrelevance in the prediction activity. However, attributes such as sex, date of birth, profession, date of last transaction, amount involved in the account opening etc have been used as inputs and left for the induction algorithm to determine their importance in the prediction task with the prediction value as the target.

3. Results And Discussion

3.1 Pattern Recognition

This is the stage where patterns representing knowledge in the mined results are identified and evaluated based on some evaluation metrics.

Evaluation Metrics

In selecting the appropriate algorithm and features that best model customers' response, the following evaluation metrics were employed:

i. **Percentage of Correct/Incorrect Classification:** This is the difference between the actual and predicted values of variables.
ii. **True Negative (TN):** Number of correct predictions that an instance is false.
iii. **False Positive (FP):** Number of incorrect predictions that an instance is true
iv. **False Negative (FN):** Number of incorrect predictions that an instance is false.
v. **True Positive (TP):** Number of correct predictions that an instance is true.
vi. **Accuracy (Acc):** Proportion of total number of predictions that were correct.

$$\boldsymbol{Accuracy = \frac{TP + TN}{TP + FP + TN + FN}} \qquad (\mathbf{2.1})$$

vii. **Error Rate (E):** Proportion of total number of predictions that were incorrect

$$\boldsymbol{Error\ Rate = 1 - Acc} \qquad (\mathbf{3.2})$$

viii. **Confusion Matrix:** The most basic performance measures can be described by a confusion matrix. The confusion matrix (Kohavi, 1988) contains information about actual and predicted classifications done by a classification system. The columns represent the predicted classifications, and the rows represent the actual (true) classifications for each of all records.

ix. **Receiver Operating Characteristics (ROC) Analysis:** The Receiver Operating Characteristic (ROC) analysis comes from statistical decision theory (Green and Swets, 1966) and was originally used during World War II for the analysis of radar images. It is a graphical plot of the sensitivity of True Positive Rates (TPR) to False Positive Rate (TFR) varying the discrimination threshold.

3.2 Performance Metrics

The accuracy (Acc) and Error rate (E) are widely used metrics for measuring the performance of learning systems. However, when the prior probabilities of the classes are very different, such metric might be misleading. Accuracy alone is not an adequate measure of performance especially where the number of negative cases is much greater than the number of positive cases.

Lift/Gain Chart: In such cases where the number of negative cases is much greater than the number of positive cases, a *lift chart* can also be used. The lift chart (also called Gain chart) gives a graphical representation of these parameters for various thresholds on the output and encapsulates all the information contained in the confusion matrix.

Table 2 presents some samples of the Naïve Bayes Prediction results obtained.

Table 2: Sample Naïve Bayes Prediction Results

Acc_No	Prediction	Remark	Probability	Cost	Rank
700002	1	Respondent	1	1.70E-08	1
200003	1	Respondent	0.9998502	3.31E-04	1
300004	1	Respondent	0.9999341	1.46E-04	1
200005	0	Non Respondent	0.72576463	0.50051033	1
500006	1	Respondent	1	6.89E-16	1
100007	0	Non Respondent	0.8347949	0.30151784	1
100008	0	Non Respondent	0.8267257	0.316245	1
100009	1	Respondent	1	1.17E-12	1
100010	1	Respondent	0.9999999	2.74E-07	1
200011	1	Respondent	1	1.29E-16	1
300012	1	Respondent	1	4.84E-17	1
200013	1	Respondent	1	1.57E-12	1
200014	1	Respondent	0.99943364	0.00125279	1
100015	0	Non Respondent	1	4.89E-09	1
100016	0	Non Respondent	0.48024932	0.9486032	1
100017	1	Respondent	1	1.20E-17	1
600018	1	Respondent	0.99988407	2.56E-04	1
100019	1	Respondent	0.99930745	0.00153182	1
100020	1	Respondent	0.99996334	8.11E-05	1
500021	1	Respondent	1	1.02E-16	1
100022	1	Respondent	0.9998686	2.91E-04	1

3.3 Summary of Prediction Report

Table 3 gives a summary of the response prediction results obtained by the Naïve Bayes classifier when applied to the customers' dataset.

Table 3: Summary of Response Prediction Report

Response	Number of Customers
Respondents	221
Non Respondents	179
Total	400

Percentage Positive Response = 221/400*100
= 55.25%

Percentage Negative Response = 179/400*100
= 44.75%

From the summary of results, the level of likely favorable response of customers is slightly above average because the percentage of likely positive response is slightly above 50% (55.25%) while that of likely unfavorable response is slightly below average because the percentage of negative response is slightly below 50% (44.75%). The implication of this is that Ebedi Microfinance bank will benefit from sending promotion offers to the potential respondents and save cost as a result of not sending promotion offers to the non respondents. This will consequently increase the bank's Return on Investment.

Naïve Bayes Response Model Evaluation and Performance

The summary results of response model evaluation are presented in Table 4. Table 5 presents the statistics of the model performance the confusion matrix is presented in Table 6

Table 4: Results of Response Model Evaluation

Metrics	Value
True Positive Rate	0.7654320987
False Positive Rate	0.4594594594
Average Accuracy	0.6529863197
Overall Accuracy	0.6580645161
Error Rate(1 -Accuracy)	0.3419354839
Cost	53
Probability Threshold	0.3787237704

Table 5: Model Performance Statistics

Target	Total Actual	Correctly Predicted (%)	Cost	Cost (%)
0	74	59.46	66.36	58.32
1	81	67.9	47.45	41.69

Table 6: Confusion Matrix

	0	1	Total	Correct %	Cost
0	44(TN)	30(FP)	74	59.46	66.36
1	26(FN)	55(TP)	81	67.9	47.45
Total	70	85	155		
Correct	62.86	64.71			
Cost	47.45	66.36			

From the confusion matrix the difference between the correctness produced by the actual classification (represented by the row interpretation) and the predicted classification (represented by the column interpretation) for the positive cases (in which we are interested) shows a higher correctness (67.9%) for the actual value which suggests the accuracy of the model.

From the confusion matrix the difference between the correctness produced by the actual classification (represented by the row interpretation) and the predicted classification (represented by the column interpretation) for the positive cases (in which we are interested) shows a higher correctness (67.9%) for the actual value which suggests the accuracy of the model.

3.3.1 Receiver Operating Characteristics (ROC) Curve

Figure 2 presents the plot of True Positive Rates (TPR) against the False Positive Rates (FPR) with the value of the TPR being 0.7654320987 (the intersection of the ROC curve and the probability threshold) and the value of FPR as 0.4594594594 (the intersection of the diagonal and the probability threshold). From the ROC result presented in Table 7, it can be observed that the model produces higher True Positive (TP) and False positive (FP) cases as the probability threshold increases. The True Positive Rate (TPR) and the False Positive Rate (FPR) produces 0.7654320987 and 0.4594594594 respectively. This implies the model produces a higher number of correctly identified respondents as against a lower number of non respondents incorrectly identified as respondents. This attests to the model's accuracy.

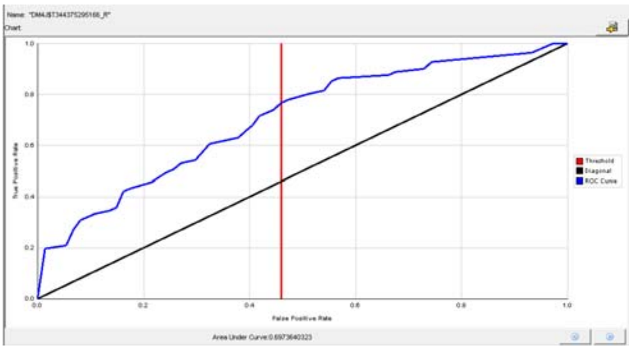

Figure 2: Receiver Operating Characteristics (ROC) Curve

3.3.2 Lift Charts

The cumulative Lift chart is presented in Figure 3. The cumulative quantile for first quantile is 1.7939814815. The cumulative positive for first quantile is 0.1851851852. The cumulative lift for the other quantiles up to the tenth quantile together with cumulative positive, threshold, target density count, percentage record cumulative, target cumulative and non target cumulative is presented in Table 8

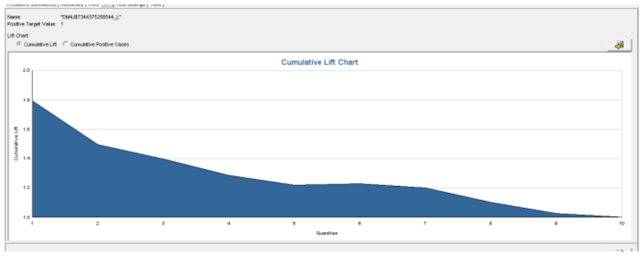

Figure 3: Cumulative Lift chart

Looking at the steeply increasing cumulative lift chart, the model was able to achieve a high degree of positive cases (1.8 out of 2.0 for 10 quantiles), that is, high number of correctly identified respondents.

The Lift result presented in Table 8 shows that the model achieves a good number of targets (respondents) out of the total count (total number of customers addressed) on the first six quantiles, sending promotion offers to more customers after the sixth quantile will not justify the cost of the promotion campaign because only a few targets are gotten from the total count.

Table 7: Receiver Operating Characteristics (ROC) Result

Index	Probability Threshold	False Positive	False Negative	True Positive	True Negative	Accuracy	Average Accuracy	Cost
0	1	0	81	0	74	0.477419355	0.5	81
1	1	1	65	16	73	0.574193548	0.592008675	66
2	0.99999994	4	64	17	70	0.561290323	0.577911245	68
3	0.999988914	5	59	22	69	0.587096774	0.602018685	64
4	0.999983847	6	56	25	68	0.6	0.613780447	62
5	0.999983072	7	55	26	67	0.6	0.61319653	62
6	0.9999578	8	54	27	66	0.6	0.612612613	62
7	0.999951899	10	53	28	64	0.593548387	0.605271939	63
8	0.999943256	11	52	29	63	0.593548387	0.604688021	63
9	0.999915421	12	47	34	62	0.619354839	0.628795462	59
10	0.999897599	13	46	35	61	0.619354839	0.628211545	59
11	0.999818563	16	44	37	58	0.612903226	0.620286954	60
12	0.999719083	18	41	40	56	0.619354839	0.625291959	59
13	0.999652863	19	40	41	55	0.619354839	0.624708041	59
14	0.9995296	20	38	43	54	0.625806452	0.630296964	58
15	0.999390244	22	37	44	52	0.619354839	0.62295629	59
16	0.771490574	24	32	49	50	0.638709677	0.640306974	56
17	0.680298388	28	30	51	46	0.625806452	0.625625626	58
18	0.581525683	29	28	53	45	0.632258065	0.631214548	57
19	0.549383104	30	26	55	44	0.638709677	0.63680347	56
20	0.460064292	31	23	58	43	0.651612903	0.648565232	54
21	0.408490717	33	21	60	41	0.651612903	0.647397397	54
22	0.37872377	34	19	62	40	0.658064516	0.65298632	53
23	0.356039256	35	18	63	39	0.658064516	0.652402402	53
24	0.25514707	38	16	65	36	0.651612903	0.644477811	54
25	0.203911513	40	15	66	34	0.64516129	0.637137137	55
26	0.154189616	41	12	69	33	0.658064516	0.648898899	53
27	0.001631624	42	11	70	32	0.658064516	0.648314982	53
28	2.07E-04	49	10	71	25	0.619354839	0.607190524	59
29	1.90E-04	50	9	72	24	0.619354839	0.606606607	59
30	8.22E-05	54	8	73	20	0.6	0.585752419	62
31	6.90E-05	55	6	75	19	0.606451613	0.591341341	61
32	1.74E-08	69	3	78	5	0.535483871	0.515265265	72
33	1.52E-08	70	2	79	4	0.535483871	0.514681348	72
34	4.10E-12	72	0	81	2	0.535483871	0.513513514	72
35	0	74	0	81	0	0.522580645	0.5	74

Table 8: Lift/Gain Result

Quantile Number	Cumulative Lift	Cumulative Positive	Threshold	Quantile Total Count	Quantile Target Count	Percentage Records Cumulative	Target Density Cumulative	Target Cumulative	Non Target Cumulative	Lift Quantile	Target Density
1	1.793981481	0.185185185	5.39E-08	16	15	0.10322581	0.9375	15	1	1.793981481	0.9375
2	1.554783951	0.320987654	5.22E-05	16	11	0.20645161	0.8125	26	6	1.31558642	0.6875
3	1.355452675	0.419753086	2.74E-04	16	8	0.30967742	0.7083333	34	14	0.956790123	0.5
4	1.255787037	0.518518519	0.0010404	16	8	0.41290323	0.65625	42	22	0.956790123	0.5
5	1.195987654	0.617283951	0.9326942	16	8	0.51612903	0.625	50	30	0.956790123	0.5
6	1.248862898	0.765432099	1.4128855	15	12	0.61290323	0.6526316	62	33	1.530864198	0.8
7	1.2003367	0.851851852	1.8286813	15	7	0.70967742	0.6272727	69	41	0.893004115	0.46666667
8	1.117530864	0.901234568	2.2117205	15	4	0.80645161	0.584	73	52	0.510288066	0.26666667
9	1.038800705	0.938271605	2.2119162	15	3	0.90322581	0.5428571	76	64	0.382716049	0.2
10	1	1	2.2119565	15	5	1	0.5225806	81	74	0.637860082	0.33333333

5. Conclusion

From the summary of the prediction results obtained, which classifies 55.25% of the customers as respondents and 44.75% as non respondents, we can conclude that Ebedi Microfinance bank can plan effective marketing of its products/services through the guiding report obtained on each of the customers. This will enable the management increase its sales by targeting the respondents and prevent wasteful expenditure that will have been incurred as a result of sending promotion offers to the non respondents. These will go a long way in increasing the bank's Return on Investment.

References

[1] Adamson C. and M. Venerable. "*Data Warehouse Design Solutions*". J. Wiley & Sons, Inc., 1998

[2] Alejandro Gutiérrez and Adriana Marotta. "An Overview of Data Warehouse Design Approaches and Techniques", 2000

[3] Almquist, E. and Wyner G. "Boost Your Marketing ROI with Experimental Design". Harvard Business Review, p.135-141., 2001

[4] Ballard. C. "*Data Modelling Techniques for Data Warehousing*". SG24-2238-00. IBM Red Book. ISBN number 0738402451., 1998

[5] Berry, M. J. A. and Linoff, G. S. "*Data Mining Technique for Marketing, Sale,* and Customer Relationship Management" (2nd edn), Indiana, Indianapolis Publishing Inc., 2004

[6] Changyun Wang. "Bayesian Belief Network Simulation", Department of Computer Science Florida State University, 2003

[7] Chen, H.Y., Chen, T.C., Min, D., Fischer, G. and Wu, Y.M.'Prediction of tacrolimus blood levels by using the neural network with genetic algorithm in liver transplantation patients', *Therapeutic Drug Monitoring*, 21, 1, 50– 56., 1999

[8] Cheung, K.-W., Kwok, J. K., Law, M. H. and Tsui, K.-C. 'Mining customer product rating for personalized marketing ', Decision Support Systems, 35, 231–243., 2003

[9] Chiu, C. "A case-based customer classification approach for direct marketing", Expert Systems with Applications, 22, 163–168., 2002

[10] Colin McGregor. Oracle Database 2 Day DBA, 10*g* Release 2 (10.2), 2005

[11] Date, C. J. *An introduction to database systems* (7th ed.). Reading, Mass.: Addison-Wesley, 2000

[12] Gang Luo. Techniques for Operational Data Warehousing. PhD Thesis (Computer Sciences), University of Wiscosin-Madison, 2004

[13] Han, J., & Kamber, M. *Data mining: Concepts and Techniques*. San Francisco: Morgan Kaufmann Publishers., 2001

[14] Ho, T.B. (nd). Knowledge Discovery and Data Mining Techniques and Practice Available on URL:http://www.netnam.vn/unescocourse/knowlegde/knowfrm.htm, 2006

[15] Inmon. W.H. Building the Data Warehouse. Wiley Computer Publishing, 1996

[16] Scalzo, B. (2003). *Oracle DBA guide to data warehousing and star schemas*. Upper Saddle River, N.J.: Prentice Hall PTR.

[17] Shapiro, G. P. 'Knowledge Discovery in Databases: 10 years after', *SIGKDD* Explor. Newsl, 1, 2 (Jan. 2000), 59- 61.

Available from: http://doi.acm.org/10.1145/846183.846197, 2000
[18] Shin, H. J. and Cho, S. 'Response modelling with support vector machines', Expert Systems with Applications, 30, 4, 746–760., 2006
[19] Ou, C., Liu, C., Huang, J. and Zhong, N. 'One Data mining for direct marketing', *Springer-Verlag Berlin Heidelberg*, pp. 491–498., 2003
[20] Paul Lane, Viv Schupmann and Ingrid Stuart. Oracle Database Data Warehousing Guide 11g Release 1(11.1), 2007
[21] Petrison, L. A., Blattberg, R. C. and Wang, P. 'Database marketing: Past present, and future', *Journal of Direct Marketing*, 11, 4, 109–125, 1997
[22] Usama M. Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth. From Data Mining to knowledge Discovery: An Overview, Advances in Knowledge Discovery and Data Mining, (1996). AAAI Press, , pp 1-34.

Eniafe Festus Ayetiran is a Lecturer at the Achievers University, Owo, Nigeria. He holds a Master of Science (M.Sc) degree of the University of Ibadan, Nigeria and a Bachelor of Science (B.sc Hons) degree of the Adekunle Ajasin University, Akungba Akoko, Nigeria. His research interests include Data Mining, Data Warehousing, Business Modeling, Software Engineering and Computer Security.

Dr. Adesesan Barnabas Adeyemo is senior lecturer at the Computer Science Department of the University of Ibadan, Nigeria. He obtained his PhD and M.Sc degrees at the Federal University of Technology, Akure. His research activities are in Data Mining, Data Warehousing & Computer Networking. He is a member of the Nigerian Computer Society and the Computer Professionals Registration Council of Nigeria. Dr. Adeyemo is a Computer Systems and Network Administration specialist with expertise in Data Analysis and Data Management.

