ELEKTROTEHNIŠKI VESTNIK 80(3): 123-127, 2013
ORIGINAL SCIENTIFIC PAPER

Improvement of the Direct-Marketing Business Process by Using Data Mining

Rok Rupnik

University of Ljubljana, Faculty of Computer and information science, Tržaška 25, 1000 Ljubljana, Slovenia
E-mail: rok.rupnik@fri.uni-lj.si

Abstract. The direct-marketing business process applied in Slovenian publishing company was inefficient because of inadequate procedure used to create a list of potential customers for a marketing campaign. Considering the nature of the problem, data mining was selected to solve the problem. The company's direct marketing business process was renovated based on the CRISP-DM methodology and by using data mining for decision-making elements of the process. The paper represents a methodology enabling implementation of data mining into the business processes and its use in a direct marketing business process for a Slovenian publishing company.

Keywords: direct marketing, methodology, data mining

Izboljšanje poslovnega procesa direktnega marketinga z uporabo odkrivanja zakonitosti v podatkih

Direktni marketing v slovenskem založniškem podjetju je bil neučinkovit zaradi neučinkovitega postopka izdelave seznama potencialnih kupcev za marketinške akcije. Glede na naravo problema je bila za izboljšanje postopka izbrana tehnologija odkrivanja zakonitosti v podatkih. Poslovni proces direktnega marketinga je bil prenovljen na osnovi CRISP-DM metodologije in uporabe odkrivanja zakonitosti v podatkih za realizacijo elementov odločanja v poslovnem procesu direktnega marketinga. Članek predstavlja metodologijo uvajanja uporabe odkrivanja zakonitosti v podatkih v poslovne procese in uporabe te metodologije na primeru poslovnega procesa direktnega marketinga v slovenskem založniškem podjetju.

1 INTRODUCTION

Data mining is used in information systems as a technology to support decision making on the tactical level as well as to enable decision activities within operational business processes. It is mostly used to support analytical decision making typically based on models acquired from large quantities of data and therefore making it possible to acquire patterns and knowledge. By using data mining methods, patterns and rules can be acquired to be used as business rules in operational processes. The rules obtained with the data mining methods can be used in business processes instead of or to support analytical decision activities. Data mining models should in such cases be acquired and used on a daily basis.

Deployment of data mining is in fact deployment of an information technology into business processes. It is therefore recommended to use the same general principles as in deployment of new technologies into business processes. It is true that data mining is in business processes used only in the steps needing analytical decisions.

We believe that the classical data mining methodologies have to be extended for the cases where data mining is deployed in business processes. A proposal of such a methodological framework is presented in the paper.

2 THEORETICAL FOUNDATIONS

2.1 Data Mining

Data mining is a process of analyzing data in order to discover implicit, but potentially useful information and uncover previously unknown patterns and relationships hidden in the data (Witten & Frank, 2005). In the last decade, the digital revolution has provided relatively inexpensive and available means to collect and store the data. The increase in the data volume causes greater difficulties in extracting useful information for decision support. The traditional manual data analysis has become insufficient, and the methods for efficient computer-based analysis indispensable. From this need, a new interdisciplinary field of data mining has evolved. Data mining encompasses statistical, pattern recognition, and machine-learning tools to support analysis of the data and discovery of principles that lie within the data.

Received 9 July 2013
Accepted 2 September 2013

2.2 CRISP-DM Process Model

The data mining process model defines the approach to using data mining, i.e. phases, activities and tasks that have to be performed. Data mining represents a rather complex and specialized field. A generic and standardized approach is needed to using data mining in order to help organizations use the data mining.

CRISP-DM (CRoss-Industry Standard Process for Data Mining) is a non-proprietary, documented and freely available data mining process model. It has been developed by the industry leaders and in collaboration with experienced data mining users, by data mining software tool providers and by data mining service providers. CRISP-DM is an industry-, tool-, and application-neutral model created in 1996 [7]. A special Interest Group (CRISP-DM SIG) was formed in order to further develop and refine the CRISP-DM process model to service the data mining community well. The CRISP-DM version 1.0 was presented in 2000 and it is being accepted by business users [7].

Even though CRISP-DM recognizes that creation of the model is generally not the end of the project and introduces four tasks within the deployment phase, it lacks more detailed directions for the deployment of the data mining results into the business process, which requires implementation of a repeatable data mining process [6].

2.3 Business Processes

Throughout the last twenty years business process orientation has been gaining importance in business community. There are several definitions of business processes. A widely used is the one by Davenport and Short [5]. It defines the business process as a set of logically related tasks performed to achieve a defined business outcome. Generally, there are two groups of the generic business processes: operational and management processes. In addition to IT renovation, business process renovation requires consideration of organizational and managerial issues, such as cross-functional integration, stakeholder involvement, leadership qualities, and employee motivation. According to [4] the high failure rates of information systems projects are often the consequence of the managers' neglect of how users will react to new ways of working.

3 THE USING OF DATA MINING TO IMPROVE BUSINESS PROCESSES

The use of data mining in business processes is increasing, but has still not reached the level appropriate to the potential benefits of its use. The literature review reveals that it is being mainly used to support decision support. There are only few examples introducing the daily use of data mining in business processes [3;6]. We can say that data mining is predominantly used to support decision making on a tactical level in the decision and business processes. What are the prospects of usinmg data mining to support business processes on operational level? We believe that using data mining in business processes can be beneficial in cases where decisions can be operationalized based on stable models representing rules, i.e. data mining models.

3.1 Related Work

Chung and Gray [6] argue that there is a lot of research done in the areas of data mining model creation, but there is a lack of research done in the area of the use of data mining models in operational business processes.

Kohavi and Provost [3] argue that it is important to enable using of data mining in business processes through automated solutions. They discuss the importance of the ease of integration of data mining in business processes. In their paper they discuss the integration of data mining in business processes as the consequence of the need to incorporate background knowledge in business processes. They state that deploying automated solutions to previously manual processes can be a rife with pitfalls and also that social issues should be duly considered when deploying automated solutions to previously manual processes.

Feelders et al. [2] discuss using data mining in business processes with the emphasis on the integration of the data mining models and solutions into the existing application systems within information systems. Authors argue that it is essential that the results of data mining are used to support operational business processes like direct mailing for the selection of potential customers.

Gray [1] considers data mining as an option of knowledge sharing within the enterprise. By integrating data mining in operational business processes, knowledge sharing is not only present in business processes on a tactical level, but also on operational level.

3.2 A Methodology to Implement Data Mining into Business Processes

Based on the approaches to BI implementation (e.g. [9], [10]), on the methodological framework to business process renovation and IS development [11], and according to our experience acquired in implementing data mining in an analytical business process, we propose a methodology to implement data mining into business processes. The methodology consists of three phases where each of them has several activities as described below.

3.2.1 Phase One: Exploratory Data Mining

This purpose of this phase is to evaluate the readiness of operational business process and people involved in it for the implementation of data mining in their operational business process. The business value of using data mining in the process and risks and opportunities are evaluated too (Figure 1).

The process model of Phase one includes the CRISP-DM process model activities and the evaluation activity, which justifies and confirms (or declines) the implementation of data mining into the operational business process. Through CRISP-DM, the problem domain is first explored and transformed in the problem definition suitable for data mining implementation through business understanding activity. After that the data needed for modeling is defined and prepared through data understanding and data preparation activities. Data mining models are then created, evaluated and deployed through the following activities: modeling, evaluation and deployment.

It is important to note that the aim of the methodology is to implement data mining for the daily use, not for an ad-hoc use. As a result data preparation activity tends to automate data preparation to the highest possible stage and implement it as daily procedure which is executed automatically during the night or on demand at any time.

The final activity is Evaluation of DM readiness, business value, risks and opportunities. This is the main activity of Phase one which carries out the mission of the first phase: it evaluates the readiness of the operational business process and people involved in it for the implementation of data mining in their operational business process. Beside that it also defines the business value of the use of data mining in the process, evaluates risks and opportunities. If the evaluation gives positive results, then the Phase two is initiated.

3.2.2 Phase Two: Deployment of Data Mining into the Operational Business Process

This phase is in responsible for deployment of data mining in operational business process (Figure 1). For successful implementation of data mining there are two key activities which must be executed more or less simultaneously: process change design and applications analysis and design. The aim of the former is to define and design the changes that will take place in operational business process in order that the use of data mining brings added value. The aim of the latter is to make analysis and design of the application that will use data mining and must be developed. There is not only new application that must be developed; there are also existing applications that must be adapted to the use of data mining models and changes in operational business process. Both activities are very dependent on each other. For example, the application must provide the functionalities suitable for changes designed in activity process change design. After the activity applications analysis and design is finished the development of application is initiated through activity applications development and integration. The aim of this activity is not only to develop application, but also to integrate it into information systems of the company.

When the activity process change design is finished the activity process renovation is initiated. The aim of this activity is to renovate the operational business process according to the changes defined and designed in previous activity.

3.2.3 Phase Three: Operational Data Mining

Phase three supports the daily use of data mining in operational business process. The first step of the phase is the re-execution of activity data understanding. This activity was already executed in the first phase. But, since between first and the third phase there can be quite a time difference, the activity is re-executed in order to adapt to the possible changes of the databases. After this activity the daily use of data mining through one or more applications is enabled.

Activity data preparation is responsible for daily creation of data sets that represents input for modeling. Data sets are either created automatically during night or by demand by business users. Modeling and evaluation are executed by the use of application developed in the second phase. The core activity for daily use of data mining in the operational business process is deployment of data mining models in operational applications. In this activity data mining models are used in various applications by business users. According to our experience in using data mining models the following scenarios are possible:

1. Business users do not notice any disadvantages in data mining models or effects of their use. In this case activity deployment of data mining models in operational applications is re-executed daily.
2. Business users notice some disadvantages in data mining models or disadvantages in effects of their use. If they know that there were no such changes done in the databases that represent sources for data preparation, then the activity models re-creation is invoked. This way the data mining models are re-created and they reflect the last changes in the contents of source databases.
3. Business users notice some disadvantages in data mining models or disadvantages in effects of their use. If they know that there were recent changes in the databases that represent sources for data preparation, then the activity data understanding is invoked. The aim of the re-invoking is to detect the effect of recent changes in the source databases and to find out if there are any changes necessary in the area of data preparation. After the activity data preparation is finished activities of modeling and deployment can be executed, i.e. the daily use of data mining models can go on.
4. Business users notice some disadvantages in data mining models or disadvantages in effects of their use. It can also happen that they know that there are changes happening in the company that affect the domain of operational business process and its mission. In this case the activity Evaluation of DM readiness, business value, risks and opportunities is re-executed in order to evaluate (and re-define) the use of data mining in the operational business process.

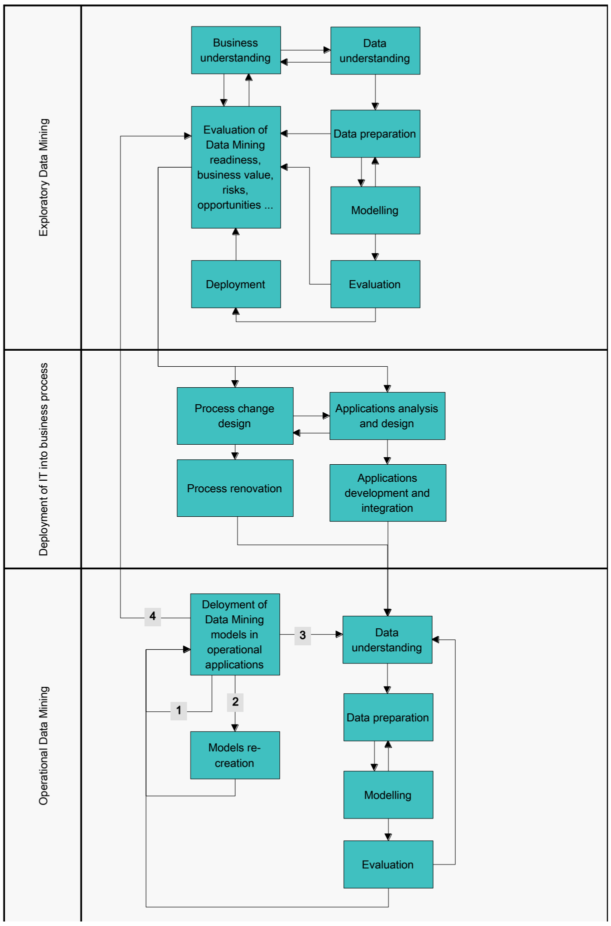

Figure 1. The methodology of implementation of data mining into operational business processes

4 USE OF THE PROPOSED METHODOLOGY: A CASE OF DIRECT MARKETING IN A PUBLISHING COMPANY

Slovenian publishing company exposed several problems with their direct marketing processes. Analysis of the most significant revealed that they could be solved by the use of data mining.

Sales marketing processes in the company are currently supported by existing application, which is in use for more than 20 years. There are several problems with it as not all the direct marketing activities are adequately supported. The main drawback is that each data processing for selecting prospective buyers from the database requires several hours. Therefore queries are run during nights and as a consequence running times for lists of potential customers are very long.

In the current (As-Is) direct marketing process a product manager first defines a target group for the book that is to be marketed. There are several possible ways or criteria to define the group:

- Through demographic data (gender, age, city) and other characteristics of customer that have bought similar titles in the past are analyzed. Demographs and geo-graphs are used for this type of analysis.
- Based on the characteristics of “similar books”. Product manager creates the list of those books according to his experience, i.e. his business knowledge. The existing application enables several ways for selecting “similar titles”, e.g. using the book classification.
- By the use of the RFM (Recency, Frequency, Monetary) method. Each customer that has bought at least one title in the last 5 years has a 3 digit RFM code.

The product manager then fills a paper form with the selected characteristic of the target group. The selection criteria are defined using a combination of data that is derived from the entire history of sales and from the knowledge, intuition and experience of product managers. It is obvious that data analysis and business knowledge are very important for direct marketing process, what makes this process rather knowledge intensive.

At this point of the process the IT staff is drawn in the process. The product manager sends them the request for preparing the prospective buyers list that includes the above mentioned paper form with the selection criteria. In most cases the attributes that are used for defining the criteria are standard and predefined. In these cases the IT staff enters the criteria in the existing application and the list is created over the night. There are some cases when the product manager wants to use complex criteria or additional attributes, besides the standard ones. In such situations a preliminary data processing is done in the earlier phase, at the time of criteria definition. This even prolongs the time required for preparing the prospective buyers list.

When the final list is prepared in the required form, it is forwarded to the printing office by the product manager for printing and personalization of the marketing materials. Responses of the customers are recorded in the database, which enables analysis of the response and the creation of data mining models to improve future campaigns.

4.1 *TO-BE Process Based on the Use of Data Mining*

The last paragraph of the previous section directly indicates the potentials of the use of data mining in direct marketing process. The use of Data mining will enable new ways for customer segmentation and discovering customer groups for marketing campaigns. The standard attribute set will not be the only way for segmentation any more, the segmentation on various attributes will be enabled. Publisher is aware that deployment of data mining would require a renovation of the direct marketing process, including changes in the employee roles and responsibilities. After all, the latter is the reason for consideration to deploy new information technologies. As with other information technologies, the data mining is an enabler for business process renovation on one side, however it also requires process changes on the other side.

Step-by-step changes of the direct marketing process are planned; first the analytical application will be deployed and later on the data mining application will be launched. Accordingly, the transition is expected to be smoother, particularly the changes related to the changing of roles. In the first phase the product managers will autonomously query and analyze the database. They will learn more about the database content and in this way they will get ready for the data mining, where the knowledge and understanding of the database content is one of the key success factors. As it can be learned from the direct marketing experiences, around 50 % of the success is in good understanding of the data. Only in the second phase, the data mining models for generating the prospective buyers list will be used inside the analytical application. An integration of both systems is planned.

The TO-BE process shows the direct marketing process flow after the analytical application and data mining have been deployed. Significant differences can be noticed when the AS-IS and TO-BE processes are compared:

- A significant shift in the workload of the IT staff toward the business users can be noticed. The number of activities performed by the IT staff is minimized, and moreover they are not involved in each marketing campaign. They have to acquire the appropriate level of new knowledge in the data mining, as they are mostly involved in building new models and re-creating of models. They are also involved in the designing of problem definitions together with business users, integration of models into the direct marketing application and into the analytical application, etc. Thus, the IT staff will have the role of a Data Mining model constructor.
- It cannot be expected that the product managers will be able to do all the additional analyses by themselves, therefore a new role appears on the business users' side: business analyst. Despite this is a business user, an advanced level of information technology knowledge is desired for such a person.
- Analytical activities are much more emphasized and the number of activities with the low added value is decreased. Moreover, the data mining supports the analytical activities.
- An increased number of possible iterations (see the closed loops in the process model) as the consequence of the increased interactivity can be noticed. This enables greater flexibility during the preparation of the prospective buyers list.

5 SUMMARY

According to the methodology presented (Figure 1) we have already finished the first phase and the part of the second phase which covers process renovation. The data mining models acquired in the first phase and test instance of direct mailing process gave positive results.

It has become clear that business process renovating and deploying automated solutions is not an easy task. Business users in general agree with the need to deploy data mining into direct marketing process. But, there is a little resistance felt and we know that the third phase will be not an easy task at all.

REFERENCES

[1] P. Gray, "New thinking about the enterprise," Information systems management, vol. 11, iss. 5, p. 91-95, 2005.

[2] A. Feelders, H. Daniels, M. Holsheimer, "Methodological and practical aspects of data mining," Information&Management, vol. 37, iss. 5, p. 271-281, 2000.

[3] R. Kohavi, F. Provost, "Applications of data mining in electronic commerce," Data mining and knowledge discovery, vol. 5, iss. 5, 2001.

[4] D. Chaffey, S. Wood, "Bisiness information management: Improving performacne using information systems," Prentice Hall, 2005.

[5] T.H. Davenport, J.E. Short, "The new industrial engineering: information technology and business process redesign," Sloan management review, vol. 31, iss. 4, p. 11-27, 1990.

[6] H.M. Chung, P. Gray, "SPecial edition: Data mining," Journal of management information systems, vol. 16, iss. 1, p. 11-16, 1999.

[7] C. Shearer, "The CRISP-DM model: the new blueprint for data mining," Journal for Data Warehousing, vol. 5, iss. 4, p. 13-22, 2000.

[8] I.H. Witten, E. Frank, "Data Mining: Practical Machine Learning Tools and Techniques", Morgan-Kaufmann, New York, 2005.

[9] L.T. Moss, S. Atre, "Business Intelligence Roadmap: The complete project lifecycle for decision-support applications", Addison-Wesley, 2003.

[10] S. Williams, N. Williams, "The profit impact of business inteligence", Morgan-Kaufman, 2007.

[11] A. Kovačič, V. Bosilj-Vukšič, "Business process management", GV založba, 2005.

Rok Rupnik je zaposlen kot docent na Fakulteti za računalništvo in informatiko. V svoji karieri je vodij več projektov na področju informatike. Od leta 2009 je imetnik certifikata PMP (Project Management Professional), ki ga podeljuje PMI (Project Management Institute). Njegova raziskovalna področja so: projektno vodenje, mobilne aplikacije, strateško planiranje informatike, obvladovanje informatike in podpora odločanju.