A Direct Marketing Framework to Facilitate Data Mining Usage for Marketers: A Case Study in Supermarket Promotions Strategy

Adel Flici (0630951)
Business School, Brunel University, London, UK

Abstract

Data mining technology can detect and predict changes in customer behaviour which facilitates the effective planning of direct marketing campaigns. Most of research on data mining for direct marketing focuses on the theoretical and computational aspects of the technology. However, the process of choosing the mining objectives and methods for data mining on a specific context is still unstructured and based mostly on user judgments. In this paper, we design a Direct Marketing Framework (DMF) using a combination of different data mining methodologies and direct marketing concepts. It is aimed to facilitate marketers with only little data mining skills to effectively plan direct marketing campaigns. A simplified case study on supermarket promotions strategy was conducted to verify the feasibility of the framework. The valuation of its performance and utility is very positive.

Keywords: Direct marketing, data mining, DMF.
1. Introduction

The dynamic nature of the marketing environment makes customers’ behaviour increasingly unsteady and unpredictable which in turn make marketing campaigns difficult to plan. In such environment, data mining techniques can detect and predict changes in customer behaviour. This facilitates the process of building long-term and strong relationship between businesses and their customers’. In addition, understanding customers’ behaviour can support managers to plan more effective direct marketing campaigns. That is to say, data mining is very important technology for efficient direct marketing (Chen, Chiu & Chang 2005); (Wang, Hong 2006).

The application of data mining models to direct marketing has been investigated by many researchers in the field. However, practitioners seem to rarely use these research findings and apply them to real-world cases. This led to ask a key question; is there a gap between what marketing managers need and what data mining modellers offer? It is not a question of the relevance of research and their theoretic aspects but rather their managerial application. Indeed, the lack of marketing managers involvement to develop and use these models to support their decisions show the urgent need for systematic, simple and well-integrated framework to help them get more involved and capable to use data mining with little help from data analysis experts (Martínez-López, Francisco J. 2008). Moreover, current data mining methodologies only offer general guidelines to carry out a data mining project and also involves high level of user judgements (Shaw et al. 2001). In this paper, we design a Direct Marketing Framework (DFM) using a combination of different data mining methodologies and direct marketing activities. It provides a structure for an effective management of small to large data mining projects in direct marketing context. It is aimed to ease marketers with only little data mining skills to use the technology. This is achieved by facilitating the overall process of marketing knowledge extraction. One of the key objectives of DFM is to reduce user judgments by leading marketers through a clear path and pre-determined phases. It is also imperative that DFM reduces marketers’ reliance on data analysts tackling issues such as time and trust. This emphasises on the urgency to provide marketers with a systematic, simple, and well-integrated framework, in order to facilitate and encourage the use and trust of data mining technology.
The reminder of the paper is organised as follows. Section 2 provides a brief overview of data mining and direct marketing. In section 3, we discuss the motivations of this research. Section four presents DMF and briefly discusses what is required in each phase of the process. Section 5 presents the supermarket promotions case study that illustrates the utility of the framework. Finally, we conclude the paper and give a brief summary of future work to be carried out.

2. Background

Data mining is one of many technologies that use advanced analytics to critically analyse vast amount of data stored in data warehouses or other information repositories, in order to discover hidden or interesting patterns, associations and anomalies which are then represented in models. In this context, interesting pattern is a combination of validity, novelty, usefulness, and understandability. The use of data mining resulting models can lead companies to new insights, and in a business context, to competitive advantage. These models also play the role of inferred knowledge. However, deciding on the usefulness of knowledge is part of a wider process for which subjective user judgements is usually needed (Fayyad, Piatetsky-shapiro & Smyth 1996); (Mitra, Pal & Mitra 2002); (Witten, Frank 2005). This is discussed in more details in section 4.

Direct marketing is a business model that uses data mining techniques and marketing databases for personalisation and business intelligence. It is a new approach that uses interactive one-to-one communication between marketer and customer. It is aimed at specific customers with personalised advertising and promotional campaigns. It involves investigating the customers’ characteristics and needs to identify customer market value and select customers most likely to respond to promotions. This method is progressively establishing itself as the preferred option for companies because of its low cost and high profitability. Recently, expenditure on direct marketing has increased dramatically i.e. 73.6% in the UK between 2001 and 2004. Furthermore, US expenditure was as high as $161.3 billion in 2005 accounting for 10.3% of its total GDP (Apte et al. 2002); (Bose, Chen 2009); (Ou et al. 2003). That is to say, direct marketing is rapidly growing in importance.
3. Research Motivations

Most of research on data mining for direct marketing focuses on the theoretical and computational aspects of the technology. However, the process of choosing the mining objectives and methods for data mining on a specific context is still unstructured and based mostly on judgment. There are several research challenges surrounding the process of discovering useful knowledge for direct marketing. This process need to be more structured in order to improve the productivity of data mining projects in direct marketing (Shaw et al. 2001). One of direct marketing key challenges is to research on which data mining methods should be used under what circumstances. This implies answering key questions before data analysis can proceed including; which customers should be selected as target for direct marketing? What data mining method should be used for targets selection? (Bose, Chen 2009). Moreover, the complexity of data mining model results like too many if-then statements make it difficult for marketers to understand them. Therefore, marketing managers are more reluctant to utilise the results due to difficulty, poor comprehensibility, and trust issues (Kim 2006). This emphasises on the need to provide direct marketers with a systematic, simple, and well-integrated framework, in order to facilitate and encourage the use and trust of data mining by marketers. It is also important to mention that factors such as market environment make direct marketing data mining project unique in a certain sense. However, there are a large set of common issues that rise in most of these projects. The developed framework should be usable by marketers with only little data mining skills (Gersten, Wirth & Arndt 2000).

4. Direct Marketing Framework (DMF)

Early studies on data mining knowledge discovery process such as (Fayyad, Piatetsky-shapiro & Smyth 1996) provided a framework for Knowledge Discovery for Databases (KDD) which integrates data mining in a simple process to follow sequentially in order to generate knowledge. On the other hand, this process involved high level of user judgements which require expert knowledge on both data analysis and the field understudy such as marketing. Many other researchers followed Fayyad steps and developed Knowledge discovery processes which eventually lead to the development of industry standard for data mining knowledge discovery projects i.e.
CRISP-DM. After extensively reviewing different academic and industry research studies on data mining process methodologies and direct marketing activities, a theoretical framework of data mining for direct marketing was designed. It is a systematic, simple and well-integrated framework that uses direct marketing concepts and data mining methods to ease the creation of marketing knowledge and support marketing decisions. A systematic approach for the data model is essential to achieve marketing objectives by leading marketers to migrate along pre-determined and desirable tracks. The main objective of the framework is to facilitate for marketers the process of finding interesting patterns. Figure 1 presents data mining framework for direct marketing which illustrates the phases of a data mining project in direct marketing context including; 1) Direct Marketing Planning, 2) Data Preparation, 3) Data mining technique selection, 4) Analysis, Evaluation & Interpretation of Model, and finally, 5) Presentation of Marketing Knowledge.

![Figure 1: Direct Marketing Framework (DFM).](image)

4.1. Direct Marketing Planning

The first step in the framework involves the marketer to identify the requirement of the project. In data mining projects, focus is really important factor because there are big amount of attributes and values to be considered and can lead to endless combinations (Mitra, Pal & Mitra 2002). Data mining application has three main purposes in direct marketing including customer profiling, deviation analysis and trends analysis (Shaw et al. 2001).

4.2. Data Preparation

The data preparation phase comprises data sampling, record and attribute selection, transformation and cleaning of data. Direct marketers have three commonly used categories of data in direct marketing including customer, transaction and product. Table 1 is a summary of data used in direct marketing models.
<table>
<thead>
<tr>
<th>Category</th>
<th>Data type</th>
<th>Importance</th>
<th>Accessibility</th>
<th>Variability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer</td>
<td>Demographic, lifestyle, socio-graphic</td>
<td>Low</td>
<td>External</td>
<td>High</td>
</tr>
<tr>
<td>Transaction</td>
<td>Transaction records, feedbacks, web browsing log files</td>
<td>High</td>
<td>Internal, accumulating</td>
<td>High</td>
</tr>
<tr>
<td>Product</td>
<td>Size, colour, price, design style</td>
<td>High</td>
<td>Internal</td>
<td>Low</td>
</tr>
</tbody>
</table>

Table 1: direct marketing data categories and types. Adapted (Bose, Chen 2009).

4.3. **Data Mining Technique Selection**

By this stage, direct marketers should have selected one of data mining main purposes in direct marketing i.e. customer profiling, deviation analysis and trends analysis. According to (Gersten, Wirth & Arndt 2000), mass modelling i.e. using many data mining techniques is a pragmatic method yielding good result in real-world marketing projects. It is important to mention that there are no universally best data mining techniques and selecting a specific or combination of method need subjective judgment on the suitability of an approach (Mitra, Pal & Mitra 2002). It is also important to mention that decision tree and association rules can be followed more easily by business users than neural networks. This is primarily due to the logic of both techniques and should be a good selection for non-experts in data mining (Ngai, Xiu & Chau 2009).

4.4. **Analysis, Evaluation and Interpretation of Model**

Marketers can use predictive accuracy for model evaluation such as lift chart. It is a graphical model where X-axis represents the percentage of data used for prediction and the Y-axis represent the percentage of prediction correctness. A typical model would include two lines; one line that will always be 100% correct and the other line represent the prediction correctness of the model (Harinath and Quinn, 2006).
4.5. Presentation of Marketing Knowledge

This step aims to ease the comprehensibility of the marketing model hence the extraction of knowledge.

Data visualisation provides marketers with various visual capabilities such as view complex patterns in three dimensions and colours. It also offers advanced manipulation functions to slice, rotate or zoom the model objects in order for marketers to have different level of details of patterns observed (Shaw et al. 2001).

5. A Case Study of Supermarket Promotions Strategy

After its theoretical presentation, we empirically experiment with the framework, using a consumer behaviour model of reference. In this part of the paper, we try to offer an overall perspective of how it works. The valuation of its performance and utility is very positive. This case study will present some of the conducted findings, but due to limitation of space only an overview of the case study will be presented.

The Meat Promotion dataset used in this case study is of a major supermarket chain in the UK. It contains five main attributes including Product Type Sales per unit, Promotion Type, Date, Lifestage, and Region. Given the information that is provided by the dataset, it can be interrogated for a variety of purposes. Consequently, the objective for the application of the framework is to facilitate marketers in identifying the opportunities that are available to a supermarket with regard to effective marketing. Thus, the direct marketing planning is achieved through the identification of relevant objectives. In this case, identify lowest sales regions for beef products?

For this purpose a sample of 4300 products was selected, with attributes of beef products and regions. After that, mass modelling have been performed using different data mining techniques to identify lowest sales regions of beef products. Decision tree has been selected as the most revealing and comprehensible model. The output model is presented through a decision tree along with accuracy lift chart in figure 2 and 3 respectively. Analysis of the decision tree led to the discovery that it is the South West region that is most poorly performing. Therefore, further analysis could be employed to increase sales in the South West. It is also worth mentioning that the decision evaluation model was the most accurate compare to other techniques applied for this sample.
The presentation of marketing knowledge phase depends on the data mining technique used. In this case, decision tree is easy to understand by marketing managers; hence further visual enhancement for more comprehensibility is not required.
6. Conclusion

Data mining is very powerful technology that can help organisations plan effective direct marketing campaigns. Most of research on data mining for direct marketing focuses on the theoretical and computational aspects of the technology. Yet these researches fail to achieve managerial application hence, creating a gap between researchers and practitioners. The key for effective direct marketing is to empower marketers with ability to independently create marketing knowledge.

In this paper, we introduced the theoretical design of a data mining framework for direct marketing i.e. DFM. It provides an overview of an ongoing research that allows marketers with only little data mining skills to build direct marketing models. The paper presented a case study in the context of supermarket promotions strategy. It demonstrates the use of DFM to create marketing knowledge. The valuation of its performance and utility is very positive. It also shows that the process of creating marketing knowledge for direct marketing can greatly be enhanced by a suitable framework.

This research has presented a key issue in the application of data mining in direct marketing context. It presented a theoretical framework and empirically tested its performance through the use of a case study in supermarket promotions strategy. Due to limitation of space only an overview of the case study has been illustrated, however the author can be contacted for detailed information regarding this research. Finally, it is intended to further verify the utility of DFM through the use of three more case studies all related to direct marketing context.
7. Reference


