Data Mining Techniques in CRM

Inside Customer Segmentation

Konstantinos Tsiptsis

CRM & Customer Intelligence Expert, Athens, Greece

Antonios Chorianopoulos

Data Mining Expert, Athens, Greece

WILEY

A John Wiley and Sons, Ltd., Publication

Data Mining Techniques in CRM

Data Mining Techniques in CRM

Inside Customer Segmentation

Konstantinos Tsiptsis

CRM & Customer Intelligence Expert, Athens, Greece

Antonios Chorianopoulos

Data Mining Expert, Athens, Greece

WILEY

A John Wiley and Sons, Ltd., Publication

This edition first published 2009

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

Library of Congress Cataloguing-in-Publication Data

Record on file

A catalogue record for this book is available from the British Library.

ISBN: 978-0-470-74397-3

Typeset in 11/13.5pt NewCaledonia by Laserwords Private Limited, Chennai, India.
Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire.

To my daughter Eugenia and my wife Virginia, for their support and understanding.
And to my parents.
– Antonios

In memory of my father.
Dedicated to my daughters Marcella and Christina, my wife Maria, my sister Marina and my niece Julia and of course, to my mother Maria who taught me to set my goals in life.
– Konstantinos

CONTENTS

ACKNOWLEDGEMENTS

We would like to thank Vlassis Papapanagis, Leonidas Georgiou and Ioanna Koutrouvis of SPSS, Greece. Also, Andreas Kokkinos, George Krassadakis, Kyriakos Kokkalas and Loukas Maragos. Special thanks to Ioannis Mataragas for the creation of the line drawings.

CHAPTER ONE

Data Mining in CRM

THE CRM STRATEGY

Customers are the most important asset of an organization. There cannot be any business prospects without satisfied customers who remain loyal and develop their relationship with the organization. That is why an organization should plan and employ a clear strategy for treating customers. CRM (Customer Relationship Management) is the strategy for building, managing, and strengthening loyal and long-lasting customer relationships. CRM should be a customer-centric approach based on customer insight. Its scope should be the "personalized" handling of customers as distinct entities through the identification and understanding of their differentiated needs, preferences, and behaviors.

In order to make the CRM objectives and benefits clearer, let us consider the following real-life example of two clothing stores with different selling approaches. Employees of the first store try to sell everything to everyone. In the second store, employees try to identify each customer's needs and wants and make appropriate suggestions. Which store will finally look more reliable in the eyes of customers? Certainly the second one seems more trustworthy for a long-term relationship, since it aims for customer satisfaction by taking into account the specific customer needs.

CRM has two main objectives:

1. Customer retention through customer satisfaction.
2. Customer development through customer insight.

The importance of the first objective is obvious. Customer acquisition is tough, especially in mature markets. It is always difficult to replace existing customers

Data Mining Techniques in CRM: Inside Customer Segmentation K. Tsiptsis and A. Chorianopoulos

with new ones from the competition. With respect to the second CRM goal of customer development, the key message is that there is no average customer. The customer base comprises different persons, with different needs, behaviors, and potentials that should be handled accordingly.

Several CRM software packages are available and used to track and efficiently organize inbound and outbound interactions with customers, including the management of marketing campaigns and call centers. These systems, referred to as operational CRM systems, typically support front-line processes in sales, marketing, and customer service, automating communications and interactions with the customers. They record contact history and store valuable customer information. They also ensure that a consistent picture of the customer's relationship with the organization is available at all customer "touch" (interaction) points.

However, these systems are just tools that should be used to support the strategy of effectively managing customers. To succeed with CRM and address the aforementioned objectives, organizations need to gain insight into customers, their needs, and wants through data analysis. This is where analytical CRM comes in. Analytical CRM is about analyzing customer information to better address the CRM objectives and deliver the right message to the right customer. It involves the use of data mining models in order to assess the value of the customers, understand, and predict their behavior. It is about analyzing data patterns to extract knowledge for optimizing the customer relationships.

For example, data mining can help in customer retention as it enables the timely identification of valuable customers with increased likelihood to leave, allowing time for targeted retention campaigns. It can support customer development by matching products with customers and better targeting of product promotion campaigns. It can also help to reveal distinct customer segments, facilitating the development of customized new products and product offerings which better address the specific preferences and priorities of the customers.

The results of the analytical CRM procedures should be loaded and integrated into the operational CRM front-line systems so that all customer interactions can be more effectively handled on a more informed and "personalized" base. This book is about analytical CRM. Its scope is to present the application of data mining techniques in the CRM framework and it especially focuses on the topic of customer segmentation.

WHAT CAN DATA MINING DO?

Data mining aims to extract knowledge and insight through the analysis of large amounts of data using sophisticated modeling techniques. It converts data into knowledge and actionable information.

The data to be analyzed may reside in well-organized data marts and data warehouses or may be extracted from various unstructured data sources. A data mining procedure has many stages. It typically involves extensive data management before the application of a statistical or machine learning algorithm and the development of an appropriate model. Specialized software packages have been developed (data mining tools), which can support the whole data mining procedure.

Data mining models consist of a set of rules, equations, or complex "transfer functions" that can be used to identify useful data patterns, understand, and predict behaviors. They can be grouped into two main classes according to their goal, as follows.

SUPERVISED/PREDICTIVE MODELS

In supervised, or predictive, directed, or targeted modeling, the goal is to predict an event or estimate the values of a continuous numeric attribute. In these models there are input fields or attributes and an output or target field. Input fields are also called predictors because they are used by the model to identify a prediction function for the output field. We can think of predictors as the *X* part of the function and the target field as the *Y* part, the outcome.

The model uses the input fields which are analyzed with respect to their effect on the target field. Pattern recognition is "supervised" by the target field. Relationships are established between input and output fields. An input–output mapping "function" is generated by the model, which associates predictors with the output and permits the prediction of the output values, given the values of the input fields.

Predictive models are further categorized into classification and estimation models:

- **Classification or propensity models:** In these models the target groups or classes are known from the start. The goal is to classify the cases into these predefined groups; in other words, to predict an event. The generated model can be used as a scoring engine for assigning new cases to the predefined classes. It also estimates a propensity score for each case. The propensity score denotes the likelihood of occurrence of the target group or event.
- **Estimation models:** These models are similar to classification models but with one major difference. They are used to predict the value of a continuous field based on the observed values of the input attributes.

UNSUPERVISED MODELS

In unsupervised or undirected models there is no output field, just inputs. The pattern recognition is undirected; it is not guided by a specific target attribute.

The goal of such models is to uncover data patterns in the set of input fields. Unsupervised models include:

- **Cluster models:** In these models the groups are not known in advance. Instead we want the algorithms to analyze the input data patterns and identify the natural groupings of records or cases. When new cases are scored by the generated cluster model they are assigned to one of the revealed clusters.
- **Association and sequence models:** These models also belong to the class of unsupervised modeling. They do not involve direct prediction of a single field. In fact, all the fields involved have a double role, since they act as inputs and outputs at the same time. Association models detect associations between discrete events, products, or attributes. Sequence models detect associations over time.

DATA MINING IN THE CRM FRAMEWORK

Data mining can provide customer insight, which is vital for establishing an effective CRM strategy. It can lead to personalized interactions with customers and hence increased satisfaction and profitable customer relationships through data analysis. It can support an 'individualized' and optimized customer management throughout all the phases of the customer lifecycle, from the acquisition and establishment of a strong relationship to the prevention of attrition and the winning back of lost customers. Marketers strive to get a greater market share and a greater share of their customers. In plain words, they are responsible for getting, developing, and keeping the customers. Data mining models can help in all these tasks, as shown in Figure 1.1.

More specifically, the marketing activities that can be supported with the use of data mining include the following topics.

Customer Segmentation

Segmentation is the process of dividing the customer base into distinct and internally homogeneous groups in order to develop differentiated marketing strategies according to their characteristics. There are many different segmentation types based on the specific criteria or attributes used for segmentation.

In behavioral segmentation, customers are grouped by behavioral and usage characteristics. Although behavioral segments can be created with business rules, this approach has inherent disadvantages. It can efficiently handle only a few segmentation fields and its objectivity is questionable as it is based on the personal perceptions of a business expert. Data mining on the other hand can create

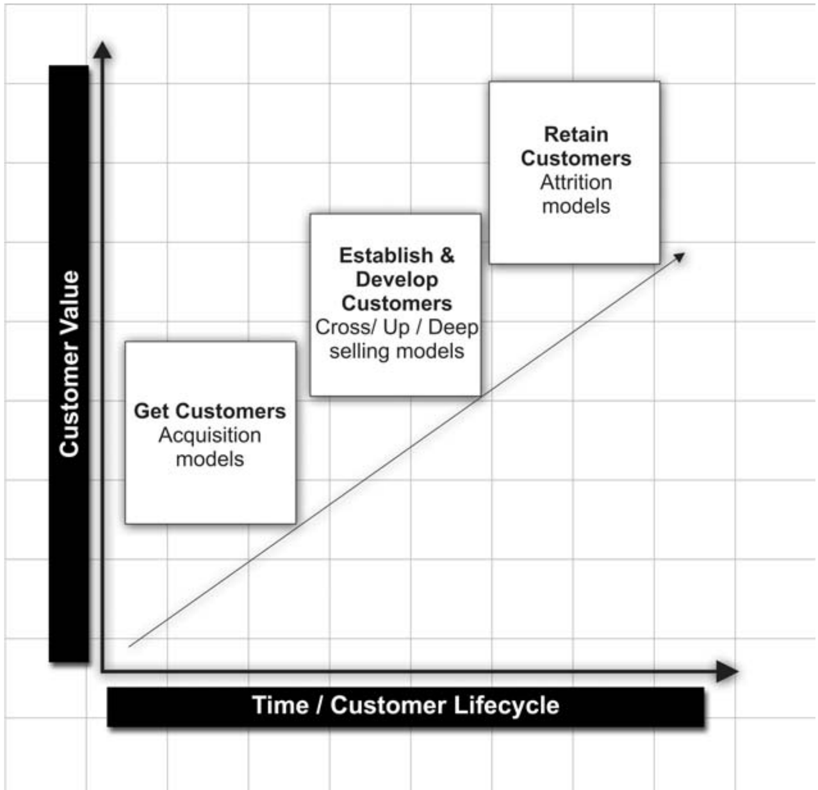

Figure 1.1 Data mining and customer lifecycle management.

data-driven behavioral segments. Clustering algorithms can analyze behavioral data, identify the natural groupings of customers, and suggest a solution founded on observed data patterns. Provided the data mining models are properly built, they can uncover groups with distinct profiles and characteristics and lead to rich segmentation schemes with business meaning and value.

Data mining can also be used for the development of segmentation schemes based on the current or expected/estimated value of the customers. These segments are necessary in order to prioritize customer handling and marketing interventions according to the importance of each customer.

Direct Marketing Campaigns

Marketers use direct marketing campaigns to communicate a message to their customers through mail, the Internet, e-mail, telemarketing (phone), and other direct channels in order to prevent churn (attrition) and to drive customer acquisition and purchase of add-on products. More specifically, acquisition campaigns aim at drawing new and potentially valuable customers away from the competition. Cross-/deep-/up-selling campaigns are implemented to sell additional products, more of the same product, or alternative but more profitable products to existing customers. Finally, retention campaigns aim at preventing valuable customers from terminating their relationship with the organization.

When not refined, these campaigns, although potentially effective, can also lead to a huge waste of resources and to bombarding and annoying customers with unsolicited communications. Data mining and classification (propensity) models in particular can support the development of targeted marketing campaigns. They analyze customer characteristics and recognize the profiles of the target customers. New cases with similar profiles are then identified, assigned a high propensity score, and included in the target lists. The following classification models are used to optimize the subsequent marketing campaigns:

- **Acquisition models:** These can be used to recognize potentially profitable prospective customers by finding "clones" of valuable existing customers in external lists of contacts,
- **Cross-/deep-/up-selling models:** These can reveal the purchasing potential of existing customers.
- **Voluntary attrition or voluntary churn models:** These identify early churn signals and spot those customers with an increased likelihood to leave voluntarily.

When properly built, these models can identify the right customers to contact and lead to campaign lists with increased density/frequency of target customers. They outperform random selections as well as predictions based on business rules and personal intuition. In predictive modeling, the measure that compares the predictive ability of a model to randomness is called the lift. It denotes how much better a classification data mining model performs in comparison to a random selection. The "lift" concept is illustrated in Figure 1.2 which compares the results of a data mining churn model to random selection.

In this hypothetical example, a randomly selected sample contains 10% of actual "churners." On the other hand, a list of the same size generated by a data mining model is far more effective since it contains about 60% of actual churners. Thus, data mining achieved six times better predictive ability than randomness. Although completely hypothetical, these results are not far from reality. Lift values higher than 4, 5, or even 6 are quite common in those real-world situations that

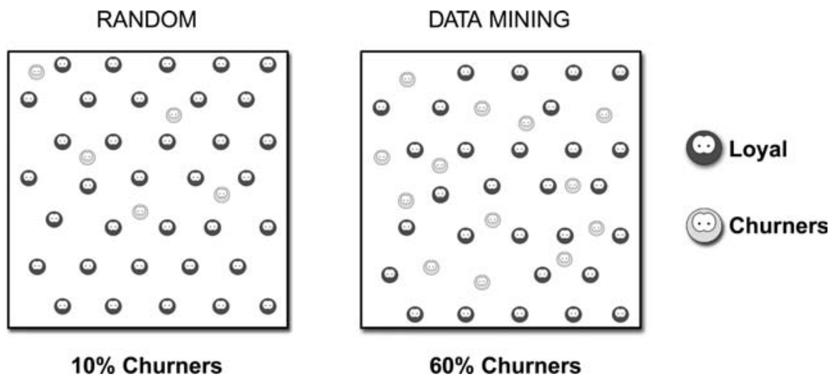

Figure 1.2 The increase in predictive ability resulting from the use of a data mining churn model.

were appropriately tackled by well-designed propensity models, indicating the potential for improvement offered by data mining.

The stages of direct marketing campaigns are illustrated in Figure 1.3 and explained below:

1. Gathering and integrating the necessary data from different data sources.
2. Customer analysis and segmentation into distinct customer groups.
3. Development of targeted marketing campaigns by using propensity models in order to select the right customers.
4. Campaign execution by choosing the appropriate channel, the appropriate time, and the appropriate offer for each campaign.
5. Campaign evaluation through the use of test and control groups. The evaluation involves the partition of the population into test and control groups and comparison of the positive responses.
6. Analysis of campaign results in order to improve the campaign for the next round in terms of targeting, time, offer, product, communication, and so on.

Data mining can play a significant role in all these stages, particularly in identifying the right customers to be contacted.

Market Basket and Sequence Analysis

Data mining and association models in particular can be used to identify related products typically purchased together. These models can be used for market basket analysis and for revealing bundles of products or services that can be sold together.

Figure 1.3 The stages of direct marketing campaigns.

Sequence models take into account the order of actions/purchases and can identify sequences of events.

THE NEXT BEST ACTIVITY STRATEGY AND "INDIVIDUALIZED" CUSTOMER MANAGEMENT

The data mining models should be put together and used in the everyday business operations of an organization to achieve more effective customer management. The knowledge extracted by data mining can contribute to the design of a next best activity (NBA) strategy. More specifically, the customer insight gained by data mining can enable the setting of "personalized" marketing objectives. The organization can decide on a more informed base the next best marketing activity for each customer and select an "individualized" approach which might be the following:

- An offer for preventing attrition, mainly for high-value, at-risk customers.
- A promotion for the right add-on product and a targeted cross-/up-/deep-selling offer for customers with growth potential.

- Imposing usage limitations and restrictions on customers with bad payment records and bad credit risk scores.
- The development of a new product/offering tailored to the specific characteristics of an identified segment, and so on.

The main components that should be taken into account in the design of the NBA strategy are illustrated in Figure 1.4. They are:

1. The current and expected/estimated customer profitability and value.
2. The type of customer, the differentiating behavioral and demographic characteristics, the identified needs and attitudes revealed through data analysis and segmentation.
3. The growth potential as designated by relevant cross-/up-/deep-selling models and propensities.

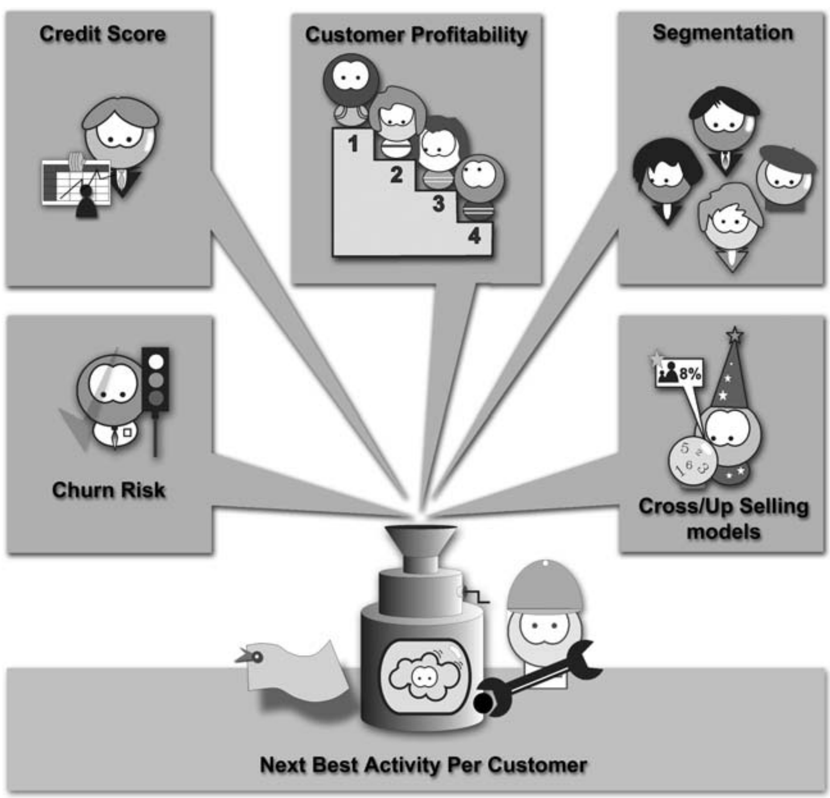

Figure 1.4 The next best activity components.

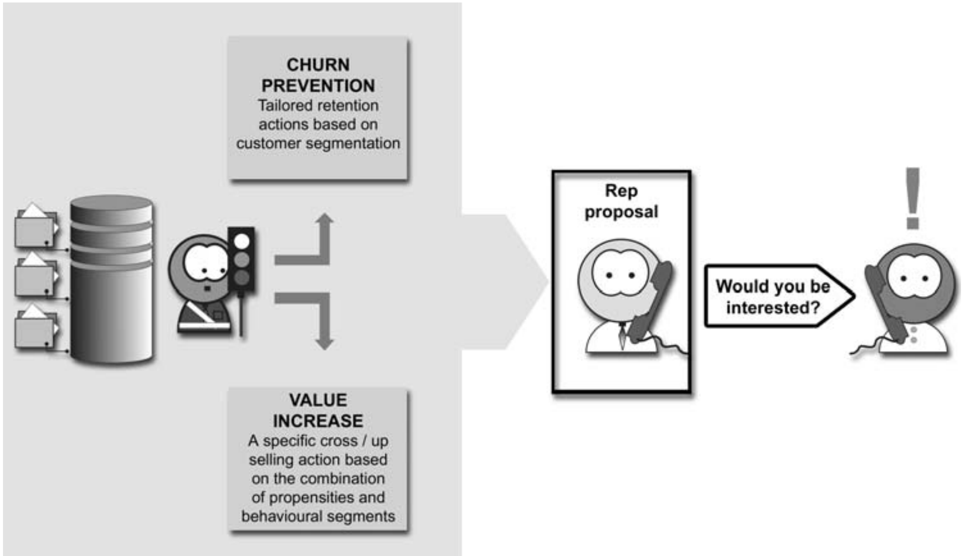

Figure 1.5 The next best activity strategy in action.

4. The defection risk/churn propensity as estimated by a voluntary churn model.
5. The payment behavior and credit score of the customer.

In order to better understand the role of these components and see the NBA strategy in action (Figure 1.5), let us consider the following simple example. A high-value banking customer has a high potential of getting a mortgage loan but at the same time is also scored with a high probability to churn. What is the best approach for this customer and how should he be handled by the organization? As a high-value, at-risk customer, the top priority is to prevent his leaving and lure him with an offer that matches his particular profile. Therefore, instead of receiving a cross-selling offer, he should be included in a retention campaign and contacted with an offer tailored to the specific characteristics of the segment to which he belongs.

THE DATA MINING METHODOLOGY

A data mining project involves more than modeling. The modeling phase is just one phase in the implementation process of a data mining project. Steps of critical importance precede and follow model building and have a significant effect on the success of the project.

Table 1.1 The CRISP-DM phases.

1. Business understanding	**2. Data understanding**	**3. Data preparation**
• Understanding the business goal • Situation assessment • Translating the business goal into a data mining objective • Development of a project plan	• Considering data requirements • Initial data collection, exploration, and quality assessment	• Selection of required data • Data acquisition • Data integration and formatting (merge/joins, aggregations) • Data cleaning • Data transformations and enrichment (regrouping/binning of existing fields, creation of derived attributes and key performance indicators: ratios, flag fields, averages, sums, etc.)
4. Modeling	**5. Model evaluation**	**6. Deployment**
• Selection of the appropriate modeling technique • Especially in the case of predictive models, splitting of the dataset into training and testing subsets for evaluation purposes • Development and examination of alternative modeling algorithms and parameter settings • Fine tuning of the model settings according to an initial assessment of the model's performance	• Evaluation of the model in the context of the business success criteria • Model approval	• Create a report of findings • Planning and development of the deployment procedure • Deployment of the data mining model • Distribution of the model results and integration in the organization's operational CRM system • Development of a maintenance–update plan • Review of the project • Planning the next steps

An outline of the basic phases in the development of a data mining project according to the CRISP-DM (Cross Industry Standard Process for Data Mining) process model is presented in Table 1.1.

Data mining projects are not simple. They usually start with high expectations but may end in business failure if the engaged team is not guided by a clear methodological framework. The CRISP-DM process model charts the steps that should be followed for successful data mining implementations. These steps are as follows:

1. **Business understanding:** The data mining project should start with an understanding of the business objective and an assessment of the current situation. The project's parameters should be considered, including resources and

limitations. The business objective should be translated into a data mining goal. Success criteria should be defined and a project plan should be developed.

2. **Data understanding:** This phase involves considering the data requirements for properly addressing the defined goal and an investigation of the availability of the required data. This phase also includes initial data collection and exploration with summary statistics and visualization tools to understand the data and identify potential problems in availability and quality.
3. **Data preparation:** The data to be used should be identified, selected, and prepared for inclusion in the data mining model. This phase involves the acquisition, integration, and formatting of the data according to the needs of the project. The consolidated data should then be "cleaned" and properly transformed according to the requirements of the algorithm to be applied. New fields such as sums, averages, ratios, flags, and so on should be derived from the raw fields to enrich customer information, to better summarize customer characteristics, and therefore to enhance the performance of the models.
4. **Modeling:** The processed data are then used for model training. Analysts should select the appropriate modeling technique for the particular business objective. Before the training of the models and especially in the case of predictive modeling, the modeling dataset should be partitioned so that the model's performance is evaluated on a separate dataset. This phase involves the examination of alternative modeling algorithms and parameter settings and a comparison of their fit and performance in order to find the one that yields the best results. Based on an initial evaluation of the model results, the model settings can be revised and fine tuned.
5. **Evaluation:** The generated models are then formally evaluated not only in terms of technical measures but also, more importantly, in the context of the business success criteria set out in the business understanding phase. The project team should decide whether the results of a given model properly address the initial business objectives. If so, this model is approved and prepared for deployment.
6. **Deployment:** The project's findings and conclusions are summarized in a report, but this is hardly the end of the project. Even the best model will turn out to be a business failure if its results are not deployed and integrated into the organization's everyday marketing operations. A procedure should be designed and developed to enable the scoring of customers and the updating of the results. The deployment procedure should also enable the distribution of the model results throughout the enterprise and their incorporation in the organization's databases and operational CRM system. Finally, a maintenance plan should be designed and the whole process should be reviewed. Lessons learned should be taken into account and the next steps should be planned.

The phases above present strong dependencies and the outcomes of a phase may lead to revisiting and reviewing the results of preceding phases. The nature of the process is cyclical since the data mining itself is a never-ending journey and quest, demanding continuous reassessment and updating of completed tasks in the context of a rapidly changing business environment.

DATA MINING AND BUSINESS DOMAIN EXPERTISE

The role of data mining models in marketing is quite new. Although rapidly expanding, data mining is still "foreign territory" for many marketers who trust only their "intuition" and domain experience. Their segmentation schemes and marketing campaign lists are created by business rules based on their business knowledge.

Data mining models are not "threatening": they cannot substitute or replace the significant role of domain experts and their business knowledge. These models, however powerful, cannot effectively work without the active support of business experts. On the contrary, only when data mining capabilities are complemented with business expertise can they achieve truly meaningful results. For instance, the predictive ability of a data mining model can be substantially increased by including informative inputs with predictive power suggested by persons with experience in the field. Additionally, the information of existing business rules/scores can be integrated into a data mining model and contribute to the building of a more robust and successful result. Moreover, before the actual deployment, model results should always be evaluated by business experts with respect to their meaning, in order to minimize the risk of coming up with trivial or unclear findings. Thus, business domain knowledge can truly help and enrich the data mining results.

On the other hand, data mining models can identify patterns that even the most experienced business people may have missed. They can help in fine tuning the existing business rules, and enrich, automate, and standardize judgmental ways of working which are based on personal perceptions and views. They comprise an objective, data-driven approach, minimizing subjective decisions and simplifying time-consuming processes.

In conclusion, the combination of business domain expertise with the power of data mining models can help organizations gain a competitive advantage in their efforts to optimize customer management.

SUMMARY

In this chapter we introduced data mining. We presented the main types of data mining models and a process model, a methodological framework for designing and implementing successful data mining projects. We also outlined how data mining can help an organization to better address the CRM objectives and achieve "individualized" and more effective customer management through customer insight. The following list summarizes some of the most useful data mining applications in the CRM framework:

- Customer segmentation:
  - **Value-based segmentation:** Customer ranking and segmentation according to current and expected/estimated customer value.
  - **Behavioral segmentation:** Customer segmentation based on behavioral attributes.
  - **Value-at-risk segmentation:** Customer segmentation based on value and estimated voluntary churn propensity scores.
- Targeted marketing campaigns:
  - Voluntary churn modeling and estimation of the customer's likelihood/propensity to churn.
  - Estimation of the likelihood/propensity to take up an add-on product, to switch to a more profitable product, or to increase usage of an existing product.
  - Estimation of the lifetime value (LTV) of customers.

Table 1.2 presents some of the most widely used data mining modeling techniques together with an indicative listing of the marketing applications they can support.

Table 1.2 Data mining modeling techniques and their applications.

Category of modeling techniques	**Modeling techniques**	**Applications**
Classification (propensity) models	Neural networks, decision trees, logistic regression, etc.	• Voluntary churn prediction • Cross/up/deep selling
Clustering models	K-means, TwoStep, Kohonen network/self-organizing map, etc.	• Segmentation
Association and sequence models	A priori, Generalized Rule Induction, sequence	• Market basket analysis • Web path analysis