Classification-Based Data Mining in Target Marketing

N. Mlambo
Lecturer, School of Information and Communication Technology, College of Science and Technology,
University of Rwanda, Rwanda

Abstract: Target marketing is the selection of potential customers to whom a business wishes to sellproducts or services. The targeting strategy involves segmenting the market, choosing which segments of the market are appropriate, and determining the products that will be offered in each segment. A business offering multiple products can determine if the various segments should receive one generic product (such as in mass marketing), or if each segment should receive a customized product (multi-segment), based upon the market'sdiversity, maturity, the level of competition and the volume of sales expected. In this paper we look at the applications of classification-based data mining techniques in target marketing and some specific models that were developed to assist in the same.

Keywords: Data mining, target marketing, classification, direct marketing

I. INTRODUCTION

Data mining refers to the extraction of hidden predictive information from large data sources, mainly data warehouses. The target of data mining [1, 3] is to find unexpected characteristics, hidden features or other unclear relationships inside the data based on techniques combinations.There are two major tasks of data mining namely; predictive and descriptive. Predictive mining tasks perform inference on the current data in order to make predictions. Descriptive mining tasks characterize the general properties of the data in the data source. Several major Data Mining techniques have been developed and used in data mining projects recently including clustering, association, prediction, sequential patterns, classification, etc [2]

II. CLASSIFICATION

Classification is a very popular data mining technique, which employs a set of pre-classified examples to develop a model that can classify the population of records. It can be used to predict categorical class labels and classifies data on training set and class labels and hence can be used for classifying newly available data [2, 5, 12]. There are many classification techniques, which include; neural networks, decision trees, Bayesian networks, k-nearest neighbor, and support vector machines. Fraud detection and credit –risk applications are particularly well suited to this type of analysis.

III. APPLICATION OF CLASSIFICATION IN DIRECT MARKETING

Almost all industries that sell products and servicesneed to advertise and promote their products and services.There are generally two approachesto advertisement and promotion: mass marketing anddirect marketing [9]. Mass marketing, which uses massmedia such as television, radio, and newspapers, broadcastsmessages to the public without discrimination. Itused to be an effective way of promotion when the productswere in great demand by the public.

Direct marketing is a process of identifying likely buyersof certain products and promoting the products accordingly. Instead of promoting to customers indiscriminatively, this marketing technique studies customers' characteristics and needs, and selects certain customers as the target for promotion [8, 9]. The hope is that the responserate for the selected customers can be much improved. Direct marketing has generated an increasing interest among academics and practitioners over the past few years mainly due to competitive market environment, advancement in technology and changing behavior of customers which are difficult to predict [8]. It is increasingly used by banks, insurance companies, and the retail industry.

3.1 Life Insurance Companies

Modeling underwriting decisions helps in selling life insurance by identifying potential customers who are more likely to qualify for life insurance products. Marketing expenses are significant portions of life insurance company budgets, and utilizing them efficiently is a key operational strategy.

3.2 Retail Industry

Classification can help direct marketers by providing useful and accurate trends on purchasing behavior of their customers and also help them in predicting which products their customers may be interested in buying. In fact data mining allows retail store managers to identify their best customers, attract customers and inform customers via mail marketing, and maximize profitability by means of identifying profitable customers.

3.3 Financial, banking and credit scoring

Classification can assist financial institutions in various ways, such as credit reporting, credit rating, loan or credit card approval by predicting good customers, mode of service delivery and customer retention (i.e build profiles of customers likely to use which services. A credit card company can leverage its vast warehouse of customer transaction data to identify customers most likely to be interested in a new credit product and target the same.

IV.   EXAMPLES OF CLASSIFICATION-BASED MODELS DEVELOPED TO ASSIST DIRECT MARKETING

4.1 Bank Direct Marketing Based on Neural Network [7]

Data mining has gained popularity for illustrative and predictive applications in banking processes. In this application the decision tree model (C5.0), and the Multilayer perceptron neural network (MLPNN) with back propagation are applied on the bank direct marketing. The C5.0 classifier is selected because it gives more accurate and efficient output with comparatively high speed. Memory usage to store the rule set in case of the C5.0 classifier is less as it generates smaller decision trees [10]. The dataset well known as bank marketing from the University of California at Irvine (UCI) [11] is used.

The objective is to examine the performance of MLPNN and C5.0 models on a real-world data of bank deposit subscription. The purpose is increasing the campaign effectiveness by identifying the main characteristics that affect a success (the deposit subscribed by the client) based on MLPNN and C5.0. The experimental results demonstrate, with higher accuracies, the success of these models in predicting the best campaign contact with the clients for subscribing deposit. The performances are measured by three statistical measures; classification accuracy, sensitivity, and specificity.

4.2 A Data Mining-Based Response Model for Target Selection in Direct Marketing [12] (Naïve Bayes Algorithm)

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 (Nigeria) 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.

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 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

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.

V.   CONCLUSION

A target marketing model helps to establish aguideline of marketing strategies. Marketing strategy for a bank should target those valuable consumers andmaintain a good relationship in order to generate greater revenue. The bank can position the target consumersprecisely and implement various marketing strategies to satisfy consumers' needs. The same should be done in companies that are in retail and life insurance industries, etc. Classification-based techniques come in handy and this explains why data mining is becoming very popular.

REFERENCES
[1]   Nikita Jain, Vishal Srivastava; "Data Mining Techniques: A Survey Paper", International Journal of Research in Engineering and Technology eISSN: 2319-1163 | pISSN: 2321-7308
[2]   BarahateSachin R. and Shelake Vijay M, "A Survey and Future Vision of Data mining in Educational Field", Second International Conference on Advanced Computing & Communication Technologies, pp. 96-100, 2012.
[3]   Bharati M. Ramageri; "ROLE OF DATA MINING IN RETAIL SECTOR", International Journal on Computer Science and Engineering (IJCSE)ISSN : 0975-3397 Vol. 5 No. 01 Jan 2013.
[4]   www.businessdictionary.com accessed on 23/11/2014
[5]   Aparna Raj, BincyG.and T. Mathu; "Survey on Common Data Mining Classification Techniques", International Journal of Wisdom Based Computing, Vol. 2(1), April 2012.

[7]   Hany. A. Elsalamony, Alaa. M. Elsayad; "Bank Direct Marketing Based on Neural Network",
        International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-2, Issue-6, August 2013.

[8]   Lilian Sing'oei and Jiayang Wang; "Data Mining Framework for Direct Marketing:A Case Study of Bank Marketing", IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 2, No 2, March 2013ISSN (Print): 1694-0814 | ISSN (Online): 1694-0784

[9]   Charles X. Ling and Chenghui Li; "Data Mining for Direct Marketing: Problems and Solutions",
        Department of Computer Science, The University of Western Ontario, Ontario, Canada.

[10]   PreetiPatidar and Anshu Tiwari; "Handling Missing Value in Decision Tree Algorithm",International Journal of Computer Applications (0975 – 8887) Volume 70– No.13, May 2013

[11]   http://en.wikipedia.org/wiki/Neural_network#History of the neural network analogy. Nov 2010.

[12]   Eniafe Festus Ayetiran; "A Data Mining-Based Response Model for Target Selection in Direct Marketing", I.J. Information Technology and Computer Science, 2012, 1, 9-18